# MAKING GLOBAL INTEGRATED ENVIRONMENTAL ASSESSMENT AND REPORTING MATTER

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

LÁSZLÓ PINTÉR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

JAMES A. PERRY, ADVISER

MARCH, 2002

László Pintér
348 words

# ABSTRACT

Producing reliable knowledge and information on the trends, dynamics and possible directions of the global environment is a precondition of intelligent policy-making. Besides assessments dealing with particular issues and sectors, there is a well-established need for regularly pulling together and synthesizing knowledge in ways that reflect the interconnectedness of environment and development. The resulting integrated assessment and reporting (IAR) systems need to reflect an understanding of cross-scale interactions from the local to the global level, link past trends and emerging issues, and cover all key sectors of the environment. The assessment systems have to be approached as a reflexive, evolving, and largely experimental process. As an important element, we need to identify and understand linkages between assessment design and effectiveness.

This study looks at the connection between design elements and effectiveness in the Global Environment Outlook (GEO), currently the most comprehensive global reporting system on the environment from the perspective of sustainability, produced by the United Nations Environment Programme (UNEP) and a global network of collaborative centres. Taking a practitioner perspective, I was interested in how selected criteria of effectiveness, including saliency, credibility, legitimacy and awareness are influenced by key design elements of GEO, including: framing, governance, participation, capacity, communication, data and indicators, and feedback to research agenda setting.

My research found that while the GEO scores well on several of these criteria, significant opportunities for improvement remain. Saliency could be increased through strengthening early and substantive participation in the assessment process, and using the process to understand and directly respond to the questions of key policy audiences. These should include the identification and systematic use of a core set of indicators. There are important opportunities to strengthen credibility by providing detailed blueprints for

assessment methods, and that both these and final results are subject to rigorous peer review. Legitimacy could be enhanced by applying a governance model that provides organizations participating in the assessment significant responsibilities, but also contributes to their increased capacity. Finally, in order to increase awareness of environment and development problems and solutions, IAR systems need to establish and systematically implement effective communication strategies.

# ACKNOWLEDGMENTS

First and foremost, thanks are due to my immediate family members, my wife Emőke, and our daughters Kamilla and Margaréta, but also my parents and sister who gave up an uncounted number of days and weeks to spend quality time together so that I can do this work.

Many thanks to Professor James Perry, my advisor, friend and colleague who was and I hope will continue to be a source of wisdom, inspiration and encouragement. Thanks to my other committee members for guidance and comments, Dr. Hans Gregersen, Dr. Peter Hardi, Dr. Allen Lundgren and Dr. Anne Kapuscinski. I am grateful for IISD for supporting my Ph.D. studies and granting a one-year leave of absence that allowed me to accept a fellowship to join the Sustainability Science project at Harvard that brought many challenges as well as inspiration. Thanks are due to Jill Jäger (IHDP/Harvard), my direct superior and William Clark (Harvard) for conversations and comments, as well as for several GEA fellows, particularly Bernd Kasemir, Bernd Siebenhüner, David Cash and Franck Alcock. I am grateful for the thoughts and invaluable support received from Marion Cheatle (UNEP), Anna Stabrawa (UNEP), Dave MacDevette (UNEP), Nickolai Denisov (GRID-Arendal), Jan Bakkes (RIVM), Paul Raskin (Tellus Institute), and Marlene Roy (IISD). I am also indebted to my interviewees and many other colleagues and friends at GEO Collaborating Centres around the world, UNEP-DEWA and IISD who expressed support for and interest in my research from an early stage. Finally, thanks to my friends, Ed Deak, Csaba Káncz, Attila Király, and Péter Nagy for wide ranging correspondence and conversations over the years that helped put my work in perspective.

# DEDICATION

Dedicated to my family, those who were around to support this work, those who already passed away but whose wisdom and love shaped me as I grew up, and those yet to be born. May all people understand the importance of taking care of the world as we take care of our loved ones.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# GLOSSARY

| | |
|---|---|
| AC | associated centre |
| AI | appreciative inquiry |
| AGU | Arabian Gulf University |
| ASEAN | Association of Southeast Asian Nations |
| BCAS | Bangladesh Centre for Advanced Studies |
| CBD | Convention on Biological Diversity |
| CC | Collaborating Centre |
| CDE | capacity development in environment |
| CEROI | Cities Environment Reports on the Internet |
| CEU | Central European University |
| DEWA | Division of Early Warning and Assessment |
| DPSIR | driving force-pressure-state-impact-response |
| EEA | European Environment Agency |
| EFIEA | European Forum on Integrated Environmental Assessment |
| ENRIN | Environment and Natural Resource Information Network |
| EU | European Union |
| FAO | UN Food and Agriculture Organization |
| GCOS | Global Climate Observation System |
| GEA | Global Environmental Assessment |
| GEF | Global Environmental Facility |
| GEO | Global Environment Outlook |
| GIWA | Global Integrated Water Assessment |
| GOOS | Global Oceanic Observation System |
| GTOS | Global Terrestrial Observing System |
| HDI | Human Development Index |
| HDR | Human Development Report |
| IA | integrated assessment |
| IAR | integrated assessment and reporting |
| IDGEC | Institutional Dimensions of Global Environmental Change |

| | |
|---|---|
| IDRC | International Development Research Centre |
| IGBP | International Geosphere-Biosphere Programme |
| IGOS | Integrated Global Observing Strategy |
| IHDP | International Human Dimensions Programme |
| IIASA | International Institute for Applied Systems Analysis |
| IIIs | international information institutions |
| IISD | International Institute for Sustainable Development |
| IMERCSA | Musokotwane Environment Resource Centre for Southern Africa |
| IPCC | Intergovernmental Panel on Climate Change |
| IREP | integrated region-oriented environmental policy |
| IUCN | International Union for the Conservation of Nature |
| IUE | information use environment |
| LRTAP | Long-Range Transboundary Air Pollution |
| MEA | Millennium Ecosystem Assessment |
| MSU | Moscow State University |
| NASA | National Aeronautical and Space Agency |
| OECD | Organization for Economic Cooperation and Development |
| PSR | Pressure-State-Response |
| RIVM | Rijksinstitut voor Volksgezondheid en Milieu (National Institute of Public Health and the Environment, The Netherlands) |
| SCOPE | Scientific Committee on Problems of the Environment |
| SEI(-B) | Stockholm Environment Institute(-Boston) |
| SEPA | State Environment Protection Agency, People's Republic of China |
| SOE | state of the environment |
| UNCED | United Nations Conference on Environment and Development |
| UN-CSD | United Nations Commission for Sustainable Development |
| UNDP | United Nations Development Programme |
| UNEP | United Nations Environment Programme |
| UNEP-ROAP | United Nations Environment Programme – Asia Pacific |
| UNEP-DEWA | United Nations Environment Programme, Division for Early Warning and Assessment |
| UNEP-DTIE | United Nations Environment Programme, Division of Technology, Industry and Economics |

| | |
|---|---|
| UNEP-GPA | United Nations Environment Programme, Global Programme of Action for Marine Pollution from Land-Based Sources |
| UNEP-RRC.AP | United Nations Environment Programme, Regional Resource Centre for Asia and the Pacific |
| UNFIP | United Nations Foundation for International Partnerships |
| UNFCCC | United Nations Framework Convention on Climate Change |
| UNGASS | United Nations General Assembly |
| UNIHP | United Nations Intellectual History Project |
| UNITAR | United Nations Institute for Training and Research |
| USGCRP | United States Global Change Research Program |
| WCED | World Commission on Environment and Development |
| WCMC | World Conservation Monitoring Centre |
| WCRP | World Climate Research Programme |
| WMO | World Meteorological Organization |
| WRI | World Resources Institute |
| WWF | World Wildlife Fund |

*"[We do not need a computer model to tell us that]*

*we must not destroy the system upon which our sustenance depends.*
*poverty is wrong and preventable.*
*the exploitation of one person or nation by another degrades both the exploited and the exploiter.*
*it is better for individuals and nations to cooperate than to fight.*
*the love we have for all humankind and for future generations should be the same as our love for those close to us.*
*if we do not embrace these principles and live by them, our system cannot survive.*
*our future is in our hands and will be no better or worse than we make it."*

D. Meadows, J. Richardson and G. Bruckmann.

# 1. Introduction

Many theses start with descriptions of problems. This one starts with a vision. The vision is about global information systems on environment and development some 4-5 decades into the future. It is a best guess of what I think will happen, but it also has a normative overtone of what I think should happen. I am not starting with a vision to sidestep problems. I do so, because I find that in order to constructively deal with challenges it is often useful to first consider positive alternatives for contrast, guidance and in fact motivation[1].

Visions are normally associated with time frames and underlying assumptions. I put the time frame of this vision at approximately 2050. My basic assumption is that as the seriousness of environment and development problems grows, the demand for relevant information on diagnoses and solutions consistent with some form of sustainable development will also increase[2]. There is a need and opportunity to establish more effective information systems that communicate issues arising from the interaction of environment and development (e.g., Carnegie Commission 1992).

What are the key attributes of global assessment and reporting systems in this positive version of 2050? Generally, they will be better institutionalized and have a more stable

role in governance. This applies not only to environmental governance, but the governance of environment / development issues that will have become even more interwoven than they are today. While better institutionalization means - among other things - more certainty that assessments can take place on a regular basis, it does not mean structural and conceptual rigidity. Rather their underlying approach is learning and adaptive, which means assessment systems are reflexive and change their focus, methods, technologies, partners, communications and other strategies in response to changing issues and requirements.

Even when coordinated by a single agency or consortium, assessments are produced by globally distributed knowledge networks or virtual communities for assessment. These include member organizations from all regions of the world that play a dual role by interpreting global knowledge for local audiences and by contributing local perspectives to regional and global assessments. Having resolved basic information access problems, the developing countries of 2050 are participants and important players in information production. Information in 2050 continues to be produced by many types of actors, including scientific organizations, public sector reporting systems, corporations, and more than ever today, by a variety of locally established and controlled autonomous, non-governmental, community-based and multi-stakeholder organizations. Many of those who contribute to global assessments also participate in similar initiatives taking place on the regional, national or community scales. The distinction between producers and users of assessment information often is blurred, as policy audiences are routine participants or co-producers of assessments.

While information production is thus decentralized, information access is possible among other ways, through interlinked Internet portals. As the amount of available information and the need for specialized information continues to grow, there also is increased demand for services to filter relevant information. Some of this is done directly by audiences through the better use of artificial intelligence and software robots. However, there also is a major industry involved in mining, compiling, organizing, interpreting, repackaging and bringing environment and development information to the attention of

the public and private audiences. Along all of these steps information gains and loses content and meaning and therefore information use by audiences is recognized as playing as much role in determining the utilization and impact of information as primary production.

Institutional stability of assessment and reporting systems is a precondition to adequately satisfy demand for two types of information. The first is the need for baseline information on issues that have strong 'staying power' on public agendas. Second, there is demand for 'rapid response' assessments associated with emerging issues and acute or critical events that have become more common with the increased vulnerability of ecosystems and human systems. In response to unforeseen events, new questions and surprises, *ad hoc* task forces are formed, assembled from members of the assessment community involved in the production of baseline information products, in order to prepare rapid assessments on short notice. From being focused on single products, assessment systems have shifted towards producing information streams and 'information on demand'. There is less lag time between gathering and making primary data available due to a closer integration of decision-maker information needs, monitoring, data analysis and communication. Assessment systems are better suited to respond to specific information demands when and where those needs arise, rather than work only on the basis of their own schedules. Trademark products and reports that are published regularly by high-profile organizations and reported widely by the mainstream media are still around, but they all have parallel and continuously updated, Internet-based equivalents.

The information produced by thematically focused science assessments serves as input into more integrated systems, and vice versa, integrated analyses helping to frame narrow assessments. For instance, an assessment dealing with urban air pollution does not ignore cross-scale effects or linkages (e.g., to energy production, transportation, public health and underlying driving forces and policies). While there continue to be separate narrow and broad assessments, there is intensive interaction and information exchange between the two layers.

Data and knowledge gaps continue to exist with regard to many environment / development issues and regions of the world. However, there is better understanding – as well as acceptance – of the unavoidability and nature of these uncertainties when dealing with complex systems, a multitude of social actors, coarse spatial scales and long time horizons. Filling data and assessment related knowledge gaps are integrated into mitigation and adaptation strategies designed to reduce the risks associated with decisions based on incomplete understanding of the world. Similar to the closer association of assessment users and producers, there is more active cooperation between monitoring systems and assessment producers as key users of primary data. Monitoring, data collection and indicator development are guided by internationally accepted standards and protocols developed since the 1990s[3]. These extend to the explicit identification of targets and sustainability thresholds identified through scientific research and consensus-seeking participatory processes, often associated with policies of particular decision-makers. As the need for new and better data and monitoring activities continues to emerge, there is an effective institutional mechanism to periodically review and adjust global monitoring systems and their equivalents on finer spatial scales.

Assessments are inherently forward looking and involve linking past trends with future decision options with regard to high priority policy issues. Rather than arising from academic interest alone, this is a response to the need of decision-makers for science-based and legitimate knowledge and guidance that has direct utility in policy planning. Given the inherent uncertainties associated with future outcomes, alternative scenarios are presented as pathways composed of a sequence of policy decisions, explained through narratives and illustrated by indicators derived partly from integrated models. The participatory character of the assessment will extend to the conceptualization and analysis of policy options and scenarios. In fact, these have become an integral part of assessment and reporting architectures on all levels.

Information overload of audiences continues to be a problem, but assessment and reporting systems have improved their communication strategies (or rather they have communication strategies). Communication strategy planning is an integral part of

assessment and reporting, led by professionals. The strategy reflects the importance of partnerships, the needs and capacities of various audiences, views communication as an ongoing process, and builds on a global network composed of assessment producers and partnerships with public and private sector organizations, such as major news agencies and publishers.

## 1.1. Purpose and research questions

Many elements of the model described in the introductory vision are visible in today's existing assessment and reporting systems. One of the first comprehensive reviews by Hodge (1995) shows that by the mid-1990s many reports of the state of the environment (SOE) experimented with integration, sustainability and other concepts mentioned in this vision. This is in contrast with earlier SOE reports that focused on the description of environmental conditions and put less emphasis on identifying causal linkages and interactions across spatial and temporal scales. Rump (1996) compiled a "source book of methods and approaches" that both generalizes and pulls together various elements into what could be considered a loosely structured model of SOE reporting and illustrates general points with real-life examples. Hardi and Zdan (1997) do the same when they illustrate the general Bellagio Principles on assessment in the context of sustainable development. GRID-Arendal essentially created a template based on which they started supporting the establishment of reporting and assessment systems on the national and community scales through the Environment and Natural Resource Information Network (ENRIN) and Cities Environment Reports on the Internet (CEROI) projects, among others (UNEP/GRID-Arendal 1998, Denisov *et al.* 2000, Simonett *et al.* 1998). This was taken a step further through the GEO capacity development programme of UNEP that added an institutional and process dimension to the previous models that focused more on conceptual issues and assessment products (Pintér, Zahedi and Cressman 2000).

The existing systems described by Hodge (1995), Rump (1996) and others are still far from what I tried to describe in the opening vision-experiment. But even our understanding of what an ideal assessment system, a possible basis of comparison, would

look like is imperfect. What we do know is that decisions involving environment / development issues are also made under conditions of imperfect information and uncertainty (e.g., Funtowicz and Ravetz 1993). In the absence of adequate science or expert advice people make up the information deficit in different ways, often using non-scientific sources, and ultimately relying on best professional judgment. It is not a surprise that decisions are often found to be subjective, fragmented, imprecise and sub-optimal. I am not suggesting that it is realistic to expect there will be integrated assessment and reporting systems that will provide perfect information, but we can probably improve on their architecture so that they are more useful. Perhaps more importantly, we are still in the early stages of understanding what makes existing systems work or fail. This is the area where my work aims to provide original contribution.

The purpose of this study is to contribute to understanding the relationship between the design and effectiveness of global scale, but regionally differentiated integrated environmental assessment and reporting systems. There are many assessments that deal with specific places or environmental problems. In contrast, my interest in this research is in global assessment systems that report on the state and dynamics of the entire planetary environment and its interaction with human development. In order to achieve that function, assessments need to integrate thematic information related to a wide variety of environment and development issues, find ways to aggregate regional trends and perspectives to meaningful global averages, and highlight local implications of global dynamics on the local scale (e.g., Cash and Moser 2000). Crucially, they also need to maintain relevance for a variety of policy audiences. The Global Environment Outlook (GEO) of the United Nations Environment Programme (UNEP) stands as a prominent assessment system with such ambitions. Among UNEP publications, GEO has been one of the most visible and GEO-2000 created the most media inquiries of any UNEP publication (Universalia 2000). In contrast with some other global reporting systems, GEO is a joint product of UNEP and a global network of collaborating institutions. These institutions bring regional and thematic expertise and connections to local networks, but when working together often they also need to bridge cultural, methodological and capacity gaps. Undertaking the GEO program involves making strategic choices in terms

of how the assessment is done and configured to deal with complex environment/ development problems. Depending on the choices made, resulting outputs of the assessment may lead to different outcomes in the world of policy, and ultimately global sustainability.

Although I build on relevant theory, the core of this research is inductive and the intended tone is practical. This is partly a reflection of my professional background, but also a response to a perceived need for practice-oriented analysis in the broader assessment and reporting community. Many assessments and reporting systems, particularly of the integrated type have not more than a few years history, and they became subjects of research relatively recently. Results of this have started to appear in the ecological economics, management science, and political science literature (e.g., Costanza and Tognetti 1996, Jäger 1998, Rotmans 1998, Fritz 1998). Bringing the results of the growing body of theoretical and academic research to bear on the practice of assessment and reporting is a significant challenge, though with potentially high pay-offs. Some of the most ambitious work in this area has been carried out in the Global Environmental Assessment (GEA) project at Harvard University and the European Forum on Integrated Environmental Assessment (EFIEA) network in Europe (Jäger and Farrell, forthcoming; European Environment Agency 2000).

As environmental issues have increased in importance, there has been growing interest within the research community in the use and usefulness of expert advice, science, and environmental information in decision-making (e.g., Jasanoff 1990, Denisov and Christoffersen 2001). It is only more recently, however, that trans-boundary, integrated environmental assessment systems became subjects of research. As the view of environmental problems has changed to include broader interactions across issues, sectors, spatial and temporal scales, as well as cultures, assessment systems also adopted broader, integrative perspectives. Many science-based advisory processes and assessment systems that focus on particular themes, such as IPCC on climate change or GIWA on water now reflect this broader view. As challenging as it is, thematic integrated assessment still falls short of assessment where thematic, where a mosaic of fine scale

spatial and temporal details must fit together to provide a view of coarse scale trends and dynamics. Regionally differentiated but global level integrated analysis is probably one of society's most challenging social and scientific enterprises, one whose building requires an institutional, systematic approach (Carnegie Commission 1992). There is a clear need for progress. With the costs of global environmental change and its local impacts increasing, both the public in general and decision-makers in particular are asking questions not only about the condition of particular regions or issues, but also how partial trends add up to a whole.

It seems fitting that UNEP as the guardian of the global environment in the UN system undertook to establish an assessment and reporting system to address these questions on an ongoing basis. Although UNEP used to publish decadal SOE reports, these were assessments focused predominantly on the natural environment (e.g., UNEP 1982, UNEP 1991). This started to change with the publication of *The World Environment 1972-1992,* that was structured to cover in detail not only issues, but causes and consequences, responses, and included a chapter on future directions (Tolba and El-Kholy 1992). Although led by UNEP the 1992 report, a substantive volume, was already a product of a global team of experts. UNEP initiated the GEO program in order to address sustainability issues from the perspective of the environment, but also to learn together with other partners in the process how integrated assessment and reporting on this level can be done. GEO is a key test case in this work.

Learning about how GEO works is important because it might offer valuable empirical evidence about the challenges involved when building an assessment and reporting system with global ambitions, regional texture, future orientation, and with the clear objective to bridge the science-policy gap. Although GEO is certainly not the only one, in the environmental field there are few other operational reporting systems with a distributed organizational structure and a nested 'panarchy' of dynamic global, regional and sub-regional processes[4] (Gunderson, Holling and Light 1995). It is far from clear and certain how such constructs can deal with the complexity arising from the issues themselves, but also from the multiple cultures of participating institutions and the very

different capacities, needs and levels of understanding.

Based on GEO as an empirical case study, this research will help understand the nature and interaction of multiple challenges associated with the construction and operation of global integrated assessment and reporting systems, the response to the challenges through assessment design, and the implications of the chosen design solutions for the usefulness of the results. A basic premise of my research is that the usefulness of environmental assessment and reporting is simultaneously determined by the ways the producers understand and address the needs of end-users and, at the same time, end-users having the interests and capacity to make use of the information provided. My goal is to further thinking on GEO, to help others deal with similar assessment and reporting systems, and to add to theory that may help other assessment and reporting systems in the future. Based on these, my key research questions are as follows:

> What are the implications of global change for the need for information and the conceptualization of assessment and reporting programs? How can assessments respond and what are the characteristics of the emerging approach? Where does GEO fit in that context?
>
> What are some of the main design elements in GEO, and how do they interact with criteria of effectiveness according to the producers and users of the assessment? What are the implications beyond GEO?
>
> Based on the findings, what alternative assessment and reporting strategies would increase the usefulness of integrated assessment and reporting in general and GEO in particular?
>
> If we were to apply the integrated assessment and reporting model in other contexts, what capacities would be required and how could these capacities be developed?

In comparison with most other work carried out under the GEA project that provided me with an opportunity to undertake a significant part of this research, I am casting a very wide net. GEO, my key case is one of the most ambitious and complex global assessment

and reporting initiatives ever undertaken, certainly in the environmental domain. Rather than focusing on one or a small set of variables, I am interested in several key design elements, their interactions with each other and ultimately their impact on effectiveness. While this reflects the complex reality of assessment and reporting system design with which practitioners need to deal, it sets a limit on the analytic detail with regard to particular design issues as compared with academic research focused - say - on participation or capacity alone.

Another limitation, which may be also a strength, is my history of close involvement with GEO, dating back to the very beginning of GEO-2000 when I took over the leadership of GEO related activities at the International Institute for Sustainable Development (IISD). This history has allowed me to contribute not only to global and thematic GEO reports, but also to contribute to capacity development activities and learn from the many colleagues involved. As a result, I had access to information, written and verbal, that would be hard for outside researchers to match. However, that experience increases the need for my objectivity to be clearly evident. Resisting the temptation to be overly descriptive and less analytic has been a challenge. Also, without any doubt, I developed biases and assumptions that are hard to get around as I write. Recognizing and, to the extent possible, getting around my biases arising from this situation while also building on its advantages has not been easy.

# 2. Effectiveness, assessment design, and the trouble with attribution

Analyzing the utilization and effectiveness of research programs has been of interest to social science at least since the early 1960s (McLean 2000). What counts as utilization and effectiveness? What makes evaluations and assessment systems useful and

effective? Why do people choose to use the findings of some assessments, and not others, even if the latter are potentially useful? The program evaluation research community has been dealing with such questions since the time when few of today's environmental assessment systems were in place (Leviton and Hughes 1981, Young and Comptois 1979). It is interesting, though not surprising that many of the findings twenty years ago seem to support the results of more current research aimed at understanding the effectiveness of environmental assessment systems.

In contrast with effects, effectiveness is not an attribute that has absolute measures one can measure irrespective of context. A recent study of the European Environment Agency provides a definition that is applicable to this research:

> "Statements about the effects of environmental measures are different from statements about their effectiveness, although the two terms are easily confused and are often used interchangeably. 'Effect' implies causality between a policy and its impact on the outside world. The process of identifying effects – both intended and unintended – is based upon scientific and social observation and analysis, and should be judgment-free. By contrast, assessing 'effectiveness' involves the further step of judging whether and how far the observed effects of a policy measure up to the explicit objectives set for it, and this involves comparing intentions with performance." (Vaz *et al.* 2001)

One of the key concepts of concern to the earlier school of evaluation research was utilization of the results of research and information. The 'bottom-line' criteria of utilization were seen as the function of information having been 'processed' or translated into their implications for relevant issues (Cook and Pollard 1977). It was also recognized that there must be some evidence that in the absence of the given information or information system, decision-makers would have acted differently (Leviton and Hughes 1981).

These criteria do not address the question whether utilization actually led to better decisions: they simple imply that the information was received, internalized and had an influence, one way or another. As environmental performance and sustainability become more widely accepted and applied evaluation criteria of policies and programs, the notion of what is a better decision and what are better outcomes is getting redefined[5]. Decision-making has probably never been as linear and straightforward as a Newtonian worldview would have us believe, but global change adds significant new elements of complexity and risk to which information and information systems need to respond.

## 2.1. Tracing the influence of information in the age of global change

According to theory, the information carried by a message that a specific event has occurred is inversely related to the probability of the event occurring. Simply put, the information value of rare events and phenomena is higher (Shannon and Weaver 1949). Many aspects of global change are not just rare they are without precedent. Some, like technological change, are the products of human ingenuity, and others, like the depletion of the ozone layer, are partly its consequences. Elements of global change also emerged at a fast pace, and typically in interaction with each other and social development. Theory suggests that under such conditions the value of information grows. But it has been also demonstrated that – at least in developed countries - technological progress and rapid change led to today's information explosion and overload. For an increasing number of people, the issue is not lack of information but separating 'noise' from message and meaning (Watanabe 1975, Wilson 1995). As important as scientific information and advice may seem to be to producers, it is increasingly competing for attention with information and influence from other sources.

The importance of information is increasing not only because of the emergence of new issues and interactions, but also because of the increasing risk and uncertainty associated with choices. Decision theory differentiates between decision under certainty where each

alternative has only one outcome, decision under risk where decision options have several possible consequences but with known probability distributions, and decision under uncertainty where each alternative has many possible consequences with unknown probability distributions (Losee 1990, Anon. 2001). Many choices associated with sustainability and global change fall in the third category. Decisions under uncertainty are likely to require subjective judgment based on expert knowledge, past experience or other factors that make decision-making as much an art as science. This is supported by critics of mechanistic decision-making models who also point out that uncertainties or not, humans tend to imperfectly understand alternatives and the connections between actions and outcomes. There are limits to rationality in decision-making – as is well known in psychology – people often fail to recognize and represent their own self-interest even if adequate information is available (Simon 1986, Jensen 1994). It appears that even if the information captures the attention of its audience, it is not at all guaranteed that it will be 'processed' and taken as seriously as its advocates might like.

The capacity of decision-makers to understand and systematically translate information on sustainable development into policies is often weak, as pointed out by Ann Dale who spent many years spearheading sustainability programs in the Canadian federal bureaucracy:

> "Another powerful barrier is the fundamental lack of ecological literacy among the bureaucracy, and most particularly at the political level. Moreover, the latter work in an environmental context of urgency, denial of alternatives and unreasonable deadlines so that sustained reflexivity and opportunities for new learning are virtually non-existent. This makes the political decision-making level even more dependent upon the quality of information they receive from their bureaucratic advisors, as well as the many external sources hoping to exert influence on their decision-making." (Dale 1998)

As Patton (1986) pointed out, researchers and decision-makers who desire to see evaluations and assessments used must take seriously the importance of identifying who

needs what information under what conditions for what purpose. According to earlier models, decision-making is a fairly linear process. Based on the definition of a problem the missing piece of knowledge is identified, the information is produced, transmitted, received, and used by the decision-maker. A classical and still often quoted model of communication by Shannon and Weaver conceptualizes the communication process as the mere decoding and transmission of a message (Shannon and Weaver 1949).

This linear and mechanistic 'transmission model' of decision-making is problematic and has received considerable criticism from social scientists of the constructivist school. Although there seems to be no single, accepted constructivist model of communication, some key attributes of such model can be deduced from the critiques (Chandler 1994, Thorngate 1996):

> communication is a non-linear process where the recipient is interpretively active and connected to the source through dynamic feedback loops;
> communication is never perfectly transparent and its outcomes in a social context can not be fully predetermined;
> information cannot be used without prior knowledge and the capacity to respond;
> communication is not about information, but meaning. Meaning is actively constructed, not passively extracted from books or other sources provided by whatever source;
> meaning is also influenced by the differing purposes of people and their power relationships. People without power and capacity may not have the means to react even when a message is perceived as relevant and timely;
> a message may represent many alternative meanings, so meaning is contained primarily in the interpretation;
> the interpretation of a message is strongly influenced by situational, social, political and other contexts;
> decisions are shaped not only by beliefs, but also values. However, values are affected by different information than are beliefs, and in different ways;

decision makers are prone to seek, ignore and use information in biased or irrational ways or change preferences without notice;
roles, purposes, relationships and contexts all change over time, favouring different interpretations at different points in time; and
the choice of medium matters because of social conventions and associated individual preferences.

The practical implications of the constructivist approach for understanding the utilization, effectiveness and influence of assessment and reporting systems are profound. They confirm that we need to look beyond information and its transmission and pay much more attention to context, process, history and other factors surrounding producers, recipients, and their interaction. The task is even more difficult in the case of assessments whose thematic scope and audiences are broad, as different segments of the population may construct different meanings around the information and thus react in different ways.

As Weiss (1977) observed, "the policymaking process is a political process, with the basic aim of reconciling interests in order to negotiate a consensus, not of implementing logic and truth." Looking at the use of social science in public policy, she found evidence that, contrary to expectations, decision-makers tend to use results of scientific research less to solve, rather to formulate problems and orient themselves with what she termed the 'enlightenment function' of research. "Knowledge creep', another term coined by Weiss, indicates the way ideas become gradually accepted and dominant in science or policy discourse. To complement this view, however, subsequent research pointed to an 'action function' because under certain conditions policy-makers do seek out and commission specific research to help answer very specific policy questions (Deshpande 1981). It appears that over the years since this early work the 'action function' of environmental information has been increasing and there are clear cases in the field of global environmental governance where the link between assessment systems and policy effects can be and have been made. Particularly after Rio, strengthening the science-for-policy aspect of assessment has become an interest not only to scholars but also to international agencies, such as the United Nations System-Wide Earthwatch (Fritz 1997).

In either case, decisions are influenced by many factors, scientific and non-scientific, information-based and not. As a result, tracing the effect of a particular piece of information and report may be very difficult; to compound the problem, sometimes even policymakers themselves have trouble identifying specific inputs that influenced particular decisions. Following the arguments of Weiss and Deshpande, measuring effectiveness based on the enlightenment function of information would require determining changes in orientation, an influence too vague and complex to measure effectiveness. Even if the assessment serves an action function and has closer ties to the conceptualization and resolution of particular problems, the decisions are likely to be influenced by a large number of factors, so that attributing the decision to a particular piece of information is problematic.

Information and assessment systems influence decision-making, but on the macro scale the expectation is that they influence the evolution of issue domains, broadly defined as the composite of (a) the participants and participant networks actively engaged in the issue, (b) the social institutions that influence the interaction among the participants involved, and (c) the policies, decisions, behaviours, and impacts of these on the environment (Clark *et al.* 2002). It has been argued that shifts in environmental issue domains have taken place over a decadal time-scale or more (Sabatier and Jenkins-Smith 1993, Ney and Thompson 2000). However, GEO has been around for only a few years. As one of my interviewees observed:

> "It turns out in the lifecycle of a repeated publication that it takes a long time to build up awareness, familiarity, credibility and distribution so that you can have some impact. And for a new report that's very difficult. … The World Resources Report has been out there for 15 years. So if you look at the impact of the report, I am not surprised if they didn't find too much, but I think it's the wrong measure." (A. Hammond, World Resources Institute, interview)

In the case of GEO and possibly other distributed global assessments, there are further complications because of the requirement to make results available at least in the six

official UN languages (Arabic, Chinese, English, French, Russian and Spanish). The first two GEO reports were produced in English, and it took long time for translations into other major languages to appear. For instance, although GEO-2000 was released in September 1999, its Russian translation was published only in early 2001, after the GEO-2000 user study was already completed:

> "In a country like Russia and the CIS, as a rule, people don't know a foreign language. (…) The Russian translation of GEO-2000 has been out just about 2 months ago. So it's still at the very beginning of distribution and one has to wait a couple of years to get any sense of impact."
> (G. Golubev interview)

Given the long wavelength of issue domain change and the short history of GEO, at this stage one cannot reasonably expect to uncover systematic evidence to convincingly demonstrate effect on decisions and decision outcomes.

## 2.2. Criteria of effectiveness

Given the difficulty of precisely evaluating GEO as a complex and relatively new assessment and reporting system, I will follow the approach adopted by other researchers who instead of looking at direct measures of success or effectiveness found it more feasible to identify criteria that can be shown to be correlated with it.

Studying the criteria affecting the utilization of evaluations, Leviton and Hughes (1981) reaffirmed that utilization is not the same as utility: an assessment can have utility to a user without being actually used. They view utilization broadly and following previous authors identify three categories analogous to the enlightenment / action function described earlier: instrumental use where decision-makers could cite and document the use of the information in specific cases, conceptual use where the information is influencing thinking and approach to a problem without specific identifiable or

documentable use, and persuasive use where the information is used to convince others to defend or support a particular position. They also identified five variables affecting use and effectiveness as follows:

> Relevance of the assessment, addressing the real needs of the audience at the time the information is required;
> Communication between producers and users of the assessment without the loss of distortion information (e.g., in executive summaries);
> Presentation and processing of information in forms most useful for the audience;
> Credibility of information in light of knowledge from other sources; credibility and independence of the producer;
> User involvement and advocacy by key individuals in getting assessments used.

A similar set of criteria, including technical adequacy (≈ credibility), value (≈ relevance), legitimacy, and effectiveness has been identified by Clark and Majone (1985). They also pointed out that partial perspectives of integrated science / policy assessments cannot do justice to their synthetic character. Such assessments have to be evaluated in their entirety, treating these criteria as a set.

Having more specifically studied environmental assessment systems and the most frequent causes of their failures, the GEA project has identified saliency/relevance, and credibility as the most critical factors (Jäger and Farrell, forthcoming):

> Credibility to reflect the technical believability of the information;
> Saliency defined as the ability of the assessment to address the needs of a particular user or users;
> Legitimacy to ensure that the assessment is carried out in a fair and politically acceptable way taking the views and values of its audience into account.

This criteria set maps well, although not entirely on the one identified earlier e.g., by Leviton and Hughes (**Table 1**). Both cases recognized credibility and saliency / relevance

as key criteria, the first set did not include legitimacy, while the GEA set does not separate out the communication / presentation aspect and the role of individuals. For the purposes of this research, I will build on the criteria identified in the GEA work, but given the characteristics and goals of GEO I borrow and to some degree reinterpret two additional criteria from Leviton and Hughes (1981). Following the GEA philosophy, I apply these criteria, saliency/relevance, credibility, legitimacy, awareness and individual advocacy as proxy or indirect measures of effectiveness. Based on this logic the design and effectiveness of an assessment can be analyzed on the basis of its potential to increase saliency, credibility, legitimacy, or awareness. I separate out individual advocacy as I consider it more an outcome of meeting the other criteria.

**Table 1:** Criteria of effectiveness.

| Leviton and Hughes (1981) | The assessment's utilization and effectiveness would be enhanced when it is… |
|---|---|
| relevance | relevant to the needs of a particular audience and made available at the right time |
| communication | characterized by close communication among producers and consumers of the information |
| information processing | clearly presented with a suitable mix of quantitative and qualitative information |
| credibility | presented by producers in high standing and resonate with information available to the decision-maker from other sources |
| user involvement and advocacy | geared to the particular individuals receptive to the information and willing to engage in persistent advocacy |
| GEA (based on Eckley 2001) | |
| saliency/ relevance | known to the user and if that user deems that assessment relevant to current policy and behavioural decision |

| | |
|---|---|
| legitimacy | conducted in a manner that allows users to be satisfied that their interests were taken into account, and that the process was a fair one |
| credibility | believable from the scientific and technical perspective to a defined user of that assessment, often in the scientific community |

In order to produce criteria that would be most useful in the analysis of GEO, the two sets could be combined. Saliency / relevance is a key concern for GEO and similar IAR systems, although there are some characteristics that set them apart from focused traditional science assessments. By definition, GEO is broad and its mandate is to provide a comprehensive analysis rather than answer straightforward and clearly defined policy questions that are of interest to a narrow set of policy audiences. Given its broad audience and its broad policy agendas, GEO reports cannot be equally salient to all, but they definitely have to be salient to their core audiences, which is UNEP's Governing Council on the global level and national governments on the lower scale. The picture is further complicated in that GEO is not simply a report, but an information system with diverse products, participants, and a strong emphasis on process, scanning multiple spatial scales. Understanding what contributes to the system's saliency is thus key, but finding an answer is not simple.

The other common criterion identified is credibility. Formally the GEO report ought to be based on the 'best science' regarding key policy issues. The process has significant elements of quality control to ensure this is the case. However, due to its broad focus and very different institutional and analytic capacities among regions, ensuring scientific credibility is a challenge. It is also challenging because GEO cannot provide detail on very specific issues; some analytic details, uncertainties, and complexities do not get as thorough treatment as they would in strictly scientific assessments. Yet, credibility is particularly important if UNEP is to fulfill its mandate as the global authority on environmental information.

An essential question regarding credibility is credibility to whom. The audience of GEO and other similar assessments is broad, possibly broader than most environment agencies.

As Rob Maas of RIVM pointed out, it is important that the environmental outlooks and research they produce under the main umbrella of environment also have to be credible to the Dutch Minister of Finance, the Employers' Union and other organizations. Although these are not normally thought of as environmental organizations, their decisions may at times have more profound and far-reaching environmental consequences than environmental policies, narrowly construed. The same logic and yardstick probably applies to GEO and beyond. For integrated assessment and reporting systems, it is not enough to be seen as credible and authoritative on the environment only among environmental groups. They also have to be credible to a broader circle of organizations and constituencies with influence on and interests in the environment.

The cross-scale aspects of GEO have significant implications for the assessment's *legitimacy.* Challenges arise from the need to ensure the compatibility of regional analyses in order to construct a globally coherent picture and messages. But GEO also must allow sufficient regional autonomy and freedom to collaborative centers to bring their values and worldviews to bear on the assessment. While the interaction of UNEP and its collaborative centers may lead to mutual legitimization (i.e., UNEP legitimizing CCs, particularly in developing countries, by providing them fora and recognition globally, and CCs playing a role in legitimizing UNEP in the regions in which they are based), making sure this is the case requires carefully choreographed interaction among the partners.

An assessment and reporting system may speak to salient policy issues, meet the most rigorous scientific standards, and represent the real and legitimate views of stakeholders, but yet it may still fail to make a difference if its findings are not brought to the attention of target audiences. These audiences can be narrow and well defined, or – as it is the case with most SOE reports – very broad. Even in traditional scientific assessments, it is no longer sufficient to assume that uncovering scientific truth will automatically lead to awareness of findings. I define awareness as familiarity of the assessment and the implications of its key results for the policies and actions of a given audience. As earlier mentioned, decision-making is increasingly characterized by information overload and

competition for attention. Awareness needs to be addressed through focused communication strategies integrated into the overall assessment and reporting program. This is not limited to the 'advertisement' of findings and products of the assessment; in fact, awareness has to be built through the entire assessment process when producers and potential audiences interact.

As making a decision often is in the domain of individuals, assessments can be more or less effective depending on the extent they influence people in power and with policymaking authority. Individual advocacy has been mentioned as a key factor beyond the direction and to some degree success of regional and sub-regional components of GEO. I do not consider individual advocacy a separate criterion as it is not an attribute of the assessment itself – it is more an outcome if other criteria are met. However, the assessment can still make an effort to influence selective target opinion leaders and high-powered individuals and thus, increase relevance, legitimacy, visibility and potentially impact.

Individual advocacy may play a role both at the producer and user end, but designers of assessment and reporting systems have more influence over those producing the information. There is some evidence from GEO that individuals with particular backgrounds, such as former environment ministers have been particularly helpful in both the production and follow-up phases of the GEO process on the regional and sub-regional level (Golubev interview). The advocacy of individuals on the user side is more difficult to pin down, particularly because user groups are very diverse. The assessment may increase its effectiveness if there is a nuanced understanding of the role particular individuals play among target audiences. The association with such individuals may be as little as making sure they receive copies of GEO outputs, or as ambitious as forming *ad-hoc* alliances to represent and leverage support for mutually advocated positions at critical fora. Such formations have been described as advocacy coalitions and shown to have the potential to play a major role in the shaping of environmental agendas (Sabatier and Jenkins-Smith 1993).

Assessment and reporting systems that seek out and cultivate relationships with key advocates may significantly increase their utilization, effectiveness and impact. However, building a coalition of advocates should be more feasible around an assessment and reporting system that is seen as relevant, salient, legitimate and familiar to audiences. Advocates may not only support, but also energetically oppose findings. Influential individuals, whether with credentials in the scientific or policy world, who hold and voice dissenting views can cause considerable damage to the assessment, although depending on who is 'right' or 'wrong' constructive criticism may be very helpful in clarifying issues and improving the assessment system in the future.

Beyond the role of individuals, producers and users of assessment and reporting systems may have different interests and perspectives on how the system works, how it has been utilized and what it has achieved. My research is biased towards the producer side given that producers are easier to identify, and identifying a somewhat, even if not statistically, representative sample of users was not feasible. Considering only the main GEO report, its user community would include not only the UNEP Governing Council, but national governments, students in academic organizations, the media and others. User groups are geographically dispersed, and not only read GEO in different languages, but also at different times, as it may take over a year for some foreign language versions to appear in print. A further complication is that the global GEO report is only one of many products and processes through which the assessment exerts influence.

The criteria I have used may not apply equally to all assessment and reporting systems, but I will argue that they do apply to GEO. I also agree with those who point out the possible interactions among evaluation criteria. In some cases, the interactions are mutually reinforcing, for instance, interaction with target audiences could lead to increased relevance. In other cases, improving an assessment with respect to one criterion may come at a price of decreasing effectiveness on another, as could be the case when increasing saliency for policymakers comes at a cost to scientific credibility. These and other tradeoffs are clearly present in GEO and need to be made conscious at the design (or redesign) stage of the assessment and reporting system.

# 3. The context of global assessment and reporting systems

Global environmental trends and dynamics are of increasing interest to a growing number of audiences, as the signals of change became virtually impossible to ignore. Parallel to the growth of the human enterprise, its impact on the Earth system grew in terms of scale, complexity and severity far beyond what it was just a century ago. While in the past local forces of change lead to mostly local impacts, today interconnected local and global forces of change associated with population increase, technological change and continuing economic growth contribute to a web of impacts whose signals are detectable in both local communities and planetary biogeochemical cycles (Vitousek *et al.* 1997). Like it or not, we have become none less than architects of a new epoch, anthropocene, with less than adequate knowledge and capacity to manage the system in a sustainable way (Crutzen and Stoermer 2000). From being an abstract concept, global environmental change has become tangible in many regions around the world. People can experience its effects in everyday life, even if the association is not always made. Scaling down, understanding local implications of global dynamics, and scaling up, understanding the contributions of local choices to global outcomes often require a level of abstraction that makes describing and understanding the connection challenging. Making this linkage is a particular challenge for coarse scale analysis where the understanding of local diversity and its relevance can easily be lost (e.g., Wilbanks and Kates 1997).

Although the situation is changing, many of today's monitoring, assessment, and reporting systems still operate on the basis of earlier, compartmentalized world-views and data infrastructure that reflect bygone interests and priorities. A recent review of the use of water related monitoring data in the EU revealed that only about 25% of what is routinely measured has relevance for policymaking, and this is true at the time when some critical variables are still unmeasured (R. Uhel interview). There is no reason to

believe the situation is much better elsewhere. In particular, developing countries continue to be inadequately covered by information systems – certainly from the perspective of what gets communicated to the local population[6]. Capacity constraints and more immediate priorities broadly associated with poverty and its consequences usually mean they can rarely maintain more than a token presence in information and assessment exercises. Given these and other systemic deficiencies, assessment and reporting requires fundamental improvement in institutional architecture, research methodologies, and links to decision-making processes to address the issue of sustainable development as an overriding social objective in its very details that affect peoples' lives. Although there is nothing new in this statement, much work remains to be done (United Nations 1993, Carnegie Commission 1992, Young *et al.* 1999, etc.).

Although the first major international conference dealing with the relationship of environment and development took place as early as 1969 in Founex, Switzerland, it was the United Nations Conference on the Human Environment in 1972, the World Conservation Strategy (International Union for the Conservation of Nature, United Nations Environment Programme and World Wildlife Fund 1980), Our Common Future (World Commission on Environment and Development 1987), and particularly Agenda 21 (United Nations 1993) that popularized the concept and framed it as 'sustainable development'. Chapter 40 of Agenda 21 pointed out that if sustainability is indeed a legitimate and feasible social goal, society must find ways to differentiate between sustainable and unsustainable forms of development. Recognizing that this required new types of assessment and information systems, several groups developed general principles to provide more detailed guidance than Agenda 21, which discussed this issue in very general terms. Guidelines in the Bellagio Principles and the Balaton Group's work on sustainability information systems, among others, suggested that in order to effectively address the trends and interactions of environment and development, there was a need for rethinking not only what information needed to be produced, but also how, by whom, and in what institutional settings (Hardi and Zdan 1997, Meadows 1998). Both these groups and others took the view that sustainability was not a destination but a road to travel, a dynamic and multi-scale process of learning and iterative adaptation to

typically imperfectly understood changes (Holling 1978, Kay *et al.* 1999, National Research Council, Board on Sustainable Development 1999). The problem really is a new condition in which misunderstood or poorly understood change is continual. It is not a poorly understood new state, it has gone from (assumedly) static to clearly dynamic.

Referring to sustainability assessment, Hodge (1995) proposed a conceptual framework with three domains: human system, ecosystem, and interaction. He also showed that by the mid-1990s, a significant number of private and public sector organizations were initiating holistic reporting on environment and development, even if terminology varied. These were often building on existing structures - such SOE reporting in government, or corporate performance reporting in the private sector. However, they also tried to go further in terms of scope, degree of stakeholder participation, use of indicators, scenario analysis, association with targets and thresholds, and other aspects. Many initiatives were requested by governing bodies and internalized the rhetoric of sustainability and integrated assessment (IA), but provided little guidance about what these terms meant in detail for the way assessments were to be carried out or how the results were used in formulating policy. In fact, there would be expectations that a better understanding of sustainability would actually emerge from assessment and reporting programs, through learning by doing (Pintér 1997). This supports a view of assessment systems as social learning enterprises that provide opportunities for the organization of existing knowledge, the development of common frameworks, and the identification of shared and individual interests to facilitate policy-making (The Social Learning Group 2001, Long and Iles 1997).

## 3.1. Towards a typology of assessment systems

Assessments are produced by many institutions and for a wide variety of audiences and purposes. I look at a specific type, and it will help later analyses to understand where my cases belong on a family tree of assessment systems. Although a widely accepted,

systematic typology is yet to emerge, some broad categories can be teased out of the literature. Clark *et al.* (2002) identified the following broad categories:

> Assessment systems producing new knowledge, characterized by norms, methods and institutional structures of basic or curiosity-driven science;
> Assessment systems closer to applied science that respond directly to the needs of decision-making, building on the findings of basic science, and often using scientific method; and
> Systems that draw on both basic and applied science and serve to organize, interpret, and communicate both scientific facts, as they relate to normative policy interests and values of social stakeholders.

If the assessments in the first group operate clearly in the domain of science, and the second on the boundary of science and policy, assessments in the third category cross multiple boundaries and an even broader circle of participants and audiences. They build on science, but also build on and reflect normative positions of various stakeholder groups. If the primary purpose of basic science type assessments can be seen as discovering truth, and the purpose of applied science is to bring the best science to bear on specific policy questions, assessments in the third category help organize, interpret and communicate scientific facts, and associate them with policy questions on environment and development that are relevant for a wide range of audiences. This view is close to what Connolly *et al.* (1998) referred to as international information institutions (IIIs) "primarily focused on causal relationships and states of the world". I put emphasis here on world versus narrow issues or a particular policy question. 'World' can mean of course, less than the entire planet (e.g., a community, a country or an ecosystem). That global and local perspectives, home and planet are both important was recognized a long time ago; in fact it was well articulated even prior to the 1972 UN Conference on the Human Environment (Ward and Dubos 1972). The emphasis is on the whole system, integration, and the entirety of a locale versus some of its narrower aspects. This approach has been suggested by several authors, labelled, among others, as 'place-based' assessment (Wilbanks and Kates 1997), 'regional integrated assessment' (Rotmans 1998;

Knight n.d.) or integrated region-oriented environmental policy – IREP (van der Kolk and Dekker 1999). There are further details of this notion of IIIs that resonate with the evolving practice of sustainability assessment: they characterize both natural and anthropogenic aspects of the empirical world; characterize the state of knowledge of those aspects; and, forecast possible futures.

In addition to the III description of what assessments do, I include how they are prepared, communicated and used and how they are institutionalized, following the broad approach previously adopted in the Global Environmental Assessment (GEA) project. The parallel emphasis on product, process and institutionalization is becoming increasingly evident and inherent in many assessments[7].

Keeping in mind that it is a combination of substantive, procedural and institutional factors that define an assessment or an assessment type, my shorthand for the assessments with which I deal will be Integrated Assessment and Reporting (IAR). Broadly speaking, these would fall under the third category of assessments connecting basic and applied science and policy, as described on page 27 and Connolly's international information institutions, wit a few amendments. The cases in this category are linked to and to some degree evolve from integrated assessment (IA) methodology, linked to non-transient institutional arrangements, and they involve publishing specific outputs on a periodic basis. The emphasis on outputs also highlights the importance of effective communication that may receive lower priority in other assessments, particularly in basic science type assessments where communication is either in the form of publication in limited circulation scientific journals or reporting to clearly but very narrowly defined policy audiences. The attributes of IAR are provided in **Table 2**; I apply them here to identify and analyze key design aspects of my cases.

Although they share common attributes, the assessment systems I include in the IAR category are not identical. The World Resources Institute (WRI) and the European Environment Agency (EEA) have attempted further categorization of reporting and assessment systems that I include in this cluster. The more systematic EEA classification

refers to environmental reports, indicator reports, statistical compendia, thematic reports, other reports, and unpublished reports (WRI 1996; Briggs, de Hoogh, and Wills 1999). While there may be practices that fall clearly into one of these categories, an increasing number of emerging systems seems to serve multiple purposes and lead to multiple products, including more qualitative and analytic environment reports, core indicator sets, and statistical compendia. This trend is apparent in the case of many national SOE reports that often serve as starting points for more holistic, 'quality of life' or 'sustainable development' reports.

**Table 2:** General criteria of integrated environmental assessment and reporting initiatives.

| CRITERION | EXPLANATION |
|---|---|
| Organizational | |
| Institutionalized | Requirement for organizational host with adequate capacity, legitimacy and mandate to carry out assessment and reporting on a continuous or iterative basis |
| Structured for stakeholder participation | Suitable fora, mechanisms and channels for stakeholder participation throughout the entire assessment cycle from problem framing, through issue identification, priority setting, development of recommendations to evaluation and redesign of assessment strategies |
| Policy relevant | Direct interface between producers of the assessment and policy communities with specific information needs and agendas |
| Content related | |
| Holistic | Coverage of both socio-economic and environmental factors and their interactions |
| Retrospective and forward looking | Explanation of past trends and information on possible future directions, often involving modeling and policy scenarios |
| Inclusive of externalities | Characterization of externalities, or at least making positive and negative externalities explicit |

| | |
|---|---|
| Quantitative and qualitative | Maximum use of data and indicators obtained through monitoring networks; integration of qualitative information as appropriate |
| Accommodating of uncertainties | Explicit recognition of uncertainties and their implications for decision-making and outcomes |
| Change-sensitive | Identification and early warning about changes in system behaviour |
| Scale-sensitive | Recognition and analysis of cross-scale linkages where applicable |
| Procedural | |
| Scientifically valid | Requirement for scientific input and preferably peer-review |
| Communication-oriented | Well-formulated and realistic communication strategies of the needs and capacities of clearly outlined audiences |
| Self-reflective and adaptive | Assessment process includes built-in periodic evaluation that feeds back to process planning |
| Sensitive | Ability to accommodate new knowledge |

The description of IAR closely fits the Global Environment Outlook and represents a test case of a novel approach. The GEO assessment uses driving force-pressure-state-impact-response as its conceptual framework, thus by definition it covers not only the environment but also associated socio-economic causes and response mechanisms. Although it does not typically commission primary scientific research, it compiles and relies heavily on the results of recent scientific studies to provide a view of the state of knowledge related to key policy issues. GEO also includes an analysis of alternative policy paths using both modeling and scenario narratives.

In summary, the constellation of several attributes make GEO a particularly interesting case of IAR. It fits Clark *et al.*'s (2002) description of assessment systems that organize, interpret and communicate scientific facts as they relate to stakeholder values and policy,

and Connolly *et al*.'s (1998) description of IIIs as constructs that are focused on the description of causal relationships and states of the world.

Although not to the same extent, at the first approximation GEO satisfies most of the IAR criteria. Its key characteristics can be described as follows:

Sustainability orientation from the perspective of the environment;
Mandate provided by the UNEP Governing Council;
Decentralized organizational structure, with semi-autonomous collaborative centres coordinated through global and regional mechanisms;
Cross-scale assessment, global overview but regionally differentiated perspectives;
Cyclical assessment process;
Emphasis on identifying cross-cutting and emerging issues that would not be visible from the perspective of a single discipline, region, stakeholder or a limited time period;
Formal consultations with policymakers and other stakeholders on the regional level in shaping the assessment, and high-level involvement in developing action-oriented recommendations;
Exposure at high-level fora such as global environment ministerial summits;
Semi-quantitative analysis of policy options and scenario paths in regional breakdown, building on the results of a retrospective analysis of the state of the environment and policy drivers;
Diverse suite of products serving different audiences;
Systematic assessment and reporting capacity development at the regional and national level.

The GEO approach and its evolution has been described in some detail in the two global GEO reports published to date and further details can be obtained from the GEO production guidelines (e.g., UNEP 1997, UNEP 2000). It is useful to review the GEO architecture, process and methodologies to place the rest of this work in context.

## 3.2. The Global Environment Outlook

As the lead environmental organization in the UN system, UNEP's mission is "to provide leadership and encourage partnerships in caring for the environment by inspiring, informing and enabling nations and people to improve their quality of life without compromising that of future generations" (UNEP 1998). As required by its mandate, the organization produced global SOE reports on the tenth and twentieth anniversary of the Stockholm conference (UNEP 1982, Tolba and El-Kholy 1992). However, in the absence of an agreed-upon integrated framework, their orientation oscillated from the socio-economic to the natural sciences (G. Golubev interview). Starting in the early 1990's, UNEP found itself in the company of an increasing number of institutional players addressing environmental issues from a thematic (e.g., UNFCCC and IPCC) or broader integrative perspective (UN-CSD), increasing the difficulty but also the importance of cross-agency coordination. Parallel to this, UNEP was facing a structural shortage of funding that constrained the implementation of its core programs (Bakkes *et al.* 1999). In response to this challenge UNEP's Governing Council issued the Nairobi Declaration that reaffirmed and further elaborated the core elements of UNEP's mandate, expressing its wish that UNEP remains the authoritative voice on the environment in the UN system. That declaration was also confirmed by the United Nations General Assembly (UNEP 1997a, UN 1997)[8]. Bakkes *et al.* (1999) also observed that UNEP's monitoring and assessment strategy should be implemented in the context of the organization's overall mandate and concluded that in GEO many elements of the strategy started to fall into place.

GEO's roots go back to decision 18/27 C of the UNEP Governing Council on May 26, 1995; that decision requested a new, integrated, forward looking, and regionally differentiated assessment on the state and direction of the global environment (**Appendix 1**). The GEO reports have become known as the 'flagship' products of UNEP, and have become a cornerstone of the organization's Environmental Observing and Assessment Strategy (UNEP 2000a, 2000b). The first report was published in 1997, and by 2001 GEO had completed two reporting cycles, the third cycle will be completed just in time

for the United Nations Conference on Sustainable Development in Johannesburg in 2002 where GEO is intended to serve as one of UNEP's main inputs (UNEP 1997; UNEP 2000). GEO's scope is very broad: it covers all priority environmental issues from the global to regional level for all regions of the world. The structure of the report evolved from one cycle to the next in response to emerging ideas, issues and policy priorities. These changes and their reasons are apparent from the GEO reports and can be followed in consecutively issued production guidelines (e.g., UNEP 1997b, 1997c, 1998a, 1998b, 2000d). GEO follows a specific thematic structure (**Figure 1**).

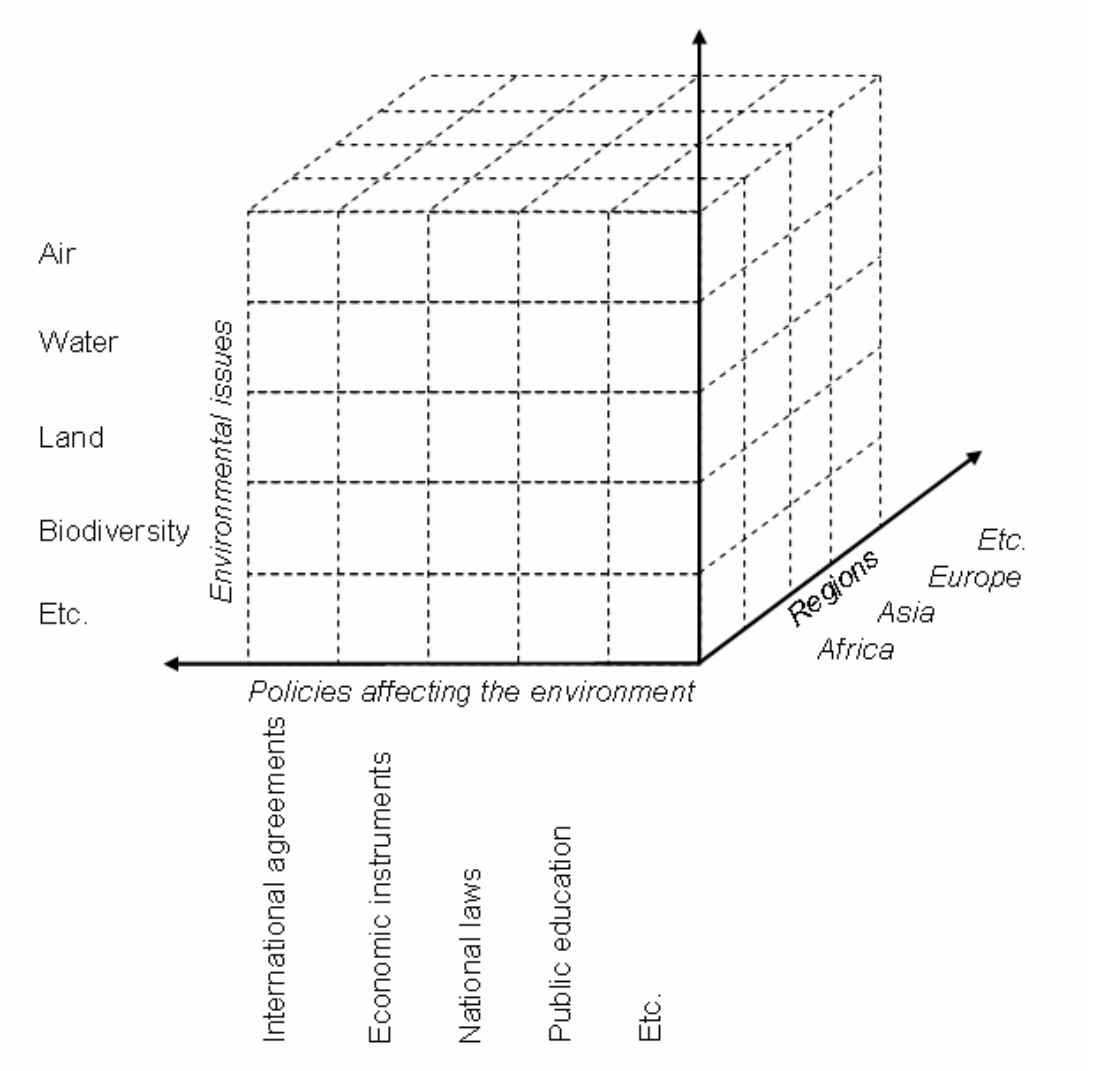



**Figure 1:** Spatial and thematic integration in GEO (modified from UNEP 2000e).

GEO's target groups or audiences are diverse. One may often find that a list of 'key' audiences easily becomes all-inclusive and thus meaningless. However, the audiences of global reports intended for public consumption *are* indeed broad, even if priorities exist. At the highest level, audiences for UNEP's monitoring and assessment strategy include the UN General Assembly and the UNEP Governing Council that as representatives of governments exercise formal oversight on the highest level. Additional priority audiences have been identified as specialized UN agencies, environment conventions, multilateral financial agencies (e.g., Bakkes *et al.* 1999). In addition, regional environmental institutions, the media, the scientific community and universities have been identified as important groups. GEO's challenge is that depending on the region, audiences may shift; while governments in Europe and North America hardly need more reports on the state of the environment, in many other regions (e.g., in Africa) GEO may well be the only integrated assessment and reporting process with significant presence and relevant products. The constant pressure for more detail, expanded products and even broader focus must be balanced with the comparative advantage and carefully defined mandate of UNEP in the UN system that does not directly extend, for better or worse, to the national and sub-national level.

GEO and many other integrated IAR-style state of the environment reporting systems are not set up as more traditional science assessments, such as the Intergovernmental Panel on Climate Change (IPCC). GEO "moves quickly across the landscape, harvesting anything it can, and has a political involvement process" (A. Hammond interview). Clarity in purpose with regard to the science / policy boundary seems to be particularly important. GEO does not at this point, respond to specific policy questions in the sense that some of the assessments carried out for example by the European Environment Agency (EEA) respond to the questions of the European Commission and Parliament or on the national scale the Dutch National Institute of Public Health and the Environment (RIVM) prepares assessments in response to requests by ministers. In contrast, its strength is clearly in linkage to policy agendas and senior and high level policymakers. "GEO should not be drawn into the science; we should get out of the science what they need in order to get the policy message" (V. Vandeveerd interview).

Although the UNEP Governing Council and the Observing and Assessment Strategy do not provide detailed instructions on how GEO should be done, the key expectations are clear (UNEP 1998):

> Provide an overview of global trends as well as regional detail;
> Respond to the information needs of key decision-maker groups, but also communicate issues in a way understandable by the general public;
> Based on integrated environmental assessment methodology, looking at interactions across spatial scale, at least between the global and regional scale (regional in terms of UNEP's regions), and across environmental issues;
> Besides a retrospective analysis, provide a future perspective;
> Utilize the best available science;
> Contribute to improved assessment and reporting capacity in developing countries.

UNEP designated its Division of Early Warning and Assessment (DEWA, formerly Division of Environmental Assessment and Early Warning or DEIA-EW) as GEO's coordinator. UNEP-DEWA maintains a key coordinating role, but the assessment process is decentralized and relies on the contribution of a network of Collaborating Centres (CC) that bring regional or thematic expertise and legitimacy to the process. CCs are not part of UNEP, but they carry out their part of the assessment under UNEP's guidance. GEO produces a range of outputs that also reflect this multi-scale character. A key product is the global GEO report that has been produced with a frequency of 2-3 years. Regional GEO reports are based on the global GEO framework, but they contain more detailed information on regional issues. Although GEO does not usually go into national scale detail, national and even sub-national data and assessments carried out e.g. by national governments are essential reference materials. GEO also publishes thematic reports that deal e.g. with methodological issues of multi-scale integrated environmental assessment. **Figure 2** illustrates GEO's nested structure, including its process and products, and the connection between GEO and issues and assessment systems below the regional level.

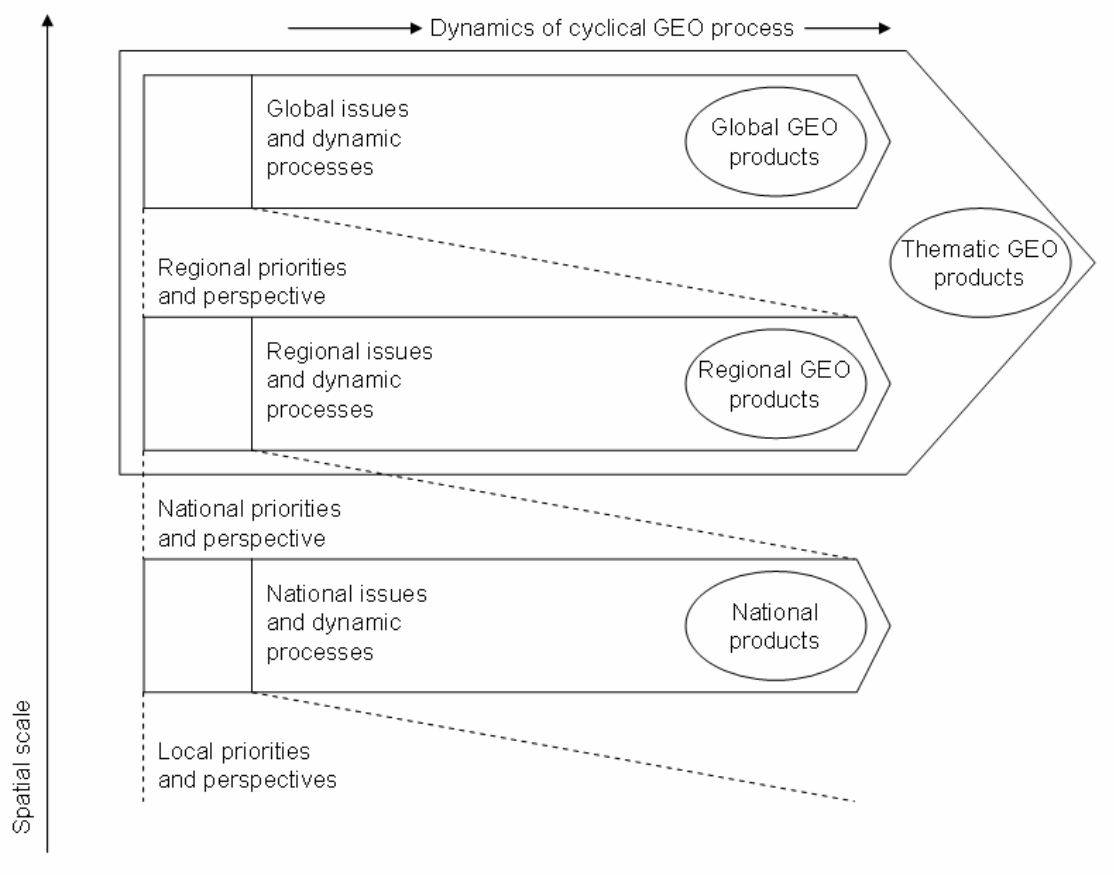



**Figure 2:** Integration of process and products in GEO across scale (Modified after Pintér 1997).

The best-known aspects of the GEO programme are its products, particularly the GEO reports in electronic or printed form. The products are intended to address particular needs and to respond to particular requirements in the Observing and Assessment Strategy. The main GEO reports serve as reference volumes with detail, indicators and analysis; an executive summary of the main report is intended to reach high-level decision-makers; a report for youth follows the logic of the main GEO report but uses language and design that is better suited to the interests of young readers; specialized reports address narrower thematic issues that emerged during the GEO assessment and may be of interest to expert audiences; capacity development materials are intended for trainee audiences. The main outputs are listed in the GEO prospectus (UNEP 2000c).

GEO, like most environmental assessment systems is heavily dependent on data. It is both strongly affected by data constraints and well positioned to report on problems from the user perspective. Core datasets are compiled both from secondary and primary sources such as the United Nations Food and Agricultural Organization (FAO), the U.S. National Aeronautical and Space Administration (NASA), the World Bank and others. In some cases, data are complemented from national databases that may be more up-to-date and accurate. By early 2002, many of the core data sets were available through an Internet data portal for the use of collaborative centers (UNEP/GRID-Geneva 2001). However, UNEP has systematically emphasized that GEO is not a data report and does not involve developing a data warehousing function or capacity other than making data essential for the assessment process available to CCs.

The GEO methodology is based on integrated environmental assessment. GEO's analytic framework builds on the pressure-state-response (PSR) model, best known from the environmental performance reviews of the Organization for Economic Cooperation and Development (Organization for Economic Cooperation and Development 1993). The PSR framework helps identify and establish linkages between pressures on the environment arising from human activities or natural processes, the resulting environmental states or conditions, and policy responses to those conditions. In GEO-2000, PSR was extended to a modified driving force-pressure-state-impact-response (DPSIR) framework, following the European Environment Agency's (EEA) lead. In comparison with the PSR approach the DPSIR model treats driving forces (e.g., industrial production, transportation or demographic change) and impacts (e.g., human health impacts, biodiversity loss, economic damage) separately (European Environment Agency 1999). The structure of the GEO reports follows, but does not in all its aspect directly mirror the DPSIR logic. It includes both stable and more transient structural elements (e.g., a planned vulnerability assessment section for GEO-3). These may appear as separate sections of a GEO report but also mean changes in the assessment process and the responsibilities of CCs with respective contributions.

The following are the stable structural elements:

Socio-economic baseline: driving forces of economic development and environmental change (e.g., demographic trends);

Integrated SOE / policy analysis: analysis of environmental trends, dynamics and interactions and associated policy pressures and responses;

Scenarios: identification and analysis of policy options;

Recommendations: actionable policy solutions related to priority issues.

Other sections that so far appeared as a separate section only in one GEO report included the analysis of the effectiveness of multilateral environmental agreements or, in GEO-3, an integrated vulnerability analysis of coupled environmental / socio-economic systems.

Ensuring the consistent application of methods is a major challenge for any polycentric assessment system such as GEO. Similar to other multi-level assessments such as the Intergovernmental Panel on Climate Change (IPCC), UNEP issued detailed production guidelines. Their preparation is coordinated by UNEP, but the guidelines have been discussed with CCs to gather and incorporate feedback. In contrast with IPCC worksheets and guidelines, the GEO guidelines contain fewer details.

In terms of institutional arrangements, GEO is a collaborative effort of a hierarchy of organizations that may serve different purposes in various stages of the process. The partnerships are organized on the basis of comparative advantage. UNEP serves as overall coordinator and the provider of methodological guidelines and standards, but from GEO-1 through GEO-3 it has served less and less as content provider. A global network of CCs covering all regions of the world brings local data and perspectives and as the GEO program evolved, contributed an increasing part of the regional analysis. Between the UNEP center and CCs are regional UNEP offices that liaise with governments and support CCs in their region.

Beyond the institutional structure which bridges from the global to the regional scale (and back), GEO also involves Associated Centres (ACs), organizations that are also external

partners like CCs but their role is defined in broader, usually thematic terms (e.g., expertise in data systems, scenarios, or capacity development) with no or only a secondary role in the assessment process in their own region. Of course both CCs and ACs have their own thematic or regional networks upon which they may rely on, with whom they work or try to influence depending on who they are and where they are in the assessment process. In addition to these groups, UNEP also cooperates on the global level with other organizations and networks with a global mandate, such as the partners associated with the Integrated Global Observing Strategy (IGOS; **Figure 3)**.

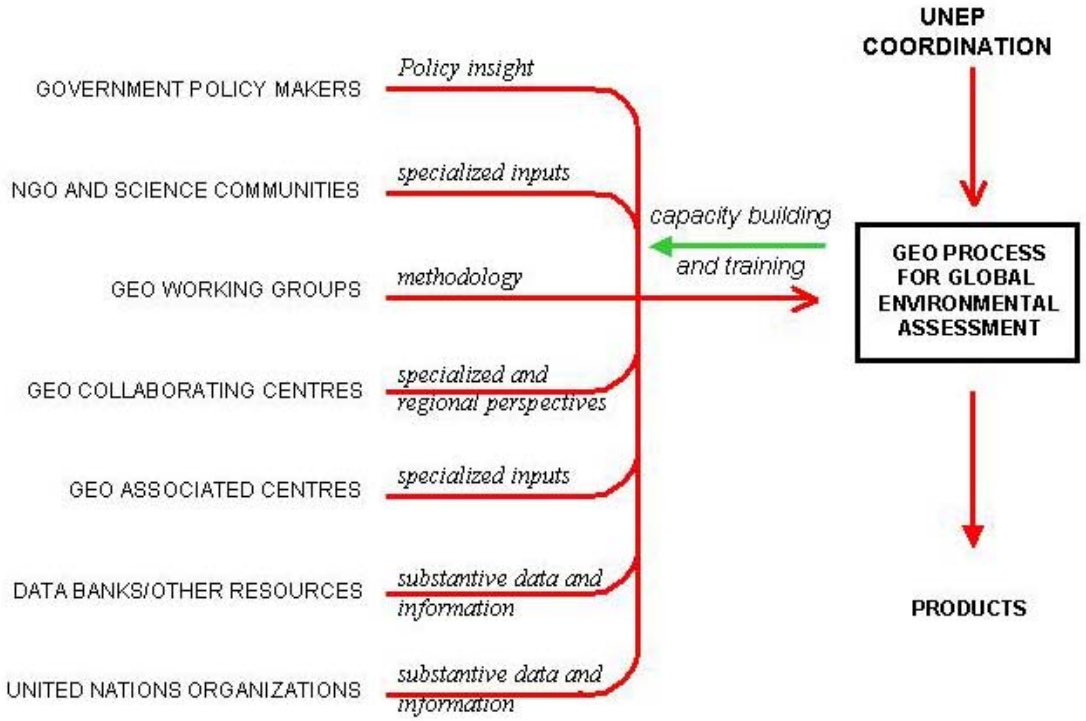



**Figure 3:** GEO's organizational structure (Cheatle 2001).

The GEO network is global, multi-level and polycentric, corresponding to the multi-scale nature of environmental and policy issues and the need for producing 'place-based' analysis following common protocols that make intra-scale comparison and cross-scale aggregation and dis-aggregation easier (though still not easy). The globally coordinated assessment process is replicated on lower scales. The purpose of sub-processes may vary

from identifying priority issues for analysis through gathering feedback on completed drafts in regional consultations with policy-makers, to more specialized activities such as capacity development for assessment and reporting on the regional, national, or even organizational level, thematic workshops, or the compilation of data sets and publication of data products.

The GEO process involves iterative interactions of GEO's participants within the global and regional scales and across scales with the purpose of identifying critical information needs of key audiences, gathering data, analysis, capacity development or communication of results. While the GEO network can be conceptualized as a hierarchy of organizations and products, it is more fitting to describe it as a hierarchy of semi-autonomous dynamic processes. Rather than being static and blindly repetitive, the cyclical GEO process offers – in principle – opportunities for learning and adaptation not only on the level of UNEP, but also on the level of individual CCs. In contrast with the static notion of hierarchy, Holling calls such systems panarchies, places where information is created and used and where adaptation can occur on multiple levels in multiple cycles (Holling 2001).

While GEO is an assessment and reporting system linking science and policy, it is probably closer to the policy side of the spectrum. Nevertheless, scientific quality control has been of concern and the GEO structure increasingly incorporates mechanisms to strengthen this aspect. There are mechanisms for both internal and external peer review, though in contrast with traditional science it may involve both scientists and policy-makers as the goal is both scientific credibility and policy relevance. More recently UNEP-DEWA has established an independent, high level scientific advisory panel led by a Nobel-laureate; creating a science-policy advisory group specifically for GEO is a possibility (D. MacDevette, UNEP-DEWA, personal communication).

Finally, integrated assessment and reporting systems are complex and in many ways imperfect, so they need to have built-in mechanisms for reflexive evaluation. GEO has been subject to an internal evaluation by an oversight unit of UNEP and an external evaluation (Attere 2000, Universalia 2001). UNEP has also gathered feedback from CCs

as part of the preparatory process for GEO-3 and devoted part of the GEO-3 start-up meeting to a discussion on strategic issues in light of the feedback received (UNEP 2000d). The challenge is to institutionalize feedback gathering, engaging CCs in a more substantive manner, and making sure there is a connection between findings of the evaluation and the next round of strategic planning and design.

## 3.3. GEO's companions

Being the flagship global assessment of the global environmental watchdog in the UN system, GEO is somewhat in a class of its own. However, on a general level, many of its aspects can be compared to other systems. Although my goal falls well short of systematic comparative analysis, where relevant, I will contrast lessons learned through the analysis of GEO with feedback on other systems.

In support of OECD's now published environment outlook (Organization for Economic Cooperation and Development 2001), the Dutch National Institute for Public Health and the Environment (Rijksinstituut voor Volksgezondheid en Milieu, RIVM) prepared an overview of nine 'review assessments' dealing with the state of knowledge on environmental issues and addressing futures aspects (Rijksinstituut voor Volksgezondheid en Milieu 1999). Although this is a relatively recent report, there are important and for my work, relevant initiatives that have surfaced since its publication (e.g., Millennium Ecosystem Assessment, the increasing attention paid to futures issues by the reports of the European Environment Agency). Four cases stand out for the purpose of comparison with GEO.

The first is the system of national assessments and outlooks in The Netherlands. The Dutch system is different in the sense that it focuses on the national scale and is comprised of three interrelated assessments: Nature Balance, Environmental Balance and National Environmental Outlooks (Rijksinstituut voor Volksgezondheid en Milieu 2000, 2000a, and 2000b). However, in terms of conceptualization and strategy it has strong

links to GEO. This is not surprising given the pioneering role of The Netherlands in IA, integrated planning, modeling and future oriented outlooks. The Netherlands and various Dutch institutions, particularly RIVM also have maintained long term and strategic interest and made a very significant mark on the GEO approach.

A second system from the supra-national scale that bears resemblance to GEO is the assessment and reporting program of the European Environment Agency (EEA). To date the EEA has published two continent-wide, integrated environmental studies known as the Dobriš Assessments plus two "Environment in the European Union" reports. The EEA also maintains a continent-wide multi-scale network of collaborating agencies actively engaged in the process. Similar to GEO, the EEA goes beyond SOE reporting in terms of following a multi-disciplinary and participatory approach and extending reporting to early warning, emerging issues and the study of policy relevant future scenarios (European Environment Agency 1999a and 2000).

Third, although it is primarily an economic organization, OECD has long maintained an environment directorate and an interest in environmental matters. The organization published its first environmental outlook in 2001, distinguished by a direct link to the policy process through OECD's Sustainable Development Strategy and a strong reliance on OECD's macroeconomic database, indicators, models and projections as a starting point for environmental analysis. The OECD outlook covers all of the member states, but it has been produced in-house by OECD staff and consultants (Organization for Economic Cooperation and Development 2001).

The fourth system is the Millennium Ecosystem Assessment (MEA). This effort is young; it started in early 2001. The MEA deserves particular attention because of the similarity of its goals, methods and structure to GEO. There are, however, also differences. In terms of its mandate the MEA is defined as a science assessment focusing on ecosystems covered by three relevant global conventions: the Convention on Biodiversity, Convention to Combat Desertification, and Convention on Wetlands (Millennium Ecosystem Assessment 2001, Millennium Ecosystem Assessment 2001a)[9].

Within the context of this mandate, the MEA looks at issues on multiple and interacting scales; operates with a polycentric network of institutions and individual researchers; includes capacity development to improve capability to carry out ecosystem assessment in the regions; and, broadly speaking, follows integrated assessment methodologies. Cross spatial-scale analysis of coupled environmental / socio-economic systems is accompanied by scenario-based assessments. Quite clearly – although not yet explicitly – the MEA is also intended to be an assessment system with the option to become iterative and ongoing in the future. In terms of institutional form, the MEA borrows rather heavily from the IPCC. While this means increased emphasis on scientific independence, it does not necessarily mean proportional representation. The developed world, particularly North America weighed heavily in the establishment of the MEA and continues to play a controlling role – a challenge to its legitimacy that needs to be addressed as the assessment matures.

# 4. Methods

The analytic framework for my work builds on the results of the GEA project described below and draws on other literature on research utilization. The framework developed in GEA has been created to address the question of effectiveness and design related to science assessments (**Figure 4**). It conceptualizes assessments as research and communication systems influencing decision-makers and the evolution of issue domains, broadly defined (Global Environmental Assessment 1997). According to this view, design elements chosen by producers impact the effectiveness of assessments where impact is measured indirectly by looking at elements of assessment design correlated with criteria of effectiveness. Based on the relevant literature and the analysis of several well-known assessment systems, the GEA project identified saliency, legitimacy and credibility as essential criteria of effectiveness.

Eckley *et al.* (2001) described the issue domain as the entire sphere of social activity surrounding a particular issue, such as climate change or the long range transport of air pollutants. An issue domain includes its own participants, and one or more institutions. It also includes the behaviours and decisions of these actors that are often influenced by their interests, beliefs, strategies and resources, and the impacts of those decisions on a particular aspect of the environment, for instance acid deposition. As the above authors further put it: "Assessments (which themselves have distinct characteristics, many of which correspond to 'design choices'), influence the changes over time in the issue domain, typically by changing the beliefs of actors, or sometimes by identifying new interests. The main purpose of this framework is that it can be used to explain how an issue domain changes over time."

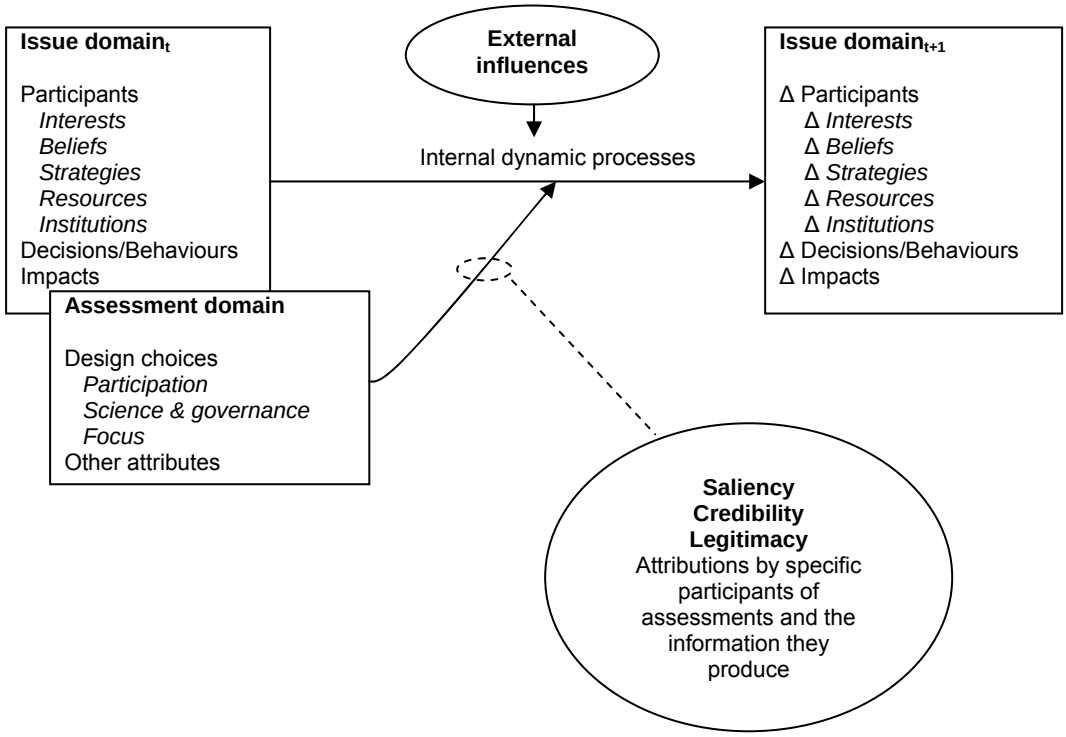



**Figure 4:** The GEA framework for understanding how assessments affect issue domains from time period *t* to *t+1* (Eckley *et al.* 2001).

Although not explicit in **Figure 1**, there is a distinction between producers responsible for assessment and recipients with decision-making authority over issue areas in question. There is also symbolic recognition of external influences that affect the decision-process. GEO and many of the systems that I classify to be integrated assessment and reporting (IAR) systems, including many SOE reports are not science assessments *per se*, but the GEA analytic framework still applies, with some modifications. First, GEO covers a very broad range of issue domains and addresses a large diversity of audiences. While one could look at GEO's influence over a particular issue domain, say climate change, it would be easy to find other thematic assessments that have more direct effect. What is different and to some degree unique in GEO, is the parallel treatment of issue domains, often emphasizing interactions that cut across not only issues, but also scales and the interests of a wide range of policy audiences.

In addition to saliency / relevance, credibility and legitimacy I also accepted awareness as an additional criterion particularly important for GEO for reasons that are explained later in the analysis. The GEA framework identified a range of design issues with implications for these criteria. Keeping in mind the attributes and goals of GEO and similar systems, I examine them in the context of several design elements. Some of these, such as participation, governance or focus are addressed in detail by studies carried out in the context of the GEA program. Others, such as communication, data and indicator systems and products involve choices that are particularly important for GEO and similar initiatives. I will justify selecting these design issues partly based on arguments found in the theoretical and applied literature, my own working experience with GEO and what I heard from interviewees during my research. I am not arguing that all design elements or assessment criteria apply the same way or with the same weight to all assessments and issue domains, but I do argue that they all have sufficient general validity to be part of the analytic toolkit with which a wide range of assessment and reporting systems can be studied.

In contrast with most thematic science assessments studied in the GEA program, overviews like GEO must, by definition cover and try to influence a large number of

issue domains. I added this aspect by showing several overlapping issue domains marked with *x,y,z(t)* and *x,y,z(t+n)* where *x*, *y*, and *z* stand for a given issue domain, *t* for a base year prior to launching an assessment and *t+n* a future date following the publication of the results of the assessment when its impacts can be already measured. As in the original GEA framework, the delta (Δ) in front of issue domain attributes in the *t+n* time period indicate that some or all of these attributes, and consequently the issue domain, has changed. Whether the change is due to the assessment and reporting system – in this case GEO – in question is a matter for further analysis. The modified conceptual framework is shown in **Figure 5**.

Although already rather complex, the diagram hides additional complexity arising from the fact that many assessment and reporting systems, including GEO, study phenomena that cut across multiple scales. The attributes of a particular issue domain can be different on one spatial and temporal scale than another. Assessment systems may match this complexity through a nested structure of assessment participants, processes and products. A global assessment system is likely to have different actors, processes and products in Latin America and the Caribbean than in Central and Eastern Europe or, further down the scale, Canada, yet it needs to have coordination and governance mechanisms that further aggregation, cross-scale comparison and harmonization if it is to say something on common issues and interests. Although many of the scale-related differences can be justified based on cultural, political, ecological or other reasons, they do not change the fact that scale is a cross-cutting issue that has implications for most assessment design elements and criteria of effectiveness.

Primary data for this study were gathered through narrative interviews. The interviews were conducted over a period of five months between November 2000 and April 2001 (**Appendix 2**). My 27 interviewees can be characterized either as senior level managers and producers of GEO, representatives of organizations not necessarily affiliated with GEO but producing reports with similar ambitions, or individuals representing key audiences for GEO type products.

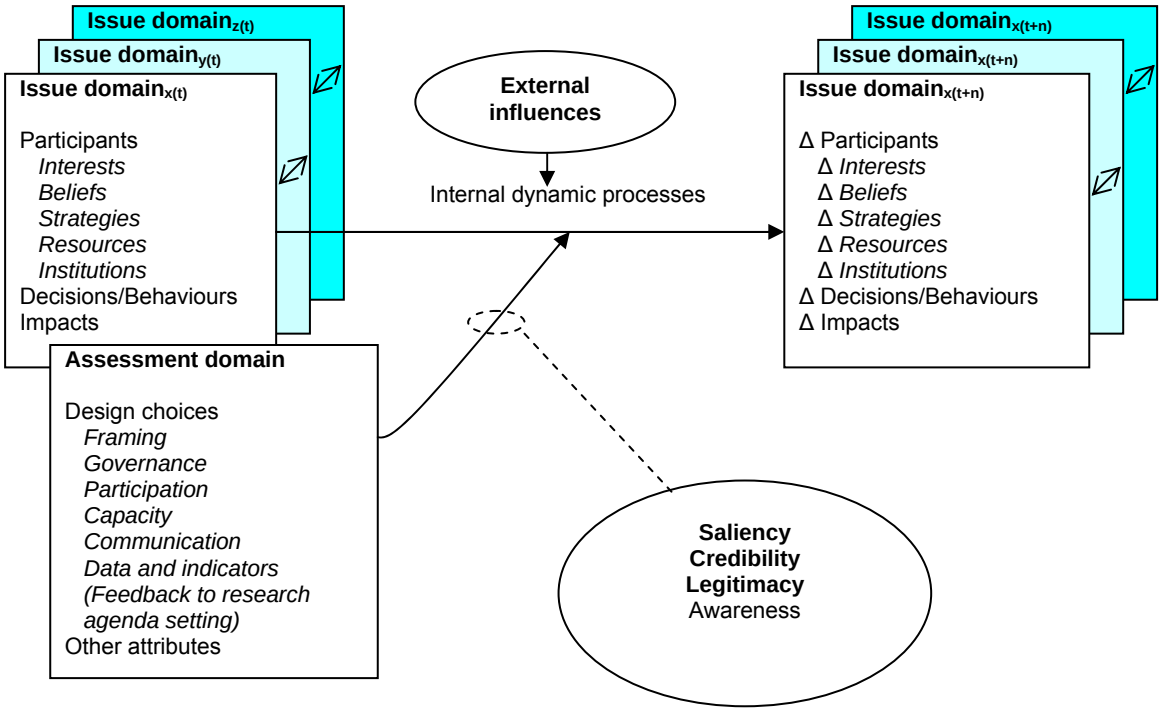



**Figure 5:** Conceptual framework for this study modified from Eckley *et al.* 2001.

I conducted the interviews using an interview protocol, which was linked to my conceptual framework and covered a small set of design elements (Appendix 3). My interview strategy was to cover key points included in my framework where a given interviewee had particular interest and expertise, while at the same time allowing exploration in new directions. As a result, in the time available for the conversations, we could not always cover in equal depth all of the design issues included in my framework. On the other hand, often we covered topics not on my list. Unevenness of raw data would have reduced the value of my data for statistical analysis, however, my intent does not include inference because of the non-representativeness of my interviewee set.

Each interview resulted in a 2-6,000 word transcript. In three cases, I lost some data due to recording equipment failure. In these cases, I had to prepare transcripts based on my recollection of key points from the interview. All other transcripts were used verbatim. Having transcribed the interviews I imported their text into NUD*IST 4.0, a software program which has a series of advanced qualitative data analysis functions based on the coding of text based data.

Besides interview data and literature, I drew on the results of the GEO user survey and evaluation; the first of those was carried out by Universalia Inc., and the second was an in-house evaluation by an oversight unit of UNEP itself (Attere 2000; Universalia 2001). The user survey relied primarily on feedback from GEO's audiences, particularly governments, GEO CCs and the GEO team at UNEP and at UNEP regional offices. Although I do not claim to have followed systematic triangulation throughout the entire research, I sought to find support for my findings from the interviews and the literature in the evaluation report and user study.

# 5. Results

This research has relied on information from the literature, the results of the GEO user survey, and the results of a set of narrative interviews. In addition, I drew on personal communication with various colleagues associated with GEO and my notes from GEO meetings over the last several years. Many of these predate the conception of this project. Out of these information sources, interviews are my main source of primary data. Beyond preparing and importing transcripts as described in Chapter 2, analysis required the coding of the resulting text to identify key variables and correlations. In order to do that, I had to create a system of codes for key ideas - in technical terms nodes – assigned to respective sections in the text. This resulted in:

- a system of codes developed on the basis of the analytic framework and the information obtained from the interviews;
- a coding pattern for each interview;
- a mosaic of contributions to each node from interviewees.

My coding structure evolved through several iterations that included a set of design elements and a set of free nodes that included topics that I found important for my analysis, but that I did not consider as being part of the design element set. Screen shots

for the entire coding tree and the design elements branch, including its place in the overall coding system are shown on **Figures 6** and **7** and an example for an interview coding pattern on **Figure 8**.

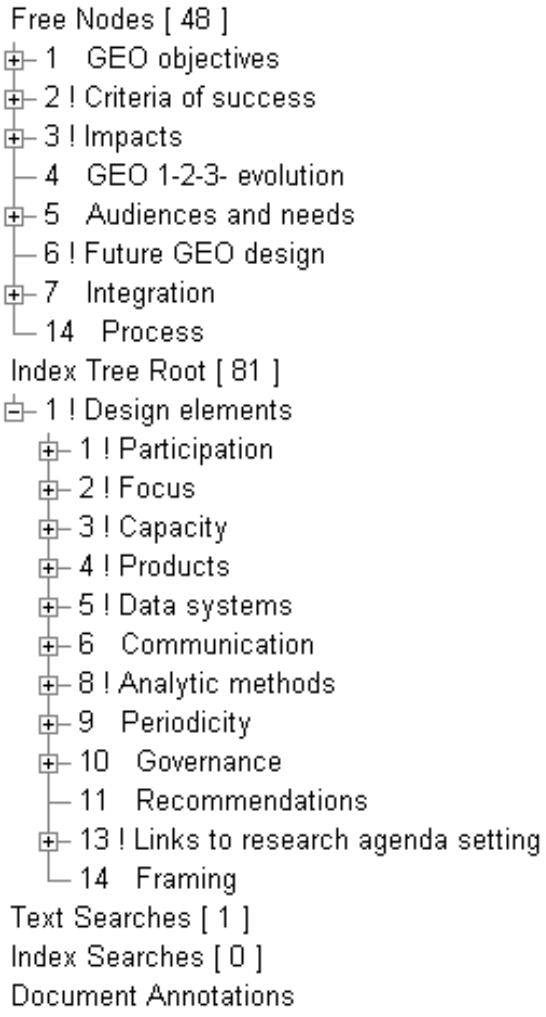



**Figure 6:** System of codes used for analyzing interview data.

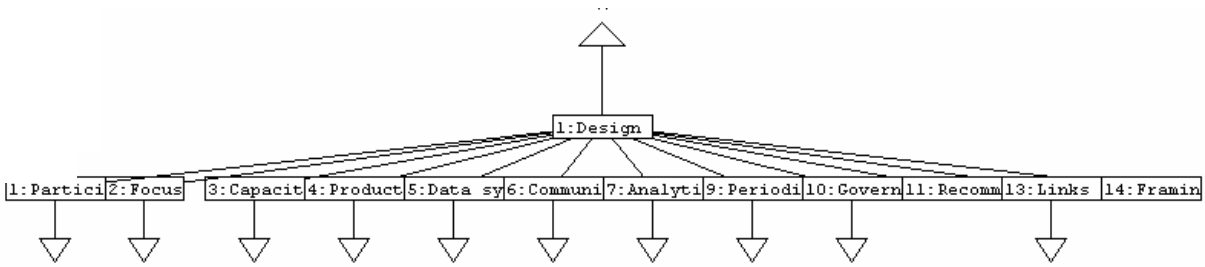



**Figure 7:** Coding tree for the design elements branch.

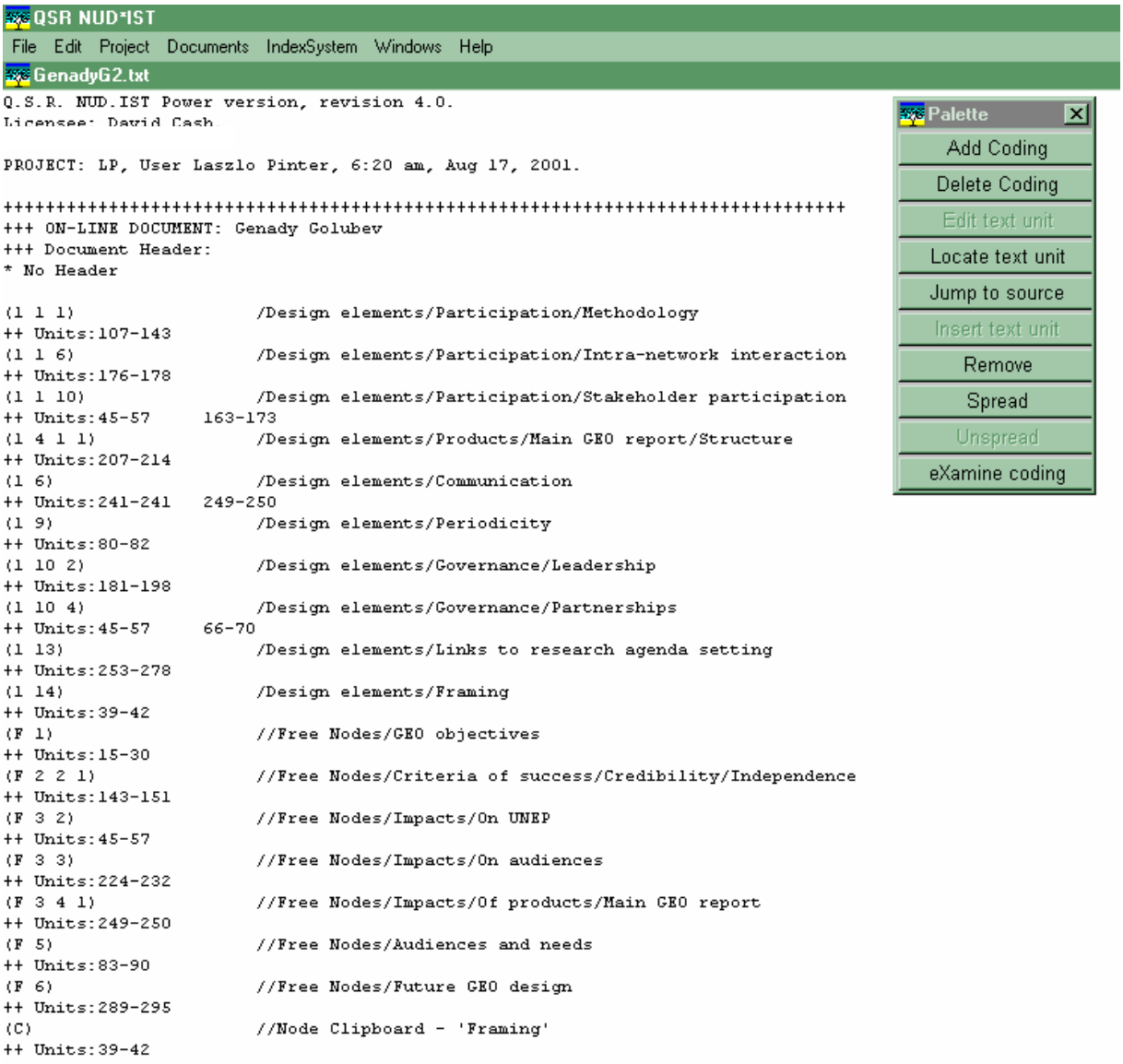



**Figure 8:** Screen shot example from NUD*IST 4.0 to illustrate interview coding pattern.

My initial design element set, borrowed partly from the GEA framework and other literature included participation, focus, governance, and links to research agenda setting. All of the others were added as the analytic framework and with that, the coding structure evolved. I anticipated some of the additions. For instance, even if it was not central in my interview protocol, I expected data and indicator issues to feature prominently in the interviews, whether I asked pointed questions or not. The same was true for communication and capacity, they were both of central concern to many. There were

other design issues I did not anticipate would receive as much attention as they did, namely periodicity and recommendations. Periodicity probably received as much attention as it did because there is ongoing discussion about the frequency of the GEO reports, so the issue is on people's mind. Similarly, policy focused recommendations are increasingly sought by key audiences. While some science assessments can distance themselves from the active domain of decision-making and continue to study and observe, policymakers seem to be increasingly demanding advice that is more directly useful in making decisions. As one of my interviewees pointed out, high level decision-makers expect more action orientation and would like to see assessment systems like GEO advise what action to take, in the form of legislation, in the form of economic instruments, in the form of preventive measures so that they do not make things worse (S. Shresta interview).

Having gone through several iterations of the design element tree, I finally chose 12 elements (shown under the 'design elements' branch of my coding structure, on **Figure 6**). As typical for qualitative analysis processes, my raw list at the end of the first iteration looked rather different and evolved through several more iterations. It was quite clear that my detailed analysis could not include all the twelve design elements without further risking losing analytic depth. As a result, I had to exclude from detailed analysis a number of design elements, or merge them with others, at least for the purposes of analysis. I made these decisions based on the importance attributed to a design element in the literature, by interviewees and by the GEO initiative.

At the end, my original list of twelve design elements was reduced to the following seven. These are analyzed and presented in seven policy papers in Chapter 6:

- Participation
- Capacity
- Data
- Communication
- Governance

- Framing
- Links to research agenda setting.

I deal with these design elements in significant detail in the following chapter, so here I discuss the reasons for leaving out others. **Table 3** lists the design elements at the end of the last iteration of coding and re-coding interview text. This could not be evolved further, but I found it sufficiently developed for the purpose of my analysis. The table also shows the contribution of individual interviewees to the discussion on various design elements according to my coding tree.

With regard to focus, I found I had relatively little interview data, but more importantly I found that many comments that I coded under framing also applied to focus. An assessment and reporting system framed around environment and development is by definition broadly focused, with all its accompanying complexities.

Products are by definition an important design element, but I found that I could also discuss key product related issues under other sections, such as participation (e.g., with regard to the role and responsibility of various network members in preparing specific outputs) or communication (e.g., the importance of product design or the phasing of product release).

Analytic methods are central to any assessment's credibility. I have written with others on methodological issues, including for instance, how to do the SOE component of an integrated report, how to construct indicators or how to do policy analysis in an integrated framework (Pintér, Zahedi and Cressman 2000). However, my interest in this work was primarily in the institutional and strategic aspects of assessment and reporting systems. Rather than dealing with the mechanics of policy assessment, I am more interested here in who decides what methods to use or to what decisions these methods should apply. Therefore, methodological issues are discussed in the context of other design elements.

**Table 3:** Interviewee contribution to the discussion of design elements.

| Interviewee | Part. | Focus | Capacity | Products | Data | Communication | Analytic methods | Periodicity | Governance | Recomm. | Links to Res. Agenda | Framing |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J. A. de Larderel | ✔ | | | ✔ | ✔ | | ✔ | | | | | |
| J. Bakkes | ✔ | ✔ | ✔ | ✔ | | | ✔ | | | ✔ | | ✔ |
| S. Barker* | ✔ | | | ✔ | | | | | ✔ | | | |
| L. Braat | | | | ✔ | | | ✔ | ✔ | ✔ | | | |
| F. Carden | ✔ | | ✔ | | ✔ | | | | | | | ✔ |
| M. Cheatle | ✔ | | | ✔ | | | ✔ | ✔ | ✔ | ✔ | | |
| M. Chenje | ✔ | | ✔ | | | ✔ | | | ✔ | | | |
| A. Dahl* | ✔ | | | ✔ | ✔ | | | | ✔ | | | |
| O. El-Kholy | ✔ | ✔ | ✔ | | ✔ | | ✔ | | | | | |
| L. Flanders | ✔ | ✔ | ✔ | | ✔ | | ✔ | | ✔ | ✔ | | ✔ |
| G. Golubev | ✔ | | | ✔ | | ✔ | | ✔ | ✔ | ✔ | ✔ | ✔ |
| E. G.-Espeleta | ✔ | | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | | | |
| A. Hammond | ✔ | | ✔ | ✔ | | | ✔ | | ✔ | ✔ | | |
| M. Hanssler* | ✔ | | | | | ✔ | | | | | | |
| R. Jolly | ✔ | | ✔ | | | | ✔ | | | | | ✔ |
| J. Loh | ✔ | | | | ✔ | | ✔ | | | ✔ | | |
| S. Luan | ✔ | | | | ✔ | | | | | | | |
| R. Maas | ✔ | | ✔ | | | | | | ✔ | | | ✔ |
| N. MacPherson | ✔ | ✔ | | | | | ✔ | | ✔ | ✔ | | ✔ |
| R. Mnatsakanian | ✔ | | ✔ | ✔ | | | | | ✔ | ✔ | ✔ | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L. Mortensen | ✔ | | | | ✔ | ✔ | | ✔ | | | | |
| A. Rahman | ✔ | ✔ | ✔ | ✔ | | ✔ | ✔ | | ✔ | ✔ | | |
| P. Raskin | ✔ | | ✔ | ✔ | | | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| W. Reid | ✔ | | ✔ | | ✔ | | | | | | ✔ | |
| S. Shresta | ✔ | | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | | ✔ | ✔ |
| R. Uhel | ✔ | | | | ✔ | | ✔ | ✔ | | ✔ | | |
| V. Vandeveerd | ✔ | | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ | | ✔ | ✔ | ✔ |

* incomplete transcript because of problem with recorder

*Shaded areas show design elements separately included in the final analytic structure.*

Periodicity and the phasing of various assessment products has many implications, for instance, for governance, participation and communication; it is discussed under those sections. Finally, I chose to address recommendations under each section, rather than addressing them in an isolated fashion.

# 6. Design choices and tradeoffs

Understanding the purpose of assessment and reporting systems is a starting point both for their design and analysis. Their purpose, increasing the awareness and understanding of environmental issues and their interactions with human socio-economic development is seemingly straightforward. It is also clear that the ultimate purpose of IAR is improving both human and environmental conditions. A closer examination, however, reveals that the parties associated in different roles with assessment and reporting systems in general and distributed systems in particular may see several purposes that they rank in different ways. These views seem to be influenced by whether one is associated with a coordinating center, a participant organization on the periphery or an outside agency; whether one represents a developed or developing country; or whether one is a scientist or policy type. Serving a variety of purposes may be the norm in the case of programs with several parallel streams of products and activities like GEO. On the other hand, diversity may also create challenges if some of the real or assumed purposes are potentially conflicting and motivate participants to press in opposite directions, reducing cohesion. Assessments need some degree of centrality of purpose(s).

Besides producing periodic reports on the global environment for broad public consumption, GEO has other associated purposes and objectives, expressed or implied. These may include raising UNEP's profile, as observed by some of my interviewees: "I think the overall morale of UNEP is better having this product" (A. Rahman, BCAS,

interview). Additional purpose and objectives include building assessment capacity in developing countries, testing and introducing an assessment and reporting approach that better reflects sustainability concerns, and setting in motion processes to improve the quality and utility of regional or global datasets. Emphasis on these may shift according to the phase of the assessment, the perspective of a given participant or the needs of a geographic region. Although the production of the global report provides an overall context, there are other activities in GEO that often take place in parallel and respond to the needs and interests of particular audiences. For instance, developing country Collaborating Centres in the GEO network contribute to the report, but some of them have also been audiences of capacity development activities.

On the broadest level, there is a view that considers GEO as part of a learning process in the context of society's scientific and policy struggle to come to terms with the nature, magnitude and complexities of global change. This is seen as a challenge for UNEP, as an organization that has a structure for, and history of dealing more with definable, discrete environmental problems, but is increasingly drawn into more integrated, holistic, and politically connected analyses:

> "I see GEO as part of the learning process of UNEP and what assessment needs, but also in the world at large driven by really grand changes and global predicament, no less than that. So my criticisms of GEO in all that are quite tempered by a recognition that I see this as a beginning of a kind of groping, a way of addressing these very difficult planetary challenges."
> (P. Raskin, Tellus Institute, interview)

This supports the view and purpose of GEO as a site for learning and experimentation, not only an analysis of where we are but also "where we might be going, it should be where you would want to go" (Ibid.). These attributes weighed heavily and consciously into the design of GEO and led to the current system, as Bakkes (1999) observed:

> "Our current achievements can be best compared to the plans for the GEO process – as it was named eventually – formulated in 1994/95. That process started off to do better than the not-so-interesting decadal SOE reports of UNEP. After ample thought, we discarded the idea of pilot GEOs and embarked on a learning-by-doing journey. That meant a choice to produce imperfect and incomplete GEOs every two years, working towards a good network, methodology and instrumentation by 2002, three decades after UNEP's start. Now that we will start preparing the 2002 GEO, that original goal is still a useful yardstick."

Given the uncertainties and the limitations of knowledge with regard to policy-relevant, future oriented integrated analysis on multiple, connected scales, experimentation is not simply desirable, it is unavoidable. While experimentation is potentially rewarding, it is also risky. The results can have significant implications for the image of the assessment institutions, in both positive and negative directions. If the assessment is done well, it can become an asset and a tool of building or rebuilding recognition for the organization. There is a perception that GEO has contributed to a stronger UNEP that was seen as struggling in the early 1990s and had no signature product on the global environment to match other global reports, such as the World Bank reports on economic issues, UNDP's Human Development Report or even some products of NGOs, such as WRI's World Resources Report series. The risk is that a weak system could potentially become a liability and be seen as a waste of resources. In order to avoid disappointment, one must bear clearly in mind the aspects of the assessment and reporting system that are experimental (e.g., the approach to constructing integrated future scenarios), and the aspects that are routine and proven (e.g., reporting of basic environmental indicators). If we recognize and admit mistakes, they are far more likely to be accepted and even seen as a strength, particularly in the growing phase of an assessment – I assume GEO is still in that phase – than self-deception induced by ignoring and hiding them.

Others claim that GEO's most important aspect may be its engagement process, particularly the engagement of developing countries, supporting GEO's role as a site of learning, experimentation and capacity development (F.Carden, IDRC; N. MacPherson,

IUCN; A. Hammond, WRI, interviews). According to this view, the GEO products themselves are of lesser importance. In fact, it is argued that given the resources, staff and technical capacity and level of political support within and outside UNEP, GEO products cannot compete with others such as the World Resources Report in terms of content and quality[10]. Determining whether this is the case or not is not the task of this research; it is worth noting, however that GEO has been evolving. GEO-2000 has been generally recognized at UNEP's 21st Governing Council session as a major improvement over GEO-1, and there are signs that GEO-3 is building on and going beyond GEO-2000 in several ways (UNEP 2001a). This suggests that the system is capable of learning and during its first three cycles, seems to be improving its process and products. Whether learning is or can become an inherent design characteristic in GEO is a question that may take more reporting cycles to convincingly answer based on outcomes. However, even today, one can look for evaluation and feedback mechanisms that are the preconditions for institutional learning.

Separation of product and process in GEO cannot be anything but artificial. The system has important audiences which view providing information on relevant policy issues and agendas as a key service, particularly so because the information comes from a process where local organizations are given the opportunity to build their region's own perspectives into an internationally recognized and sanctioned report. Regional level information seems to be particularly important:

> "… ministers from the more developed regions of the world don't look at GEO for national level guidance on policies, but they do greatly appreciate the overview of the regional overviews and the thematic overviews. This was said to me several times, and in the future they would like to see those kept, because it gives them a very concise overview from what they consider to be an authoritative and credible source."
> (M. Cheatle, UNEP, interview)

Overall, I find two sets of design questions with regard to purpose particularly relevant: niche and feasibility. Where is the niche of needs the particular assessment and reporting system wants to fill, what are its demonstrated or potential advantages in filling this niche when compared with other systems, and of course, to what degree are these potentials fulfilled? The second series of questions is related to the feasibility of organizations to effectively manage complexity arising from multiple parallel purposes assigned either formally or informally to assessment and reporting systems.

Due to globalization and the increasing scale of human activities, local environmental problems are increasingly viewed as interconnected and having planetary consequences. This created a niche for regionally differentiated, integrated environmental assessment and reporting on the global scale. Filling this niche requires an approach that articulates its purposes both as processes, aimed at creating scale, theme and region-specific content, and products communicating findings to key audiences in all relevant contexts. GEO, correctly, does that, even if there is little doubt it can do better. Whether the program can handle the complexity arising from multiple and dynamically shifting purposes is a continuous management and design challenge involving tradeoffs.

The following sections will look at a design attribute or element each. I selected these elements because of indications that the way they are treated in an assessment can be correlated with effectiveness, or more specifically the selected criteria of effectiveness. The indications came partly from the literature, but also from information gathered in the course of this research. I will look in detail at the following design elements:

framing;
governance;
participation;
capacity;
communication;
data and indicators; and,
feedback to research agenda setting.

The list is not exhaustive, but as I will show, it offers more than enough detail to analyze the selected criteria of effectiveness. I consider scale and cross-scale connections a cross-cutting issue that assessments need to respond through several of the listed design elements. So scale issues are discussed not in a separate section but woven into several. Each section dealing with a particular design element includes three subsections: one reviewing the design element's importance and justifying its selection for the analysis of IARs in general; second, the analysis of GEO-like assessment systems from the perspective of the given design element; and third, the implications of the way the particular design element is treated for effectiveness.

## 6.1. Framing

### *6.1.1. Why framing?*

Framing has been shown to play an important role in determining the effectiveness of science assessments (Clark *et al.* 2002, Schön and Rein 1994). For the purposes of this study, I define framing as the application of one's worldview to the conceptualization, analysis and presentation of environment / development problems. By analogy, I compare framing to setting the terms of reference in principled negotiation, reaching tacit or explicit agreement on perspectives acceptable to all involved (Fisher and Ury 1983). By framing issues one way versus another, certain positions and their advocates by definition may be excluded, while other positions can become legitimate almost by default. Framing is a subtle but very effective tool to exert power, to control agendas and thus to control the range of acceptable outcomes well before analysis could even start in earnest. There are always tradeoffs, as narrow framing may exclude relevant variables and thus preclude truly integrated solutions, while broad framing may lead to complexities with which the assessment system may not be prepared to deal. Framing is not static: both people and to a lesser degree organizations are capable of major frame- or paradigm-shifts in response to a wide range of forces ranging from political changes to new scientific knowledge or extreme natural or social events. Reframing brings risks and opportunities, analogous in the social and cognitive domain to the opportunities arising

from the reorganization of ecosystems after disturbance, as per the classic model of Holling (1978). It is commonly argued that sustainable development emerged as a potential paradigm shift or reframing on a grand social scale, even if did not lead, at least not yet, to a Kuhnian paradigm shift in science (Kuhn 1962).

Self-reflection and a consequent ability to adjust the frame of analysis if necessary may be important to ensure that an assessment responds to new knowledge or emerging needs and perceptions of its audiences. However, constant frame-shifts in response to a real or perceived pressure to constantly produce new ideas can also become a problem (Pintér, unpublished data). It is necessary to understand when reframing is necessary and when changing the questions, methods or some other details of the analysis without changing its framework could be part of the answer. GEO has maintained some elements of its original structure, such as the core of the pressure-state-response causal framework or the coupling of SOE, policy analysis and scenario sections. However, some aspects have changed, as GEO has evolved (e.g., inclusion of a vulnerability analysis section in GEO-3 or the amalgamation of the SOE and policy sections). None of those changes would qualify as major frame shifts, but the evolution is clearly noticeable.

Framing is a particularly important question in multi-scale, distributed assessment systems that both need to ensure some commonality of analytic frameworks to allow cross-scale aggregation where needed, and must permit some flexibility so that the assessments that are reporting on fine-scale issues (e.g., biodiversity hotspots in Costa Rica or the same questions in Gabon) can fine-tune their analysis according to the local context. If fine-scale frameworks are unique, analysts will experience problems with coarse-scale aggregation and cross-scale analyses; on the other hand, forcing use of the same framework in the same way by all centers in the network reduces local saliency.

At the heart of the framing issue is that even a narrowly construed environmental problem can have many alternative framings and analyses and, consequently, many solutions or attempts at solutions. During the early days, environmental problems were considered mainly as technical matters in industrialized countries, problems that

produced debates and solutions predominantly on technical terms. During the 1990's and particularly after the publication of Our Common Future and Agenda 21, a broader, more integrated framing of the environment became accepted. GEO fits into and builds on this more integrated tradition.

### *6.1.2. Framing in GEO and distributed systems*

GEO at its core is a system dealing with sustainable development from an environmental perspective. Implementing a very broad and integrated analytic framework is GEO's key challenge, yet it is also its very niche. As has been pointed out, one can go to the Intergovernmental Panel on Climate Change for integrated information on climate change, soon to the Millennium Ecosystem Assessment for information on biodiversity, wetlands, or desertification, or to the United Nations Development Programme's Human Development Report for information on poverty or human health trends. A fairly comprehensive view of global environmental trends is available from the World Resources Institute's World Resources Report or for those with a stronger heart from the State of the World series of the Worldwatch Institute. However, these are the products of small teams of professionals who may be the best experts in their field, but who are possibly not as close to ecological and policy realities in the regions as many CCs involved in the production of GEO. An increasing number of global science assessments follow integrated assessment methodologies from the perspective of their particular issue. As difficult as these assessments are, however, they are far less challenging than pulling together numerous sectoral and regional pieces. Integration and pulling together has a value by itself, as both the public and policymakers increasingly grapple with issues that defy conceptual and geographic boundaries traditionally followed by focused assessment systems (A. Rahman, BCAS, interview).

From the beginning GEO has been positioned not as a report "on the butterflies and the flowers", but as an authoritative volume on the critical policy issues and decisions that influence the environment (V. Vandeveerd, UNEP-GPA, interview). Very often, decisions of that scope are made not by environment departments, but agencies dealing with trade, technology, investment or other matters and they are made with no or only

token reference to the environment. As described in Chapter 3, GEO uses the DPSIR as its underlying analytic framework. The DPSIR framework is inherently complex, and the GEO approach adds further challenges, namely the need to look at issues across spatial scale, and to link past trends with policy paths and scenarios in the future. The participatory GEO process is yet an additional framing element.

This broad framing and its implementation in GEO have been contested at times. Those critiquing this approach included even some countries on the Governing Council, but also other UN organizations that would like to see UNEP more strictly adhering to its environmental mandate rather than going in a direction where it could be seen as infringing on the territory of others whose responsibility is human development, economics or sustainable development (Pintér, unpublished data). This challenge is commonly faced by broadly framed assessments, as they are more likely to intrude on other interests. Consequently, broad integrated assessments need to find mechanisms to prevent and manage conflict arising from overlaps. Even if conflict is hard to avoid, it only reflects an underlying reality of interconnected issues and decision problems.

> "I think ever since we combined environment and development this was bound to happen. Eventually I think we will have to face up to it and we have to find a way of addressing the issue in the context of sustainable development. … I think this is bound to happen and we need to find a way of doing it." (O. El-Kholy, AGU, interview)

Our reaction to that challenge should not be to retreat to the narrow view of the environment or 'box in' the broader assessment, but rather to build alliances with those responsible for adjoining issue areas, including human health, economic policy and so on. GEO has taken important steps in this direction by inviting representatives of global organizations responsible for these areas to contribute to the GEO process.

Beyond the environment / development dichotomy, another important framing issue in GEO involves presenting environmental issues both in a spatial, 'localised' context and

in a thematic context. “Local’ in this context implies a UNEP region or subregion, where the logic of the analysis remains the same as in community or ecosystem scale place-based assessments (i.e., looking at the entirety of multiple ecological and socio-economic causes and outcomes as they interact in a spatial unit on the landscape) (Wilbanks and Kates 1997). Thematic analysis draws together environment / development issues related to particular sectors, such as forestry or fisheries management, across all regions of the world.

### *6.1.3. Implications for effectiveness*

Framing can be an explicit part of the assessment and reporting process or at least part of the dialogue. Clearly, one cannot open framing issues at every step, but at some critical points in the process framing issues can and probably must be asked in order to learn if participants still have a consensus and a common platform, or adjustments are needed. This is easier if framing is formalized as a concrete framework, preferably in the early phase of the assessment and preferably with the consensus and input of key participants and audiences.

Integration of ecological and socio-economic aspects provides an overall framework in GEO, but there are of course sub-products (and corresponding sub-processes) which deal with narrower issues in greater detail. The underlying assumption is that among GEO’s audiences, some will be interested in, say, a sectoral angle such as the status of global forests, while someone else dealing with Northwestern Ontario, Canada may want to know something about forests as they interact with climate change and land-use in that part of the world. Yet another reader may have interests strictly in the application of economic instruments to environmental problems. Clearly there are many ways to dissect and integrate issues, and some of GEO’s framings may contribute more to effectiveness than others, depending on who is asking the question.

While the dual thematic and spatial framing can be justified by the needs of different audiences, implications for the organization of the assessment processes and products are significant. In principle, the same information, say water shortage in the Carpathian Basin

would appear in two contexts, first in the context of Central European issues where it is presented and analyzed along with other regional matters (or regional manifestations of global problems) such as climate change, land-use or agriculture. Second, it would be covered in thematic sections dealing with the global water cycle, where water issues in the Carpathian Basin may or may not be featured, depending on the severity of the water problem as compared with other regions of the world. The implication for governance is that the assessment needs both an organizational structure and a report structure that takes the dual thematic / spatial framing into account.

One way to assess the saliency of the overall GEO framework is to find out if there is interest to emulate it at finer spatial scales. Successive assessments predicated on its framework has been also identified as a measure of success in the Millennium Ecosyste Assessment (MEA 2001a). According to this measure, GEO scored quite well. The GEO user survey conducted by Universalia found that "the most immediate identifiable impact of GEO-2000 to regional and national policy making is the growing adoption of the GEO methodology by regional governmental forums and national governments for the production and/or improvement of their state of the environment reporting" (Universalia 2001). While some elements of the GEO method are gaining acceptance even without GEO *per se*, there are many cases where the connection is obvious. Besides the global GEO process there were other, regional GEOs for instance for small island states (UNEP 1999, UNEP 1999a). These are mostly initiated and supported by UNEP, but also are supported by regional organizations of small-island developing states. However, in some cases, national governments have decided to adopt the GEO framework to be used in their national reporting. The most striking example is the Latin American and Caribbean region, where several countries, including Panama, Honduras, Nicaragua, Chile and Cuba have adopted or are planning to adopt the GEO framework in their national reporting. In addition, at their summit in February 2001, the regional Ministers of the Environment meeting in Barbados adopted a declaration expressing their full support for the adoption of the GEO approach in their region at the national scale (Gutierrez-Espeleta, Observatorio del Dessarrollo, interview). The Libreville and Douala Declarations for the Central African region (Universalia 2001) have expressed support for the regional, sub-

regional and national adoption of the GEO methodology. In other cases, for instance in the People's Republic of China, governments looked at the GEO methodology, adopted and integrated particular aspects into their own SOE reporting and then called it their new approach. This can occur even below the national level: the Chinese State Environment Protection Agency (SEPA) organized a capacity development program to enable provincial environment agencies adopt a modified version of the GEO approach (S. Luan, SEPA, interview). Also, GEO has been one of the assessment systems looked at as an example when OECD set up its own outlook system (L. Mortensen, OECD, interview; RIVM 1999). GEO is not only a source of influence, but also an assessment to be influenced and improved. The assessments carried out by the European Environment Agency are among the possible sources of such influence (R. Uhel, EEA, interview). These are solid proofs of impact and confidence in the overall GEO framework and its ability to produce useful information. Adopting elements or the entire framework on national or sub-national scales not only shows the framework's cross-scale applicability, indirectly it also helps the global process because general aspects of the framework will be the same on the national, regional and global level.

| Recommendation |
| --- |
| Global assessment systems should examine and refine their frameworks from the perspective of cross-scale linkages and the needs of lower scale applications. |

With regard to salience and framing, an analysis that is not mandated to go into national details and comparisons, such as GEO, is by definition less relevant for policy-makers concerned with national issues, strictly speaking (R. Mnatsakanian, CEU, interview). This has been also observed in the case of EEA reporting. "The report would have been more useful and attracted greater media/public attention if there had been more consistent information on national circumstances" (Ferguson 2000). Of course, for GEO, going to the national level is not a real design option, although it can be still useful and salient when people look at environment / development issues their country shares with its neighbours in the region. In those latter cases, GEO could and does provide useful information and is seen as salient as well as legitimate (Attere 2000).

Recommendation

Even when doing coarse scale analysis, assessment could increase relevance for local audiences if it researches and correctly detects their interests in transboundary or cross-scale issues.

With respect to scientific credibility, a broad framework is inherently more of a problem, at least from the perspective of traditional disciplinary science. Part of the challenge is that because of time and resource constraints it is difficult to organize a quality control process that involves a wide range of disciplines. Also, in a distributed system there are always members with less than adequate capacity in one area, and more in another. As a result, the weak and strong points of the assessment would differ depending on the member of the network from which it came. However, one must keep in mind that GEO and other outlooks are not necessarily traditional science assessments. They identify, interpret and integrate scientific knowledge, but apart from integration they create relatively little new knowledge. Therefore, their credibility depends more on the protocols followed when identifying key issues and related scientific information, combining separate pieces of analysis into meaningful wholes, and bringing scientific facts and integrated analysis to bear on policy issues without distortion or loss of significant content. Therefore, these assessment and reporting systems have vital and direct interest in advancing research in knowledge integration.

Recommendation

GEO and similar IAR systems should actively encourage research into the theory and practice of knowledge integration and integrated assessment, and serve as testing grounds for emerging approaches.

Communicating information about complex issues, even if those issues strike a chord with people, is always more difficult than dealing e.g., with an environmental problem in a narrow context. When discussed in the abstract, many of the environment / development interactions will be difficult to recognize. One of GEO's and similar integrated assessment and reporting systems' core challenge is to be able to understand

complexity not only on the conceptual level but to be able to apply insights gained through theory to the analysis of issues relevant for policymakers.

## 6.2. Governance

### *6.2.1. Why governance?*

I view the role of governance in assessment and reporting as providing both intellectual guidance and leadership through coordination, consultation and management. A central question of governance is to whom are the producers of the assessment accountable and what the power structures are. Governance of scientific input is an important, though not the only governance issue. Governance requires liaising with bodies that exercise oversight, envisioning the way the assessment can fulfill its goals and translating the vision into an operational system that delivers the results. This involves creating an institutional strategy and structure, setting objectives, creating organizational structures and routines, mobilizing resources, building and cultivating partnerships, managing the analytic process and the production of information products, finding ways to protect scientific integrity, disseminating results, and finally ensuring there are mechanisms to collect and process feedback. As assessment and reporting should be continuous, these functions require organizational structures and institutionalization that can provide leadership on an ongoing basis (International Institute for Sustainable Development 1997).

From the perspective of governance, and particularly from the perspective of governing participatory assessment networks, it is important to understand where accountability ends and the line of influence begins (MacPherson, IUCN, interview). While GEO CCs and ACs are partly instruments of change and not parts of UNEP, their closer association with GEO means that they have responsibilities and they are to a degree accountable to UNEP based on short-term Memoranda of Understanding. Participants of regional consultations, however, who may contribute to the preparation of the assessment through

making comments, reviewing drafts or providing data are outside of the formal institutional structure and can be better classified as audiences.

The GEA effort put particular emphasis on the relationship between the organizational architecture of assessments and their ability to maximize trust between producers and audiences, scientists and policymakers. Central to this idea is that some 'boundary organizations' serve and are uniquely positioned to connect with the scientific world to produce or harvest policy-relevant knowledge, but also connect with policymakers and other audiences to better understand their information and research needs (Guston 2001).

Translating sustainable development into a governance model is a central challenge not only for assessment and reporting systems, but society in general. Many organizations, within or outside of government are still entrenched in what Ann Dale (1998) beautifully describes as "deep solitude", narrow technical conceptualization of problems and correspondingly constrained solutions. Governance is an area where progress is urgently needed, as indicated by the intense debates about global governance of environment and development in the running up to the World Summit on Sustainable Development in Johannesburg in 2002 (e.g., von Moltke 2001). The same debates illustrate that there is no full agreement on the nature of the problem, and consequently there is no unanimous agreement on the solution. While there is little disagreement about the linkage of governance and sustainability, the relationship is still insufficiently understood and, to be sure, whatever is understood certainly is not being translated very rigorously into practice. Projects such as the recently started Institutional Dimensions of Global Environmental Change (IDGEC) of IHDP hopefully will fill some of the knowledge gaps (Young *et al.* 1999). Despite the proliferation of theoretical literature (e.g., Moldan and Billharz 1997), governance and institutional components are typically the weaker parts of sustainable development indicator systems, a fact which itself is an indication of gaps in monitoring, data and probably understanding.

There is no generally accepted governance model for assessment and reporting systems. In some cases, the task is given to existing organizations, but it is common to set up new

ones, often with an ad hoc, multi-stakeholder character. A generalized governance model for a sustainable development reporting initiative is shown in Figure 9. The role of the stakeholder group in this case was to help identify key questions and issues from the perspectives of key audiences, while the technical advisory group was responsible for compiling relevant data and carrying out the analysis. The core team was responsible for overall project coordination and management and included representatives of a variety of stakeholders.

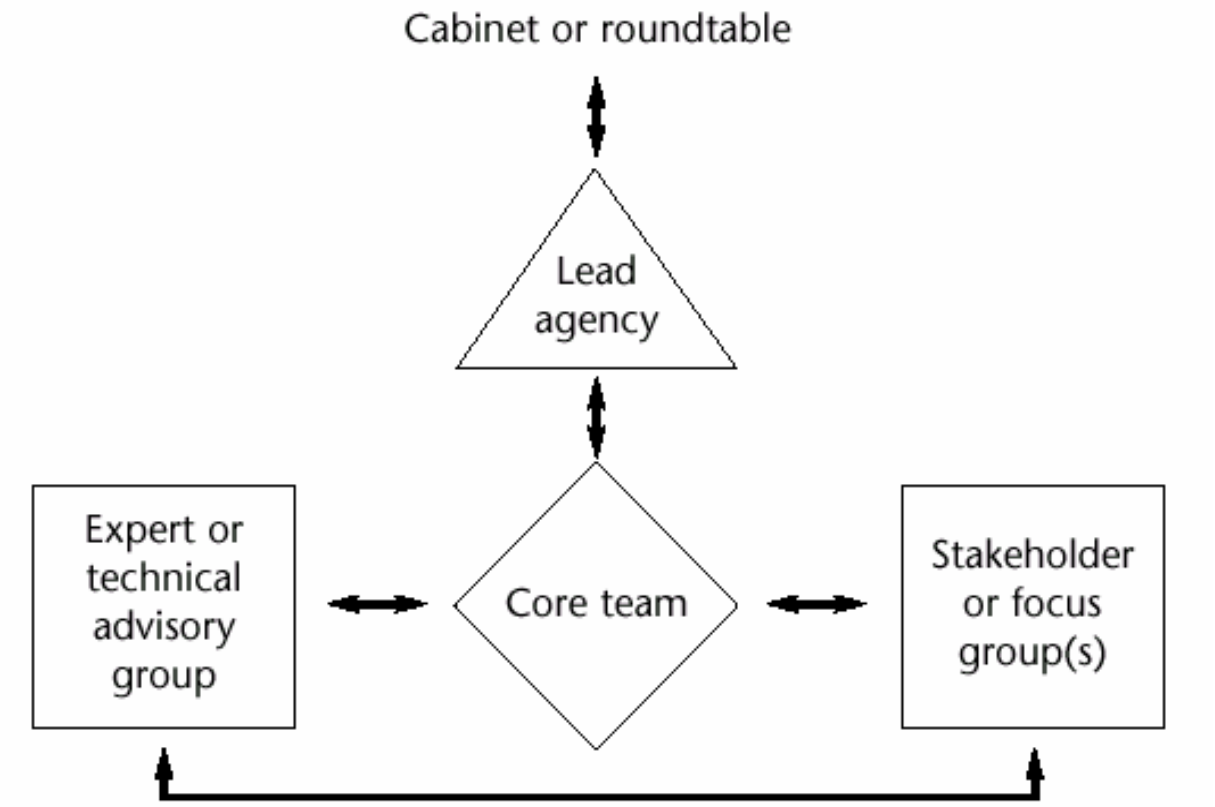



**Figure 9:** General organizational chart to manage an integrated and participatory assessment and reporting process (Manitoba Environment, as reported in Pintér, Zahedi and Cressman 2000).

Managing the interaction of science and policy is a key governance challenge, with major implications for credibility and saliency. Policy or political influence is a particularly

sensitive issue in traditional science assessments where the emphasis is on the production of primary data and new knowledge rather than meeting the information and knowledge needs of particular audiences. Even if there is less emphasis on – and capacity to produce – primary knowledge in GEO or IAR style systems, it is essential that the selection of issues not only accurately reflect policy priorities but also scientific understanding. This can be difficult, as the perspectives of science and policy are typically different. Through balanced governance, the assessment must find ways to address policy concerns but at the same time involve mechanisms to maintain the credibility and truth-seeking of science.

### *6.2.2. Governance in GEO and distributed systems*

Governance of multi-scale, thematically broad integrated assessment and reporting systems is a formidable challenge. It is particularly so because from the review of GEO's purposes, it is clear that beyond the goal of integrating the results of scientific research and policy perspectives, GEO is also a process with an agenda to help transform the way environment and development information is produced, presented and applied, not only on the global but regional, sub-regional and even national level. GEO's governance system must not only ensure that the expected outputs are indeed produced, but also that it provides an assessment (including assessment governance) model that can be adopted by others. UNEP's challenge is to provide leadership and marshal the necessary intellectual and organizational capacity and resources required by a cohesive global reporting system without the convenience of being a specialized UN organization[11]. This must include providing methodological guidance that UNEP provided mainly through production guidelines and interaction with GEO CCs though the GEO process, as previously discussed. As has become clear during the GEO-2000 process when GEO's management team in Nairobi was cut in half, a strong centre is vital to cultivate relationships with network members, provide methodological guidance and feedback, develop tools, raise funds, and represent the project to the outside world. Good governance requires not only a solid conceptual base and good intentions, but also staff and resources.

GEO being a typical polycentric system, some governance responsibilities are devolved to lower level nodes in the hierarchy. For instance, UNEP regional offices coordinate work in the GEO centers of their region, on lower spatial scales. Although not all of them get involved in producing content, some do, particularly in developing regions where CCs may lack adequate capacity and need more external assistance. In other regions such as Europe or North America, the regional centre's role is more focused on liaising with UNEP HQ and governments on the political level. With regard to the direction of the entire GEO process, CCs have lower degrees of influence, but there are again differences. Some CCs, typically in industrialized countries that have well-regarded assessment and reporting expertise, such as RIVM, the EEA or WRI are likely to play an advisory role.

In the case of GEO, it is particularly challenging to find a governance model to correspond with a structure that calls for analysis and integration along different axes: by region, by environmental theme, and by strategic issue. Addressing the regional or place-based domain is perhaps the least problematic as many centres from the same region know each other and UNEP Regional Offices provide natural focal points. Environmental themes are more difficult as they require a structure that pulls together centers with similar thematic expertise and interest, but potentially different culture and capacities. Rather than pursuing this option, UNEP in the past has chosen to carry out thematic integration more in-house, while for GEO-3 they have hired outstanding individual experts to provide global overviews of themes such as forests or freshwater.

Strategic issues may include data and indicators, capacity development or scenario analysis. These cut across both regions and themes, but they are central to the integrity of the assessment and its results. The results of this work do not necessarily appear in the main assessment reports, but might be published as separate technical publications. These are also issues that can be highly problematic, from the scientific, technical or political point of view that require a long-term endeavour of the broader scientific and policy community. GEO provides more than a forum for theorizing, it provides a 'laboratory', a forum for learning by doing. In order to facilitate the learning process, GEO includes Working Groups composed of CCs from different regions of the world that have interest

in a strategic issue and that are able to offer guidance to the rest of the GEO enterprise. Although overall guidance is provided by UNEP HQ, working groups can be led by a UNEP office (e.g., the Data Group led by the GRID Center in Geneva) or by ACs (e.g., the Scenario Group led by SEI-Boston or most capacity development activities spearheaded by IISD).

As a result of these parallel tasks and activities, GEO's governance model is complex with multiple overlapping hierarchies. A single member of the organizational hierarchy can play a leadership role in one context and be a contributor in another. In fact, the various working groups can have different dynamics of their own within the constraints of the overall GEO production. As earlier suggested, GEO can be classified as having a panarchic structure where central dynamic assessment processes and governance structures are accompanied by – and hopefully harmonized with - nested processes and corresponding governance architecture on finer scales. In principle, the outputs are information products that are based on the same framework and thus facilitate cross-scale global level aggregation and intra-scale comparability, but on lower levels are fine tuned and contextualised according to the needs and conditions of 'local' or sectoral policy audiences. In practice, the assessment requires a constant governance effort to keep semi-autonomous working groups adhering to common frameworks, methods and agreed upon production formats and guidelines.

### *6.2.3. Implications for effectiveness*

The strength of multi-scale, distributed assessment systems is partly in building on the initiative, capacities and creativity of individual members in the network. This strength could be seriously compromised by excessive centralized management, where a lead agency wanted to retain power and decision making authority on details it could not possibly understand or follow. This would not only require tremendous resources and staff capacity, but it would also make network members feel disempowered and over time disinterested. Therefore, in multi-scale, distributed assessment and reporting systems, the principle of subsidiarity should apply, whereby decision-making powers concerning a

particular problem domain are devolved to that level of the hierarchy where knowledge and capacity are at a maximum with regard to that domain.

GEO production guidelines are some of the most important tools through which GEO can both enable individual regions or CCs to carry out their work semi-autonomously, while also ensuring that their process and products are synchronized and meet the necessary scientific standards. However, production guidelines should be improved, more detailed and customized to better suit the needs of particular participants (O. El-Kholy, AGU, interview). One may also find that their consistent application throughout the network and through the reporting process is less systematically enforced; in fact, inadequate compliance with guidelines in the early drafting stages can be a common problem (UNEP 2000d). While this may not significantly influence the effectiveness and quality of the final product if appropriate correction and quality control mechanisms are in place and there is sufficient time, it makes the integration of pieces prepared separately more difficult.

How can these problems be addressed through changing production guidelines? As the guidelines have to be fairly general, there is plenty of room for different interpretation by its primary audiences – the producers of the assessment. In order to prevent the unhelpful variability arising from that and increase the consistency of output, guidelines may have to be more closely integrated into the actual process of content preparation. Guidelines could be more step-by-step and, as Osama el-Kholy pointed out in my interview, have a feel similar to the worksheets made available to the participants of IPCC assessments. Another possibility is a better use of assessment examples to help better understand the type of output expected[12]. Ultimately, production guidelines, worksheets or workspace, and examples could be integrated in an electronic template where some structural elements, analytic steps and formats can be made mandatory [13].

| Recommendation: |
| --- |
| Strengthen production guidelines and ensure their more systematic and rigorous application in order to increase analytic rigour and improve consistency of outputs. |

GEO and similar overview reports are accountable primarily to policy audiences. UNEP and the GEO team have significant freedom in terms of determining assessment and reporting strategies, but the power to set overall direction ultimately rests in this case with the Governing Council, a political body. GEO is in interesting contrast with other, finer scale assessments where governing bodies and key policy audiences may ask specific policy questions to which the assessment must respond. For example, RIVM is required to respond to specific questions of Ministers in the Dutch government (R. Maas, RIVM, interview). The National Research Council (NRC) in the United States and to some degree the European Environment Agency serve a similar purpose, bringing high quality scientific expertise to bear on priority issues that concern policymakers.

At the international level there are very few fully developed science assessments. Because of their cost and complexity, it might make better sense to focus on one or two regimes versus scattering attention and resources widely (von Moltke 2001). An architecture that involves in-depth science assessments clustered around one or two highest priority issues would, however, require a second line of defence, both partial science assessments and GEO-like reporting mechanisms that harvest and compile the results of other research programs, and interpret them for policymaking in the international context. While GEO is not able to respond with primary research to policy questions, better orientation from the policy community concerned with global environmental governance could increase its saliency.

The governance structure of GEO makes provisions for formal stakeholder consultations on the regional level, typically involving government representatives, external scientists, NGOs and others, but the groups that show most interest are typically government departments. While the mechanics of how interaction takes place is an interesting question from the perspective of participation, from the perspective of governance what matters is who sets the rules that determine by whom and how questions can be posed and how those questions will be answered:

> "If a group like GEO is going to engage in a participatory monitoring system or reporting system, to me there is a moral issue, who decides who participates, and who gets left out, who is not invited to participate? So that creates one level of dilemma. A second level of dilemma to me is having the potential to make reports or discover data which suggests that a program is on the wrong track or should be doing other things as a first priority." (F. Carden, IDRC, interview)

The legitimacy of GEO is particularly sensitive from the perspective that it is requested to put increasing emphasis on 'actionable items' or policy recommendations. Providing recommendations can be risky for science assessments that need to maintain neutrality and integrity, but for 'boundary processes' like GEO this is a must. To put it another way, policymakers seem to be looking not only for neutral science assessments that leave drawing conclusions entirely to the reader, but also and perhaps increasingly to processes that use the science then, based on the understanding of what the science says and what the policy needs are, to weigh risks, uncertainties and benefits and lay out alternative courses of action in the form of recommendations.

> "Immediately after GEO-2 was launched, our (UNEP's) Executive Director held a meeting with environment ministers of ASEAN … and they said, Dr. Töpfer, you know this is a very good report, it's a nice coffee-table book, it will be good for our students and people wanting to understand what is happening at the global level. But at our level we can't read this thick document. Can you summarize for us as ministers of ten countries, what do you want us to do? … Tell us what action you would like us to take, in form of legislation, in form of economic instruments, in form of what kind of preventions you would like so that we don't contribute and make things worse. … You have all these detailed things behind you, what would be some areas where we countries can make a positive contribution. And that's the question." (S. Shresta, UNEP-ROAP, interview)

The governance structure that makes these recommendations, arguably one of the most important outputs of the entire GEO system, must remain legitimate and salient in the

eyes of key audiences and needs to balance very carefully between science and policy in order to avoid exceedingly compromising on either. For now, recommendations are prepared by UNEP staff and advisors, in consultation with high-level policy experts. While this process ensures a high degree of salience, its legitimacy may be constrained by political feasibility or convenience.

| Recommendation |
| --- |
| As an experiment, GEO could involve an independent, arms-length representative body of experts – policymakers, scientists and others – to study GEO's findings and draft a set of practical recommendations that could either be part of the main GEO report or published as an accompanying document. |

Independent expert contributions or at least independent peer review could enhance the recommendations' credibility. Peer review is essential for the credibility of not only science assessments producing primary data, but also for systems that interpret scientific findings for policy audiences. Quality can be compromised in many ways. Partly there is a risk that important details may get lost in the 'translation' process but scientific data may also get misrepresented in the process of integration. Most reporting systems have some form of quality control in their governance system. This may range from independent peer review to sending integrated reports back to the producers for primary scientific data to make sure their findings are properly represented, as practiced in the Nature Balance program of RIVM (L. Braat interview). The GEO process involves reviews in several iterations, part of it involving review by policy audiences in regional and global consultations and partly review by scientists.

Currently there are no provisions for anonymous and thus fully independent peer review and there is not always sufficient time for review and subsequent revision cycles to properly run their course, even though the importance of the issue is recognized in the assessment system structure. The implementation of peer review could probably be more systematic, and I would add that GEO could also establish or strengthen mechanisms for independent and anonymous peer review (M. Chenje interview, Bakkes *et al.* 1999).

Recommendation

GEO and similar assessment and reporting systems could increase their credibility by strengthening the governance of quality control and scientific, anonymous peer review mechanisms. With regard to particularly sensitive sections this could also include a review whether assessment producers properly responded to the questions and concerns of peers in the revisions made.

Broadening the circle of reviewers in IARs is a particularly important necessity because of the interdisciplinary character of the assessment. This leads to some unique challenges, As in the case of traditional disciplinary research expertise brought to bear is often narrow, there are well outlined groups of researchers who can claim the necessary expertise required to undertake peer review. IARs that cross not only disciplinary and spatial boundaries, but also the boundaries of science and policy, reviewers may well need to include not only the scientist, but the policy expert or the 'stakeholder' whose perspectives need to be correctly reflected in the analysis.

## 6.3. Participation

### *6.3.1. Why participation?*

Opening up decision-making in sustainable development in general and in assessment systems in particular for broader participation had been argued for many years and for many reasons (e.g., Funtowicz *et al.* 1999, Jäger 1998, Ravetz 1999, Shindler and Cheek 1999, Tóth 1989, van Asselt 1994). There are also many experiments to integrate broader participation into processes that used to be previously thought of as essentially scientific and in the exclusive domain of experts (Dürrenberger *et al.* 1997, Vennix 1996). On the legal front, the requirement or need for public participation has been integrated into national laws in some cases and into international conventions (e.g., Economic Commission for Europe, Committee on Environmental Policy 1998).

At its heart, the need for participation can be argued on epistemological grounds. Expert knowledge – say scientific or engineering – can be successfully applied to study phenomena that have either limited number of variables or their complexity can be reduced for analytic convenience without significant loss of information. Reductionism is characteristic of the Newtonian approach, the 'clockwork universe', the world of engineering, controlled laboratory experiments or neoclassical economics. It is argued that the ability of this approach to deal with most environment / development dilemmas in real world settings is limited (e.g., Funtowicz and Ravetz 1993, Kay *et al.* 1999). Environment / development issues are inherently complex and management is always constrained by emergent properties, uncertainty and surprise. Any given expert entering real-world environment / development dilemmas from a disciplinary solitude would quickly find her/himself in the company of a great many others, who all have legitimate claims to truth, yet who all bring different values to the table, view the world through different "perceptual lenses" and thus offer alternative explanations and solutions to problems (Allison 1971). In fact, the very meaning of expertise and who is allowed to define it is increasingly contested (e.g. Jasanoff 1990)[14]. Traditional knowledge embedded in indigenous cultures is a case in point, where valid knowledge and expertise are derived through processes that are not connected to the institutionalized form of science. Unless the legitimacy of these different 'lenses' is recognized and addressed, fairness and equity, both in the inter- and intra-generational sense can suffer. The practical implication is that the assessment may well fail to identify unsustainability problems – or sustainability solutions – that are visible only to external participants and whose cooperation would be essential in any meaningful response strategy.

I agree with the position that the diversity of perspectives is not an "unfortunate accident" that could be fully resolved through better science, more research or the designation of some perspectives as irrelevant (Funtowicz *et al.* 1999). Rather it is an aspect of reality that assessments need to recognize, make explicit and address. In order to achieve this, assessments need fora where constructive dialogue can take place and where 'overlapping consensus' shared by all stakeholders can emerge (Rawls 1987). Process is

key, for the way participation is run typically has a decisive influence on results. Running the participatory process involves many choices, such as:

Who is invited to participate;
Who decides who to invite to participate and who sets policies for the mode of interaction;
In what do people participate and in what stages (e.g., before or after the problem is fully framed);
Are all participants given equal rights, access and opportunities;
Are there efforts to enable legitimate, but for whatever reason handicapped (e.g., due to lack of funds, language, physical or cultural barriers) participants to enter and meaningfully contribute to the process;
How does the process identify and keep focused on priority issues without the risk of ignoring relevant concerns and perspectives;
How are different, often mutually exclusive views reconciled;
How is the feedback received processed and reflected in policies and products;
Are participants given opportunities to review and comment on products as they are being developed and finalized;
Who presents results to key audiences and how;
How is participation institutionalized;
How is quality control applied to the participatory process?

The checklist is based on my experience taking part in the design and management of participatory assessment and reporting processes dating back to the early 1990s. It is not exhaustive and many other variations exist (e.g., Kluck 1995). It shows, however, that there are many options associated with the way participation is organized and many opportunities to influence outcomes and effectiveness. Depending on the choices made, from the perspective of decision-making and control, the outcomes may range from manipulation at one end of the spectrum to citizen action at the other with several intermediate stages (Arnstein 1969). The model's underlying logic is the relationship between participation in decision-making and participation in implementing decisions. At

the manipulation end of the spectrum, decision-making and implementation are detached and those implementing are in fact used to further the goals of those making the decisions. At the citizen action end, decision-making and implementation are fully intertwined: the locus of decision-making and control is coincidental to the locus of action. In assessments carried out or coordinated by specialized agencies, there is always some separation – in fact separation is seen as desirable from the perspective of ensuring the independence of science. There is, however, also an essential need to make assessments and the underlying science speak to the needs of policy-makers.

Participation in assessments plays an important role in shaping the interplay of science and policy. While on balance, participation may be strengthening assessments, it also involves tradeoffs. In general, through participation science can gain in policy relevance but risk credibility and independence. The policy side can gain scientific advice it can actually use at the cost of opening up the decision-making process to outside scrutiny and longer deliberation. Scientists would also find that substantive policy-maker involvement turns the assessment process quintessentially political and interest driven. Even if the results of research are scientific, the conclusions drawn become negotiated and to some degree negotiable, particularly if stakes are high. As MacNeill (1990) observed, even sympathetic governments, politicians and corporations are often found to influence assessments whose findings would require politically and economically costly responses. Broad participation in 'negotiated science' should not mean giving up on objective scientific criteria, but increasing the link between the science and the social or political side explicitly, thus improving both (Forrester 1999). To quote MacNeill (1990) again, "the result may be a scientific assessment, but it is also a political report".

### *6.3.2. Participation in GEO and distributed systems*

Participation has been an essential feature of the GEO process with implications for and links to several other design elements. The reference paper for UNEP's Environmental Observing and Assessment Strategy makes a clear statement on this:

> "UNEP's assessment activities in general should be based on a participatory approach. To secure both political support as well as the scientific credibility of the results, the assessment and observing activities should be an open and participatory process." (UNEP 1998)

In GEO, participation is not really a choice but a necessity. Although UNEP has a global environmental assessment and reporting mandate, without partners its capacity would be inadequate to undertake multi-scale, integrated assessment and reporting. Most of the issues with which GEO deals also require integration of information and knowledge from a wide array of disciplinary and geographic perspectives, so GEO also needs to connect with key policy audiences, preferably throughout the entire process.

I find it helpful to separate participation in GEO into two categories, building further on data from my interview with Nancy MacPherson at IUCN[15]. The first category includes involvement of individuals in organizational relationships and processes where formal accountability to the lead agency, in this case UNEP, dominates. I will call these internal participants, where internal means not only internal to UNEP, but institutionally internal to the GEO process. These serve to decentralize the production of the assessment and ensure the analysis reflects regional and thematic concerns, not imposed or predetermined by a global center, while keeping to common guidelines. The process also serves as a capacity development / learning mechanism for members of the network and may provide them with an opportunity to bring some of their existing products and expertise to bear on a global assessment. The latter is done more often by groups, say, with thematic expertise such as WRI expertise in data systems or RIVM and SEI in scenario analysis.

The second category that for convenience I call external participants and that includes parties involved in organizational relationships and processes which do not involve formal accountability to UNEP. These may also contribute relevant regional, thematic or sectoral information and perspectives to the assessment, but the organizations or individuals involved are also key audiences. It is equally important that they become familiar with assessment methods, results and underlying assumptions during the process.

External participants may include representatives of non-governmental or governmental organizations consulted by GEO CCs during the preparation of regional chapters. The same groups may also be invited to attend regional consultations to review and comment on completed drafts. To some degree, the line of accountability / influencing can get blurred, so the two categories can overlap.

Not only do participants – external or internal - play different roles, the roles are also changing according to the phase or aspect of the GEO process (**Table 4**). As the assessment proceeds, the intensity and nature of the work require changes, some participants become more involved than others. For various reasons, high activity periods can be separated by months of low level or no activity, which had been identified as a source of frustration by several of my interviewees. Uneven process intensity is a management challenge both for UNEP and the centers involved. During one stage in the GEO-2000 process, UNEP has lost much of its GEO staff, and the remaining core team worked in continuous peak mode. With regard to CCs, during peak load they may need to hire temporary help who are hard to keep later because of funding constraints.

Without more continuity, institutional memory and capacity built up in one cycle of the assessment can easily be lost - a concern shared by some of my interviewees (e.g. E. Gutierrez-Espeleta, Observatorio del Dessarrollo, interview).

GEO hasn't had much history of ordinary citizen participation, although many of the key outputs, particularly the main global or regional reports are supposed to be written in non-technical language, with the general reader in mind. There are pragmatic reasons due to the cost and logistics of organizing a participatory process open to the public in all regions of the world. Although not without problems, the approach taken by GEO is often followed by other public participatory processes. Rather than invite the public at large to contribute, representatives of key stakeholder or interest groups are identified who can speak for and represent a wide variety of legitimate interests. This can work, but constraints associated with the design of participation as identified in the bulleted list in section 3.2.1., may also apply.

**Table 4:** Participants and their roles in GEO to date (based on interviews conducted).

✔ Primary role

(✔) Secondary role

| | | Strategic planning | Fundraising | Scoping | Data gathering | Analysis | Review | Evaluation | Methodology development | Capacity development | Communication |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Accountability dominates | UNEP HQ | ✔ | ✔ | ✔ | (✔) | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| | UNEP RC | ✔ | | ✔ | (✔) | (✔) | ✔ | (✔) | (✔) | ✔ | ✔ |
| | CCs | | ✔ | (✔) | ✔ | ✔ | ✔ | (✔) | (✔) | ✔ | (✔) |
| | ACs | (✔) | ✔ | (✔) | ✔ | ✔ | ✔ | (✔) | ✔ | ✔ | (✔) |
| Influencing dominates | Outside scientist / expert | | | (✔) | (✔) | (✔) | ✔ | ✔ | | | |
| | Policymaker | | | (✔) | | | ✔ | | | | |
| | Other UN | | | | (✔) | | ✔ | | | ✔ | |
| | Media | | | | | | | | | | ✔ |
| | Business | | | | | | ✔ | | | | |

The interviews suggest that there is a majority view that the current funding structure falls short of what would be required to maintain a real and truly active network with increasing and sustainable capacity – a link to the capacity design element to be discussed later. In order to address this issue, UNEP and the CCs could work together on a model that approaches this problem from the long-term institutional development perspective, keeping in mind the interests of UNEP and GEO, but also the interests and incentives required by network members as internal participants.

The participation of external stakeholders or audiences, although systematic, is less intense and more uneven. It can be considered systematic because of the institutionalization and formalization of consultations with these stakeholders through written peer review and particularly through formal regional consultations. However, contributions can only be invited, not required, and many participants may choose not to make substantive contributions.

From the management point of view, this is probably less of a problem as the effort external participants invest in GEO is not as substantive and they do not require funding. However, it is important that in the current design there is little external involvement in the scoping phase of the process, so at least initially issues and foci are defined without truly substantive stakeholder input. Authors of first drafts may consult organizations and experts in their region as they scope out respective sections of the assessment, but there has been no budget to call up-front stakeholder meetings in every region, or to work closely and systematically at this stage with decision-makers to clarify policy questions.

Both external and internal participation involves a wide variety of stakeholders. In practice, government representatives, NGOs and academics are the most active while the business community has been missing from most of the consultations (L. Pintér interview data). Given that the UNEP regional offices play a significant role in organizing regional consultations, much depends on the approach and activism of these centres to reach out to and engage relevant groups and facilitate the process. They also need to make sure feedback is taken seriously into consideration and that participants know it. Much can

depend on whether participation is perceived as a bureaucratic or political necessity or whether relationships are properly cultivated and involve real engagement, often on the personal level, with key groups. My overall conclusion on this point is that while UNEP definitely seeks real and not token participation in GEO, the strategy and implementation of partnership-building was weak. In recognition of the need to strengthen this aspect, early in the GEO-3 process UNEP brought on board a partnership officer.

### *6.3.3. Implications for effectiveness*

Participation is embedded in GEO and influences parameters of effectiveness in many ways and through many stages of the process. Although there is little doubt participation is an important variable of effectiveness, it would simplistic to assert that more participation, as a rule, leads to better outcomes and increased effectiveness (Eckley 2001). Almost by definition, increased participation adds to the diversity of perspectives and complexity of assessment that can bring existing conflicts to the fore. Without mechanisms to address conflicts and the often underlying differences in worldviews and interests, the assessment process may become stalled or in the course of searching for a compromise solution to handle complexity produce results that are neither representative nor scientifically credible. It is interesting to contemplate what would guide assessment and reporting in the absence of strong participation. Without participation, one likely outcome would be less emphasis on issues and perspectives relevant for policy-makers and the public and more emphasis on data and issues of interest to the scientific community involved in the assessment process (Ronan Uhel and Nancy MacPherson interviews).

Research from the GEA program has identified several cases and pathways where substantive participation in an early phase increased the relevance of the assessment for its intended audiences (Biermann 2000, Eckley 2001). GEO can use both internal and external participation to identify critical issues, regions and policy with which audiences are most concerned. CCs are based in the regions and as they are usually connected to policy audiences for many other reasons besides GEO, they are in a position to control the direction of their regional sections and report on matters of policy relevance as far as

they comply with production guidelines. External participants usually get involved at a later stage than do reviewers.

In a global distributed assessment, balancing developing and developed country participation is a particularly sensitive issue that can either add to or seriously compromise legitimacy. Cultural diversity has been identified as a significant challenge to participation even in cases, like the EEA's integrated assessment, where the diversity of stakeholders and issues is more limited (European Environment Agency 2000). For GEO and other global integrated assessments, this can be a major hurdle that may require focused efforts to improve cross-cultural communication.

Restricting developing country participation or endorsing it in principle but implementing it only through token measures can backfire on several accounts. It can mean passing up not only opportunities to identify and address issues that are truly of interest to developing regions – versus issues thought to be important for them by Western 'experts' – but also creating serious image and political problems. It has been all too common in global assessment systems to see overwhelming representation of developed countries, and developing countries represented only on a symbolic level, underrepresented in steering groups with real decision-making power, or limited to developed country academics and researchers who have worked for a long time in the West but who can be displayed as representatives of the developing world because of their passport.

A related question is in what do stakeholders participate? As raised by Fred Carden in our interview, is participation limited to a mechanical contribution of data and information based on predetermined guidelines or are issues of substance and policy opened up for discussion that participants are invited to contribute to? While the first option does little to increase legitimacy, the second does far more to build ownership, buy-in and support for the process, although it might also result in severe dissonances or conflicts (B. Siebenhüner, personal communication). Unless mechanisms in the assessment exist to handle such conflicts, the benefits arising from a genuine participatory process may be hard to realize.

From the beginning, GEO has been planned with a balance of developed and developing country participation formally in mind (UNEP 1997b). The declared goal was genuine and proportional representation of developing countries in decision-making, giving genuine consideration to concerns, and providing them with the freedom to identify and address *their* priority issues in the assessment. While it would be unreasonable to claim this model has been perfect and particularly that it has been thoroughly implemented, both developing country governments and CCs expressed support for the approach to participation (Universalia 2001):

> "…that was the first time where a group of people was asked what was the best way to do environmental reporting on the global level. As I said, before a group of international research institutions, basically from the North got together and prepared an environmental report for the world. And in Groningen another message came out, well, this is one planet, this is just one vehicle and the South has the right to say how we see and envision the environmental problems that we have. And from then on I fully supported the idea of GEO to develop and to elaborate a global environmental report where feedback from the South was planned to be considered in the report."
> (Edgar Gutierrez-Espeleta, Observatorio del Dessarrollo, interview)

Representativeness and the legitimacy of the participatory process is important not only in the developed / developing country context, but also in terms of sectoral participants. GEO has been strong on involving participants from the government, academic and inter-governmental sector, it has been weaker in terms of involving NGOs outside of the CCs and it has clearly lacked substantive involvement of the business sector (Marion Cheatle interview). This may have implications for to whom the assessment will be legitimate and salient.

Much depends on the choice of participating institutions and individuals. Even if there are strict and consistently applied selection criteria for participants, major differences among collaborating institutions are unavoidable in any multi-scale, polycentric

assessment, given the primacy of adequate regional representation. In many regions, even the best institution can be far from meeting the standards that would make it minimally suitable for a truly credible assessment: given that no representation is not a choice, these weaknesses need to be addressed through capacity development.

Experience shows that the success of multi-scale networks is strongly influenced by the individual commitment of participants (Martinos and Humphreys 1992). Within organizations, the level of preparation of individuals assigned to contribute to the assessment is critically important. The goal of GEO is institutional participation, engaging CCs as organizations and not individuals. Yet, the success of the work with a CC often comes down to the commitment and expertise of a single person in the given centre (Marion Cheatle interview). Even if well prepared, junior bureaucrats or technical specialists can hardly offer the same perspectives as senior policy-makers. Who leaders of GEO CCs assign to preparing the actual analysis depends on many factors, including perceived significance of the project and resources available. The lower the level of available funding, the higher the probability that writing will be assigned to junior officers (M. Choudhury, personal communication). External participants do not receive funding, but particularly in developing countries they would require that GEO cover the costs associated with their participation. Although participation can also take place through correspondence, it is not a substitute for the personal interaction necessary for building and maintaining interests and involvement.

Other, more traditional science assessments, such as the IPCC or the more recent Millennium Ecosystem Assessment employ a selection process to identify contributors to the assessment based on their past achievements and scientific track record. Although the GEO approach of working with pre-selected collaborating centres has advantages e.g., in terms of facilitating long-term institutional capacity building, ideally it would be combined with engaging the best qualified individuals at least within those organizations.

Recommendation

Besides its emphasis on institutional participation, GEO could increase its effectiveness, particularly its credibility by placing equal or more emphasis on individuals representing institutions, taking part in the process and actually contributing to the report.

Participation serves different purposes depending on the phase of the assessment process. Although in principle nothing prevents a GEO CC from involving external participants in identifying critical policy questions up-front, formal consultations in the early phase are not built sufficiently into the process and not built into budgets either. Tight time-lines and limited capacity are additional constraints:

> "We're missing a scoping phase in GEO and that's because by sheer lack of capacity UNEP cannot start thinking about GEO X+1 before delivering X." (Jan Bakkes interview)

Without formal requirements for early participation, a lot depends not just on UNEP's guidance but the views and approach preferred by the individuals responsible for preparing first outlines at CCs. Of course, this doesn't necessarily mean early participation and scoping doesn't happen, but if it happens, it does not happen systematically throughout the network. Without consulting participants on the demand side of the assessment, producers may easily identify and get attached to issues that have no relevance for or actually be harmful to audiences. Echoing Miller (1998) and others, I support the view that both saliency and legitimacy would increase if provisions were made for systematic stakeholder involvement early in an assessment's scoping phase.

Recommendation

IAR systems should make provisions for substantive stakeholder participation in the very early, formative phase of the assessment process.

The pathway to the legitimacy of the assessment comes partly through the legitimacy of the participating institutions. However, legitimizing is a two-way process. On the one

hand, locally known and respected institutions can be seen as representing and bringing to the assessment genuine and legitimate local concerns. On the other hand, they also benefit from having taken part in an internationally sanctioned process. Many organizations find it attractive to participate not only because of the opportunity to obtain additional resources –an important incentive – but also because of increased regional and international visibility and legitimacy (Atiq Rahman interview, Appendix XI in Universalia 2001).

Participation may not only increase legitimacy and saliency of an assessment but also play a crucial role in 'marketing' it – another tradeoff that may counterbalance the risk of lower scientific credibility and quality associated with a broad participatory process.

> "I think it's a much more complicated process of course if you want to do it participatory and the (end product) is not going to be as good I would guess. But the whole role of involving people in the process serves a purpose in itself. And that's why GEO is so well-known, it may not be because it is the best report around, but because people are actually involved in producing it and I think that makes a change in itself."
> (Lars Mortensen interview)

The assessment system may be content that participation, focused on production of knowledge alone, can indirectly increase awareness. Alternatively, participation can be seen as a strategic element in the assessment's communication strategy and dealt with accordingly. Some of he practical implications of this would be, e.g., the conscious inclusion of opinion leaders – e.g. journalists - in the participatory process who are in a position to publicize the process and results of the assessment.

| Recommendation |
| --- |
| Integrating participation and marketing strategies throughout the assessment process and across Collaborating Centres across all scales could contribute to increased awareness of the results and improved effectiveness. |

While talking about marketing strategies in the context of science assessments or GEO-like reporting programs would have raised eyebrows a few decades ago, it has been recognized that these processes and products are good candidates for social marketing, defined as "the application of commercial marketing technologies to the analysis, planning, execution, and evaluation of programs designed to influence the voluntary behaviour of target audiences in order to improve their personal welfare and that of their society" (Andreasen 1995).

## 6.4. Capacity and capacity development

### *6.4.1. Why capacity and capacity development?*

Building on Sen's work, for the purpose of this analysis and in the broadest sense I define capacity as a property of a system or its component to function according to internally derived or externally imposed expectations (Sen 1999). Taking this definition further, capacity should be interpreted in the context of the functions a person or organization can be expected to perform. Sen's mention of functionings – in plural – indicates capacity's multiple dimensions in terms of functions performed, but also imply multiple preconditions or enabling factors for capacity itself to be available. At the simplest, capacity is determined by attributes that are internal to a system or external, i.e. can be associated with a system's environment.

Problems associated with uneven and inadequate capacity are almost unavoidable when working in the context of distributed networks that include members from developing or transition economies. The challenges are not unique to assessment and reporting. They are associated with most initiatives, particularly those involving the public sector, NGOs or local communities (Grindle 1997). Apart from the work of Miller (1998), capacity has not been a central consideration in the GEA program, yet it affects most other design elements and depending on the particular capacity problem, it may affect any one of the effectiveness criteria. Practitioners must not only understand the nature and context of capacity problems in theory, but also find ways to address them through capacity

development. Both capacity to do the assessment and capacity development as an almost unavoidable and associated activity have implications for effectiveness.

The causes of capacity problems are often historic and deeply embedded in the socio-economic character of a given country, region or organization. Lack of capacity is usually a systemic problem that cannot be 'fixed' quickly and easily. An iterative, evolutionary approach and perspective is essential. Without a systemic approach sensitive to context, capacity development can be no more than a public relations exercise, without lasting positive impact. It may actually lead to major and irreversible harm. In order to avoid the repeat of past failures, scholars concerned with the underlying flaws in the dominant capacity development paradigm point to the importance of mechanisms to derive capacity needs from the conditions, aspirations and interests of those short of capacity rather than linear projections of those thought to have adequate capacity available (Miller 1998). Indeed, while developing or transition nations may lack capacity to carry out scientific assessments, many donors and developed world institutions have themselves often demonstrated a fundamental lack of capacity to understand real needs, the importance of context, and the implications of ill-conceived capacity development efforts for sustainability.

Replacing a participant without capacity in the assessment with one that is better prepared is a possible shortcut, though does not necessarily lead to increased capacity and may lead to many other (e.g., political) ramifications. It is common to find cases where regardless of the issues in question, international agencies work with the same individuals. While this might be good for producing predictable results, it does not broaden the circle of contributors with expertise and experience, and it may lose an opportunity to discover new and relevant information. At the same time, it may actually narrow the political capital base of the efforts.

Formally, the GEO network is a network of institutions, although individuals often play a key role. Even if network members are carefully screened for existing capacity and experience, in truly global assessments there will be cases when ideal institutional

candidates are simply not available. One must make hard choices between, say, not having representatives from particular regions at all or working with centres who need to learn and develop to be suitable for the task. For the sake of overall long-term success, there also must be mechanisms to replace network members whose capacity constraints are too difficult to overcome. This strongly suggests there is a need for criteria or at least clear mechanisms for judging when a capacity constraint exceeds the value of keeping a participant involved as well as an ongoing process to measure the impedance factor of that constraint. Keeping the contextual factors strongly in mind, some of the criteria to assess CC performance might include the following:

- Quality of contribution;
- On-time delivery;
- Region-specific or thematic scientific and policy expertise;
- Substitutability;
- Cost of involvement; and
- Organizational mandate.

To turn problem into virtue, assessment systems may actually be set up to develop capacity, 'learning-by-doing', while also fulfilling their assessment and reporting mandate. This may involve compromises in terms of process and quality of product, and will require patience and firm commitment to sustained and sufficient investment, financial and other, into the weaker but promising members of the assessment network. Capacity development has two sides: those in demand of strengthened capacity, and those who are – in principle - in a position to pass on required skills, knowledge or resources. This is not necessarily to say there are two parties in a capacity development process. There may be many, depending on the type of expertise required. As the long history of failed capacity development efforts illustrates, the aspirations, ambitions and needs of the capacity developer and recipient side do not necessarily match. It is probably far more common to find capacity development programs with little understanding of systemic constraints, with short time horizons and no follow-up. Such interventions may do more harm than good by alienating participants from future capacity development activities.

After decades of rather troubled history, there are signs of an emerging capacity development approach that reflects a synthetic and more realistic perspective. Even terminology is changing. Referring to capacity development reflects the increasingly common position that what was referred to as capacity building tended to ignore the fact that some capacity exists in all situations and serve as a basis for any subsequent development (UNEP Division of Environmental Policy Implementation n.d.)[16]. Beyond semantics, the terminology reflects the need for a different underlying approach. The particular concept I find of most relevance here is Capacity Development in Environment (CDE), defined as:

> "…the ability of individuals, groups, organizations and institutions in a given setting to address environmental issues as part of a range of efforts to achieve sustainable development. (…) The key underlying principles of the CDE concept are that it integrates environment and development concerns at all levels, aims to strengthen institutional pluralism, belongs to, and is driven by, the community in which it is based and involves a variety of management techniques, analytical tools, incentives and organizational structures in order to achieve a given policy objective."
> (Organization for Economic Cooperation and Development n.d.).

This reflects an understanding that capacity development must be based on a thorough, unbiased and mutually shared understanding of not only capacity needs, but also how capacity once available can be sustained, by whom, based on what incentives, and to what end. The architecture of distributed integrated assessment and reporting systems is complex and capacity constrains may arise on many levels. As network members may fulfill different tasks and they may be of different character, they may have unique capacity demands and require custom-tailored capacity development strategies. Capacity constraints may also arise either on the producer or recipient side of the assessment. The assessment would be more concerned with capacity gaps involving CCs, but it may not be able to ignore problems affecting external partners. Capacity building is an implicit but important goal of involving policy audiences in the assessment process.

Usually the questions regarding capacity to do what cannot be reliably answered without the demand side thoroughly understanding requirements and systematically reviewing its existing capacity, presumably in close cooperation with those being able and willing to help. Keeping this approach strictly in mind, the exercise can be considered a needs assessment. This cannot be superficial. It is far too common to find that organizations skip or glide over this phase quickly with rather predictable results – recommendations for the application of generic and spectacular off-the-shelf 'solutions', pre-packaged and already in the tool-kit of donors, but little or no customization to actual needs and context. These are doomed to be typical case studies of failure.

**Table 5:** Capacity development efforts (Grindle, 1997).

| Dimension | Focus | Activities |
|---|---|---|
| **Human resource development** | Supply of professional and technical personnel | Training, salaries, conditions of work, recruitment |
| **Organizational strengthening** | Management systems to improve performance of specific tasks and functions; microstructures | Incentive systems, utilization of personnel, leadership, organizational culture, communications, managerial structures |
| **Institutional reform** | Institutions and systems; macrostructures | Rules of the game for economic and political regimes, policy and legal change, constitutional reform |

Integrated assessment and reporting requires many types of capacities. Grindle (1997) suggested a broad framework for the classification of public sector capacity issues and capacity development efforts (**Table 5**). The typology is general, so it also applies to assessment and reporting systems. Given the scope of capacity development in GEO, however, I am more interested in this study in the human resources and organizational dimensions. A significant amount of direct capacity development efforts in GEO would fall into the human resource category, while UNEP also provides assistance to CCs that

would fall in the organizational strengthening category. Because of the capacity issues particularly important for GEO, I put special emphasis on the treatment of particular activities in capacity development strategies, including tools and methods, expertise and experience, and resources.

Assessment and reporting requires tools and analytic methods that strengthen the quality and impact of the products and increase the efficiency of the process. IAR and GEO rely on a series of common software and hardware tools such as presentation, database-management and web-design packages, GIS and statistical software. In some regions, equipment or access to reliable communication systems continues to be a problem. While these are constraints that would affect assessment, they often have locally available solutions. However, IA also requires more specialized tools and related skills where specialized capacity development and training may be necessary. These include modeling and scenario analysis, Artificial Intelligence (AI), indicator development and multivariate analyses. There is an increasing trend to bundle applications in multi-functional 'expert systems' that support integrated assessment and reporting from the early stage of project conceptualization through issue identification, data collection, analysis, and product development, all the way to publishing on the Internet[17]. In many of these cases, the problem may not be simply lack of access to software (although that is common) or the lack of skilled personnel (although that can be expected with highly specialized systems), but also underlying constraints related to cultural or institutional factors.

IAR also relies on – or is composed of – a series of generic and specialised analytic methods, such as the application of the DPSIR framework in the integrated analysis of key environmental issues. Constructing indicators and indices and linking them to targets are examples of relevant areas. Because of explicit and high visibility emphasis on policy performance, integrated policy analysis methods are particularly important. As documented by the production guidelines, the policy analysis component of the GEO template changed continually from GEO-1 through GEO-3, and there is today a need both for further methodology and associated capacity development work.

Due to the large diversity of tasks, IAR requires a very wide range of scientific and non-scientific expertise. While some of the expertise is very specialized (e.g., in-depth familiarity with a particular scenario tool), there also is a need for IAR generalists who can lead synthesis. While traditional academic institutions are well suited to train specialists, the training of integrated assessment – or sustainability - generalists is not as well developed. It is also unclear what the best strategy for training integrated assessment generalists would be. Should we continue training specialists who may develop into generalists later through their career, as was the case for many current practitioners who started out in fields as diverse as applied mathematics, political science or agronomy? Is it more effective to develop specialist training in integrated assessment as part of academic programs, granting generalist degrees as some universities now do? Should we continue with specialist training but integrate generalist subjects into specialist curricula? The debate has been going on for some time and unlikely to end in the foreseeable future. In the meantime, there are experiments in all three directions and new practitioners enter the integrated assessment arena with diverse backgrounds. This has been serving us well so far, but making the transition to sustainability would require a massive pool of well-trained and dedicated generalist professionals who clearly will not be delivered without changing the current configuration of mainstream higher education.

In addition to expertise, I also include in the notion of capacity the experience of people and institutions working in an IAR process and environment. Due to the complexity of the process, academic expertise alone has its limits; familiarity with the process, conflict resolution and negotiation, the ability to make decisions under pressure, and balancing the many aspects of the science and policy process become crucial skills. Combining practical experience with post-graduate or mid-career training are promising directions for the collaboration of integrated assessment and reporting programs and institutions of higher education. GEO could, for instance, have an associated internship program for promising fresh graduates to train the next generation of assessment professionals, or an exchange fellow program among GEO CCs that would both strengthen the GEO network and contribute to the professional development of experts already involved in the program.

Availability of resources is a key precondition often bundled with the notion of institutional capacity. Depending on the definition one chooses to use, this may include staff, equipment and other tools, but most importantly it includes funding. While availability of funding is far from being a guarantee of effectiveness, lack of funding usually predetermines the level of effort that can be put into capacity development or the assessment itself.

### *6.4.2. Capacity development in GEO and distributed systems*

Capacity development in the environment has been a core function of UNEP since its inception. This has extended to most activities of the organization, including early warning and assessment. The goals of UNEP's capacity development efforts included, among others (UNEP Division of Environmental Policy Implementation n.d.):

> facilitating and supporting environmental institution-building by governments at regional, sub-regional and local levels;
> developing and testing environmental management instruments in collaboration with selected partners; and
> promoting public participation in environmental management and access to environmental information.

Most capacity development programs are cooperative efforts involving UNEP and other partners either as co-leaders or donors. In many cases, partnerships involve mutual learning, as in the case of the GEO training program developed jointly with IISD. By choosing capacity development partners strategically, UNEP can also bring in specific expertise or use partnerships to influence other organizations and audiences. Working with donors is particularly important not only to ensure funding but also to make the CDE approach better accepted.

Although capacity development was not part of the very first design of GEO, some of those involved did recognize early on that it was necessary given that UNEP would be at the mercy of CCs in terms of producing the assessment (O. El-Kholy interview). As

pointed out before, some CCs obviously need thematic and regional expertise, but few of those who need it have both. Some tried to fill the gap through short-term solutions by hiring consultants. If the only role of the consultant is to deliver a product the CC is bound by contract to deliver to UNEP, this is a stopgap measure and contributes little, if anything, to learning and continuity between reporting cycles (E. Gutierrez-Espeleta, Observatorio del Dessarrollo and S. Shresta, UNEP-ROAP, interviews). On the other hand, consultants may play a more constructive role if their terms of reference include co-producing GEO input with CC staff. Due to the extended and in-depth interactions required in the latter model, this approach may in fact be a rather effective form of capacity development. Distributed assessment systems may even offer the possibility to find consultants who are actually associates of other members of the network. Besides delivering GEO output this would lead to the added benefit of strengthening the network.

Capacity development took place in GEO in three main contexts (**Figure 10**):

> 'learning-by-doing', capacity development through participation in the GEO;
> training modules to strengthen CCs and the GEO process (referred to as "the UNFIP project" after the United Nations Fund for International Partnerships, the key funder); and
> training programs to develop IAR capacity of organizations not directly associated with GEO but strengthening the underlying knowledge base.

The first two categories of activities focus on strengthening GEO itself, including strengthening the skill base and the GEO network; the later focused on widening the circle of organizations using the GEO approach and focused mostly on the regional and national level. All of them are based on the GEO approach. Indirectly, GEO could also benefit from capacity development efforts of UNEP partner organizations, particularly those lead by the GRID centers. Some of the GRID work is directly focused on capacity development for SOE reporting on the national level and closer integration with GEO is a distinct possibility (e.g., UNEP/GRID-Arendal 1998, Simonett *et al.* 1998).

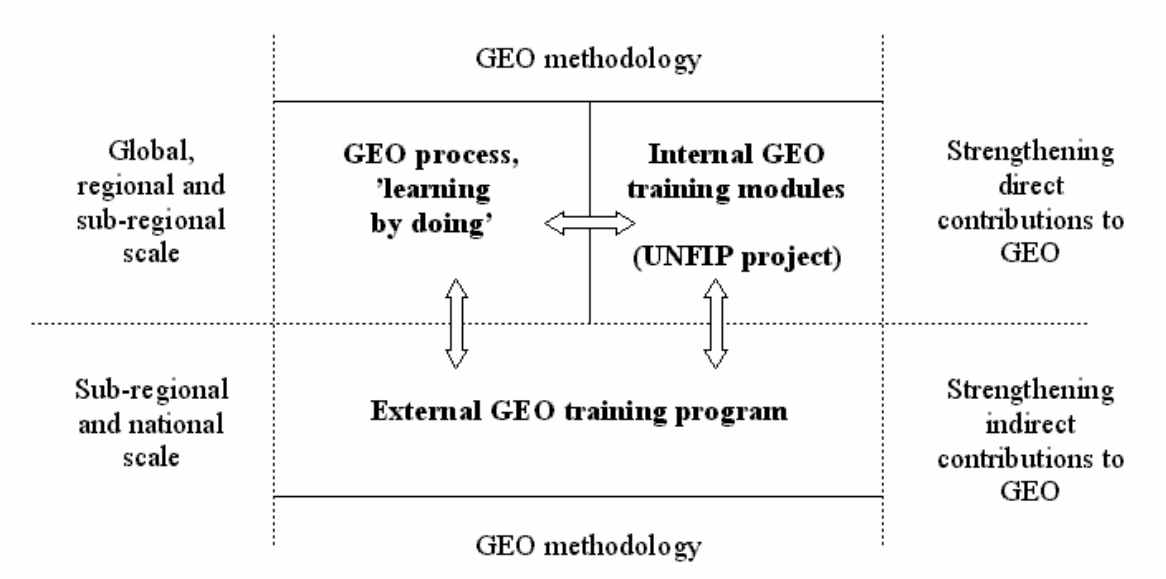



**Figure 10:** Capacity development activities in the GEO program.

Because it is an evolving system, GEO involves some learning for all participants involved, although it certainly involves more learning for some than for others. The key question with regard to learning-by-doing is whether by simultaneously trying to both produce an assessment and develop capacity, one risks achieving neither. I found some evidence from other assessments to the contrary, suggesting that capacity development taking place in the context of assessments is actually one of the ways to create impact. For example, the involvement of regional institutions in the preparation of UNDP's Human Development Report (HDR), the subsequent publication of many national HDRs, a journal and a human development forum can all be considered results but also capacity development tools of the HDR process (R. Jolly interview). In comparison to training courses, Jolly also pointed to intensive, in-depth and enduring interaction during the collaborative preparation of reports as a mechanism through which effective learning can take place:

> "Emphasis is on people learning by doing, coming back with drafts to share, themselves saying, my God, that's a fantastic idea, so they're feeling a certain sense of good competition with what some other countries have done. They're

> saying trouble is we haven't got the sort of data we need, and they find some other country didn't have the data, but in the last six months have done a deal with the statistics office or with some environmental groups to check the data etc. We used to say in UNICEF we don't want to write reports, we want to set in motion a process."

There is also support for the view that capacity development tied to actual assessment systems may be more successful and effective in creating actual capacity than detached programs:

> "Capacity building is much more strongly fostered when the process has some reality tied to it, because if the process is somewhat arbitrary and not connected to reality, there is less assessment learning that's going on. And to me doing assessment for the sake of doing assessment, for what it would become, would be less than interesting, particularly when you are doing something that is so difficult and so nebulous to do anyway."
> (F. Carden interview)

The possibility that learning-by-doing works does not mean that it unconditionally does. Throwing participants into a network and letting them learn without targeted assistance is less effective than having a parallel training program that provides assistance in critical areas, based on needs agreed upon both by capacity development recipients and providers. As Carden pointed out, rather than completing exercises for the sake of getting a certificate and pleasing donors and instructors, it is more motivating and satisfying to apply newly acquired knowledge to a real program with real stakes. Assessments produced by networks, like GEO, may build on comparative strengths of their members. For instance, centres with complementary needs and expertise, say, one having regional expertise in the Indian Ocean region and another having thematic expertise in indicators or scenario analysis could be twinned and enabled to work together over an extended period of time. Besides increasing the indicator or scenario analysis capacity of one and refining the reflection of small island issues in the methods of the other, twinning would

also grow the cohesiveness of the network. Twinning arrangements have been discussed in GEO, but not yet implemented. There are other possibilities for assisted learning-by-doing. The point is not to assume learning-by-doing will work only by going to meetings and trying to deliver assessment products based on available guidelines, but to invest in understanding real needs and custom design sustained and substantive assistance to facilitate the development of real capacity.

The learning-by-doing approach alone, however, has significant limitations, as also pointed out by some of my interviewees (e.g., A. Hammond and J. Bakkes). Currently capacity development activities are still not integral parts of the annual programs of CCs but more short-term add-ons. Because of the complexity involved, integrated assessment and reporting is a major challenge from the didactic point of view. In addition to learning-by-doing, real capacity development would require repeated and in-depth exposure of the same pre-selected trainees to methods, tools, case studies and other modules and to provide detailed feedback on progress. It also requires custom-built capacity development materials and qualified trainers. On the practical level, this requires time, staff and resources – all of which were in limited supply during GEO-2000. Due to short-term terms of reference and limited existing capacity, some CCs hired consultants to prepare outputs, acting more as subcontractors rather than genuine members of a network; hardly a sustainable strategy to develop capacity even if understandable given the need to fulfill the organizations' contractual commitments to GEO.

There is also an imbalance in the sense that most thematic CCs are in developed countries, potentially fuelling suspicions about who speaks for global humanity. Bakkes *et al.* (1998) noted that equity considerations would require that in some thematic areas, developing countries build up centers of excellence. Gradually shifting from North-South capacity development (where the North – my shorthand here for developed countries – is invariably considered the source of legitimate knowledge) to South-South and even South-North. Some South-South cooperation and knowledge transfer has taken place in GEO, for instance the possible use of data knowledge developed for Latin America to other developing regions, but the mechanisms are still too ad hoc. Of course, one has to

remain realistic. While most Southern institutions certainly can provide valuable and legitimate perspectives, far less have the in-depth and substantive expertise and resources that would make them competent agents for capacity development. The decision to strengthen this base is a strategic and political decision that has been taken in agriculture (through the Consultative Group on International Agricultural Research network) and a similar arrangement is being proposed for the energy sector (Goldemberg *et al.* 2002). There are not yet signs of similar intentions to connect and increase the capacity of the institutions dealing explicitly with the environmental components of sustainability.

In order to address the needs of selected CCs, early in the GEO-3 process UNEP started a capacity development initiative that became known after its donor, the United Nations Foundation for International Partnerships, as the UNFIP project. The project was designed to target particular capacity needs of selected CCs in developing countries and transition economies. The United Nations Institute for Training and Research (UNITAR) was invited to serve as lead executing agency. As capacity development was not yet fully designed into the GEO process when opportunity for the project arose on short notice, up-front participation in conceptualization and needs assessment has been rather limited.

The third element of the GEO capacity development effort to mention here is the training program aimed at external organizations, primarily on the national and regional level, to carry out integrated assessment and reporting. Being a high-level system but relying on data and knowledge created on finer scales, GEO (and UNEP) has fundamental interests in these monitoring, assessment and reporting systems having adequate capacity. GEO also is interested in facilitating the adoption of tools, frameworks and processes, synchronized to the extent possible and necessary with the global approach. Building on the GEO framework, this training program provides an overview of the integrated assessment and reporting approach including underlying conceptual frameworks, institutional design choices, entry points to the use of analytic methods, and selected tools, such as indicators or scenarios. The training materials for this program have been developed through an iterative process and used in many regional workshops around the world (Pintér, Zahedi and Cressman 2000; Universalia 2001).

Although a comprehensive evaluation of the GEO training programme is yet to be prepared, having taken part in both developing the program and delivering many of the training courses I offer some qualitative observations[18]. As pointed out by Fred Carden, capacity development efforts taking place in the context of actual, real-life assessments are seen as more relevant and valuable to participants. While most participants in these training programs are potential contributors to GEO, more importantly most sessions had several representatives of new or well-established assessment and SOE reporting programs in their countries. Because of the interactive and practice-oriented nature of the program, there were examples when built-in exercises focused the strategy or more technical aspects of the integrated assessment framework were turned into real-life planning for national programs. The need for assessment and reporting capacity building has been emphasized for example in the Libreville and Douala Declarations (Universalia 2001, Appendix VIII) drafted and endorsed by participants of training workshops in the Central African region. While they were generally well received, stand-alone training events are less likely to lead to lasting capacity improvement than programs that are more closely connected with (or lead to) national assessment and reporting activities. Even if some participants were directly involved in setting up national assessment programs and found the training a useful starting point, UNEP had neither sufficient resources nor, at least initially, focused strategy for sustained follow-up. Some of the required follow-up such as national scale organizational development is probably on the boundary of UNEP's and some other organizations', such as the United Nations Development Programme's (UNDP) mandate, creating complications but even more than that opportunities for cooperation. While some training sessions were co-hosted with other organizations like the IUCN – The World Conservation Union, UNDP or regional groups like the Indian Ocean Commission (IOC), extending the cooperative approach beyond single training events and beyond co-hosting is a challenge for the future.

Funding needs to be highlighted as an issue with overall bearing on institutional capacity in GEO. As Bakkes *et al.* (1998) observed, the GEO network is maintained through incremental funding, reflecting the reality of UNEP's access to resources and the nature of donor commitment. This is problematic in the sense that it makes planning horizons

short and limits the ability of CCs to fully integrate GEO into their work programme. It also makes sustained commitment to capacity development efforts difficult.

### *6.4.3. Implications for effectiveness*

In order to contribute to an effective GEO process that includes capacity development activities, the CCs themselves need to have adequate capacity. As pointed out above, capacity is still an issue for some members of the GEO network, and resolving them requires a long-term, institutional approach.

One possible systematic way to tackle these problems in the future is through more attention on the part of UNEP to mutually agreed formal and long-term designation of CCs with a broad definition of roles and responsibilities and the possibility of task-specific MoUs. This would not only avoid potential capacity problems at the CCs, but strengthen GEO's governance and participatory character.

| Recommendation |
| --- |
| Framework agreements confirming CC status and providing a basis for task-specific, short-term MOUs and funding arrangements could strengthen the assessment network and the participatory aspect of the assessment and reporting system. |

It is useful to contrast GEO with the CGIAR network that is sustained through structural funding. Although there are important differences (e.g., CGIAR has a narrower sectoral focus, it is more a scientific network, GEO works more on the science/policy interface, CGIAR is well established, GEO is more recent), the two are similar in terms of ambition and cross-scale, distributed design. While environment / development problems are hardly less challenging than those associated with agriculture and food security, there is no institutional network dealing with these issues that has the required geographic and thematic coverage comparable to CGIAR's coverage of agricultural research. Politically controlled government agencies, university departments, and predominantly Western-based non-governmental organizations are hardly a substitute. It is striking that at the time of political and academic interest in creating a centralized world environmental

super-organization (e.g., Biermann 2000), there is virtually no discussion about investing in a globally distributed network of organizations with the ability to bring global messages to local audiences and also serve as an interface to bring relevant local knowledge to bear on global assessment systems and policy. A comparatively weak global environmental watchdog would find it hard to create a strong network, but that's only half of the point. The other half of the message should be that strengthening the central agencies of global environmental governance should be accompanied by strengthening – often creating – more substantive and more sustainable capacity in the regions, North and South.

Capacity development for assessment and reporting can be an objective in and of itself, but its ultimate goal is to make assessment and reporting systems more effective. Although my primary interest here is not in the effectiveness of capacity development *per se*, but its influence on the effectiveness of assessment and reporting, the two are inextricably linked. Any one of the capacity development efforts undertaken in the context of GEO adds to the program's 'culture' and influences the perceptions, experience and attitude of its participants. Capacity development can do a lot of damage if its substance is wrong, but even if it is right, poor organization and execution can harm the overall assessment's image and perception in the eyes of participants, audiences and funders. Conversely, a capacity development program with strong participant support can be an important asset, seen as a sign of long-term interest in the process and commitment to the cause of the participants. This has the potential to increase the overall assessment's legitimacy.

One of the key questions that arose from my interviews had to do not so much with any of the criteria of effectiveness, but the overall adequacy of the level of effort put into capacity development. Some thought the investment and management effort was so low that it could not feasibly lead to a major and sustainable impact on the organizations involved, or presumably GEO itself. Consider the following statement on this issue:

> "You could take data and say OK, what does it really mean to be able to have better national and sub-regional data and to have strong centers in the regions that can really do analysis. Well that takes some real training it takes some equipment, some exposure to standards, some practice, some real projects that people do, but that takes real money. So I basically think that without that investment they are not going to in fact increase the capacity very much."
> (A. Hammond interview)

Even if the investment made through the GEO program is lower than necessary, GEO is not the sole program investing in capacity. Other capacity development programs supported internally, other multilateral donors or national governments may well complement whatever is provided through GEO. This assumes at least some coordination between donors and donor programs – a rather generous assumption. Even if resources are available, there is another argument regarding the ability of some CCs to absorb and manage large amounts of funding. As Veerle Vandeveerd observed in my interview:

> "One thing we did when we did GEO was really look at absorbing capacity of the CCs and try to build on that capacity in a gradual way. In networks you always have to look that you don't give too much so that you overload the system, but you have to give enough."

The operational variable of interest is of course not funding but capacity and as previous discussions have shown, capacity development often needs longer-term commitment. The issue is to find funding strategies that serve the needs of workable capacity development strategies well. One-time capacity development efforts and injections of large amounts of funding beyond the absorptive capacity of an organization may be as problematic as under-investment. The best funding strategy may involve adequate investment, longer-term commitment, a phased approach, and frequent mid-course, participatory evaluation and feedback to modify funding strategies as necessary.

Recommendation

In order to produce sustainable results, capacity development taking place in the context of IARs should have a long-term perspective and funding strategy, and take existing institutional capacities, including the ability to absorb funding and capacity development programmes, into account.

Although capacity development is firmly included among GEO's strategic objectives and structure (e.g., the Capacity Working Group), in practice most capacity development efforts, in particular the UNFIP project and the GEO training workshops were concentrated efforts with little follow-up.

At some point externally funded (and often conceived) capacity development programs come to an end. In a successful case, either capacity is no longer a constraint or responsibility for capacity development is taken over by other organizations. This could mean shifting at least some of the responsibility (and resources) for capacity development to developing country institutions, or in the case of the GEO network possibly to developing country CCs (i.e., South-South cooperation as argued before). Capacity development carried out by in-country or in-region organizations could increase both saliency and legitimacy. The need is clear that in order to make particular assessment and reporting strategies more widely and quickly introduced, requires moving from capacity development and training to broadening the range of qualified trainers who can lead particular capacity development activities, such as training workshops and follow-up. However, integrated assessment and reporting being a new and evolving field there are only few people qualified to serve as trainers; it is even far less clear who could serve to train trainers (P. Raskin interview). Developing train-the-trainer strategies would need to build on several types of knowledge, including: (a) experiential knowledge gained through the delivery of previous training programs on IAR and related topics; (b) didactic and curriculum development knowledge; (c) substantive knowledge of IAR, including its institutional dimension. The importance of developing a more distributed capacity development and training base has been recognized in GEO, but as of late 2001, a strategy is yet to be implemented.

| Recommendation |
|---|
| IAR could increase their effectiveness by broadening the circle of professionals qualified to lead capacity development programs through train-the-trainer activities and paying particular attention to the need for in-country training expertise. |

It seems quite obvious that training materials need to be developed and tested before methods and materials for training trainers are initiated. A possible (and ambitious) configuration of training programs, including a train-the-trainer component are shown in **Table 6**. Due to the complexity and inherent uncertainties of the subject and its delivery in diverse geographic and cultural contexts, building in feedback loops, pilot testing, evaluation and revision are essential.

**Table 6:** Possible phases, participants and roles for a train-the-trainer program.

| **Training phases and activities** | | | **Participants and roles** | | |
|---|---|---|---|---|---|
| Training program development | Train-the-trainer program | Training delivery | Trainers of trainers | Trainers | Trainees |
| Conceptualization | | | ✔ | | ✔ |
| Training material development | | | ✔ | | |
| Pilot testing | | | ✔ | | ✔ |
| Evaluation | Conceptualization | | ✔ | | ✔ |
| | Train-the-trainer material development | | ✔ | | |
| | (Pilot testing) | | ✔ | ✔ | |
| | (Revision / Finalization) | | ✔ | | |
| | Delivery | | ✔ | ✔ | |
| | Evaluation | Delivery | | ✔ | ✔ |
| | | Evaluation | ✔ | ✔ | ✔ |

Because of the increased consequences and the multiplier effect of training trainers, understanding the characteristics of what makes an individual a qualified 'trainer of

trainers' and what makes good train-the-trainer materials and strategies is critically important (e.g., Measurit 2001). At least in terms of capacity development, GEO has really been a learning-by-doing project for the small number of experts involved. There are few shortcuts: while the pool of training specialists is fairly large and the number of integrated assessment practitioners is also growing, the pool of individuals with expertise and experience in both is tiny. In addition to training and integrated assessment and reporting expertise, developing train-the-trainer programs would also benefit from didactic expertise as applied to teacher training. Given the diversity of expertise and experience required, train-the-trainer program development is inherently a participatory group project with a role not only for experts, but also representatives of future audiences.

Both science/policy and scientific assessments must be subject to peer review as a quality control mechanism. Peer review increases complexity, time and resource requirements, but without it, credibility suffers. Peer review (though not anonymous) has been used to improve the GEO training manual in several iterations. This has been less the case with the training modules prepared under the UNFIP project, mostly for lack of time and perhaps planning. There is no reason why peer review and other quality control mechanisms should not be used with the same rigour with regard to capacity development programs and thus strengthen their credibility as well as potentially saliency. VanDeveer and Dabelko (2001) also point out that capacity development programs taking place in diverse contexts should identify suitable institutional forms and mechanisms of 'organized scepticism' that serve monitoring and evaluation purposes. In addition, credibility may also benefit from having external advisory bodies representing scientific expertise, policy knowledge and practitioner experience with the design of capacity development programs.

| Recommendation |
| --- |
| In order to strengthen their saliency and legitimacy, training and train-the-trainer strategies, programs and materials should be developed in close collaboration with would-be trainees and subject to peer review. |

Multi-scale, decentralized capacity development that follows a broadly comparable framework and strategy at different locations is inherently challenging, because of major differences between the underlying realities and perspectives of participating regions and organizations. The differences may be in the domain of, say, cultural norms and thus affect local attitudes with regard to participatory methods; technical and thus affect the diffusion of necessary computing knowledge; or institutional and thus limit what organizational forms would be appropriate for hosting and supporting the assessment and reporting process. The GEO training program materials have been developed with a compromise that involved a generic training template. This made sense as there indeed is a generic template underlying GEO and because it probably reduced the up-front cost of developing customized products. On the other hand, it only defers customization that becomes necessary prior to the program's launch in any given region.

Capacity development can add much to the awareness of integrated assessment and reporting. While it could strengthen a given assessment like GEO internally, it can also be part of a mobilization process in partnership with other global organizations and regional or national partners (R. Jolly interview). Some of the GEO training project received considerable national press coverage, but this was not systematic; it depended on the affinity of the local organizers with the news media.

The long-term success of GEO will be defined not only by the success of its products, but also by the acceptance and successful application of the GEO approach. GEO can and in some cases already does, play an important role in strengthening assessment and reporting capacity. That is bound to help not only other organizations set up their own assessment programs, but there will be positive feedback and better awareness of GEO. Although the GEO capacity program has scored some successes, it needs to evolve. Capacity development could be better integrated into the GEO process with longer-term perspectives, more substantive investment, improved quality, and more up-front involvement of recipients of capacity development, particularly Southern organizations. In addition to the capacity development program aimed at external organizations, GEO could strengthen capacity efforts aimed at its CCs that, with capacity development

efforts, twinning arrangements and similar efforts could in time turn into a genuine network of highly competent organizations able to connect global and local perspectives.

## 6.5. Communication

### *6.5.1. Why communication?*

The power of packaging and presentation in maximizing the impact of information has long been known and applied to promote political ends (e.g., Ellul 1973). Communication is an essential, yet often overlooked design element of many science-based assessment systems. To some degree, this is understandable. Scientists and science assessments are held to account first and foremost for the scientific value of their work and less so for its interpretation, use and impact[19]. In fact, according to the purist view, going beyond the mandate of finding truth, strictly speaking, makes maintaining science's position as detached and objective observer harder and may even draw criticism from peers. Formal expectations are usually fulfilled by publishing scientific papers, books, and final reports to funders. This is in contrast with the conceptualization of assessments as interface and communication processes where generating knowledge and getting the message across are equally important. There is often a common – but incorrect - assumption that decision-makers will actively seek out information they need to make more effective decisions. Webber (1991-92) emphasized that this is not the case and recalls the conclusion of the knowledge utilization literature that knowledge has to be actively brought to the attention of user communities. In addition to supporting the same point, the 1999 science-policy dialogue on sustainability indicators also emphasized the increasingly important role the mass media has played in environmental communications over the last decades (Bettelli and Pintér 1999).

The need for more active communication is supported not only by theoretical arguments but increasing actual demand for the results of assessments. This is particularly the case for assessments that are commissioned by political bodies to answer specific policy questions, where presumably the audience has *a priori* interest in the results. This seems

to be the case of some products of RIVM or the EEA that are truly science-policy boundary products and processes (R. Uhel and R. Maas interviews). With the increasingly political nature of many environmental issues, settling for less than effective communication is less and less acceptable. As Denisov and Christoffersen (2001) put it, "Even high quality information that is not communicated has little or no possibility to reach decision-makers. It thus remains in ivory towers of laboratories and think tanks and results in no impact. While the quality of the content is essential, active communication is another vital ingredient that needs to be added."

Decisions affecting the environment are made on many scales and by many actors. In principle, distributed, multi-scale assessments respond not only to diverse issues, but also to the needs of multiple, equally distributed audiences. Can we reasonably presume that issuing scientific publications alone will get messages across under such conditions? Are not policy-makers, educators, corporate chief executive officers and others all different when it comes to information needs and channels of communication? Should not assessments offer more diversified communication strategies to reach them?

In order to realize the magnitude of the challenge, one need only compare the marketing and advertising budgets, organizational professionalism and staff expertise available to corporations with the resources and mechanisms available to inform the public about environment / development issues. Advertisers, political campaigners and lobbyists promoting special interests of many types use the most diverse tools of communication to ends often diametrically opposed to causes so carefully (and so esoterically) argued by scientists. What is available as communication strategies, methods, resources, tools and efforts to match the influence and impact of an overwhelming and very effective propaganda machine promoting essentially unsustainable ways of life? I believe this is a central, yet still underestimated and unaddressed issue in the transition to sustainability.

Designing communications strategies for integrated assessment and reporting systems is not the same as designing a product marketing campaign, even if the two may be competing for the same attention span of often the same audience. There is a significant

body of knowledge related to communication and marketing strategies that integrated assessment and reporting systems could tap, but they do not. The point has been noted in the GEO training manual that identifies some key communications strategy issues assessment design should address (Pintér, Zahedi and Cressman 2000)[20]:

Defining communication roles

Although the need for communication strategies in science based assessments does not mean scientists must do it, they must play a role. "The people who create environment reports do not have to be communications experts, but they should be able to make strategic communication decisions and guide the communication experts who prepare and deliver messages" (Ibid.). We must ask what is the role for the coordinator of the assessment (e.g., UNEP), for associated producers (e.g., GEO CCs), for external partners, for hired experts?

Identifying and characterizing audiences

Although the question of audience is important throughout the assessment process, clarity on this is particularly important with regard to communication. The question is not only who the audiences are, but what are their interests in the issue, how do they obtain information and who are key individuals who should receive information?

Language

It is critical to ensure that the information is made available in a language most commonly understood in a given region. There are cost, time and impact considerations. Is there an adequate budget for preparing translations? How much of potential impact is lost by spending time for translation? How many people speak a second language in which the results of the assessment are already available?

Working with the media

Working with the media – science and policy publishers as well as mass media – is key. What elements of the media do key audiences commonly use? Who are the key media partners? How can they be engaged and used most effectively?

Branding and trademarks

What are the characteristic features that can be easily recognized and create associations with the assessment throughout a range of diverse outputs, published in different forms, at different locations and over extended periods of time? What are the icons that can help create and illustrate key messages in cover page-friendly ways?

Process construction

Many global scale, integrated assessments have long life-cycles ranging up to a decade. In such cases, it would be highly inefficient to concentrate communication efforts around peak events; a phased approach would be more effective. How could a phased communication strategy be built? What are the implications for knowledge generation, staff, resources, and products?

Product design

Although some IARs, including GEO concentrate on low frequency, key products, this does not have to exclude the release of carefully prepared intermediate products. Can the assessment release early, partial results or regional products? What side products are necessary to reach particular audiences in particular regions? What are the implications for knowledge production, resources, and expertise?

Depending on the approach to communication challenges, one can differentiate between a traditional and flexible communication model, implying that the flexible model is more effective and suitable for the purposes of GEO-style integrated assessment and reporting systems (**Table 7**). There is a strong parallel between the flexible and the earlier mentioned constructivist model of communication.

Communication requires careful planning and integration into the assessment and reporting strategy from an early stage. It is typical to find communication as an add-on once the assessment is complete (e.g., organizing a press conference, issuing a press release, drafting a mailing list for final products, building a website are some common elements).

**Table 7:** Communication planning in a traditional and flexible model (Modified from Pintér, Zahedi and Cressman 2000).

| Traditional model | Flexible model |
|---|---|
| Management and experts decide there is an issue. | Management and experts decide there is an issue. Build a communications plan. |
| Determine position and performance. | Create an advisory group: multi-stakeholder, collaborative, solution seeking. Set long-term goals. Refine goals. |
| Select the audience. | Identify stakeholders and audiences. |
| Decide what people should know. | Determine their knowledge, beliefs, opinions, where they get information and who they trust. |
| Select key concepts, messages and decide on form and content. | Research what communication is being done by various parties now. |
| Prepare the messages. | Develop first message, based on research. Build on existing credible messages. |
| Produce material that reflect their opinion. | Pre-test message. Does it make sense? Train communicators in workshops. |
| Publish, disseminate, train and lobby. | Deliver messages. Help others to deliver compatible messages. |
| Determine success without formal evaluation. | Consult, survey, and determine effectiveness of messages. (This testing process establishes a feedback-loop.) Refine message, based on feedback. Modify messages. Develop other messages as necessary. Retrain communicators as necessary. Advise others on their messages. Continue to deliver and deliver messages over time. |

This also was observed by Denisov and Christoffersen (2001), who note that "many information systems and publications, at least in the public domain, seem to be designed

with no usage perspective in mind, on a completely supply-driven basis". If it is supply driven, it is enough to 'tell' audiences about results, at best produce an executive summary, with little interaction with audiences before. Communication, however, could happen throughout the process from an early stage and include not only and maybe not even primarily the communication of final results, but working with key audiences in identifying critical issues, having them review and comment on drafts, interact in meetings and so on. The same strategies that are discussed under the production and participatory aspects of the assessment in fact, also serve communication purposes. I emphasize the need for meaningful dialogue versus one way 'transmission of the message', communication as part of an open, learning process rather than something that has a static place in the assessment strategy and provides little feedback. Although this model does not exclude the need for a designated communication team and a focused strategy, it shows that communication is an open process, taking place in many stages, on many levels and in many ways.

### *6.5.2. Communication in GEO and distributed systems*

GEO communicates through a suite of products and through the interactions taking place between producers and audiences during the assessment process. Formal responsibility for coordinating communications activities rests mostly with UNEP, although CCs may also play a role on a less frequent basis. The communications strategy applied in GEO to date is closer to the traditional communication strategy shown in **Table 7**, though it incorporates elements of the flexible model. There is relatively little up-front and formal involvement in crafting a communications plan, even though the CC network would be well positioned to contribute. Time and resources are some of the real constraints, but awareness of the importance of the issue and bureaucratic hurdles may also play a role. There has been also little up-front effort to assess the actual information needs of key audiences and there are no formal provisions to test and fine-tune key messages before full-fledged release. These are untapped opportunities to be considered more intensively in the future. Although through the GEO user survey there has been an effort to obtain feedback on the usefulness of the information, the evidence gathered on the success of communication strategies is limited. In the absence of objectively measurable targets, it is

hard to determine to what degree particular elements of the communications strategy work, if at all.

**Table 8:** Phasing of product-related communication efforts in GEO.

| Phase | Strategic element |
|---|---|
| Pre-release | Product design: layout, copies, design elements, colors, icons and logos, etc. |
| | Vital graphics: charts, maps, models, etc. |
| | Editorial work: professional science writer / editor to ensure readability |
| | Languages: translation available at least in the main UN languages; usually not available at main press release |
| | Executive summary: foreign language versions available at main press release |
| | Websites: products available as PDF and/or html documents through several regional UNEP websites |
| | CD-ROM: main or regional products, statistical annexes etc. |
| Release | Press-kit: press release, executive summary, letter from senior officials etc. |
| | Global and regional press release events |
| Post-release | Presentations to professional and general audiences |
| | Follow-up interviews |
| | Evaluation and feedback gathering |

The GEO products, as described earlier, include a range of print and electronic publications. The publications are released at the end of a production cycle, and traditionally have attracted most of the communication effort and resources. Apparently, the investment paid off. Quoting T. Brevik, UNEP's Director of Communications and Public Information, the GEO user survey notes that "GEO-2000 has generated the most media inquiries of any UNEP publication and that, as of June 2000, was still major currency with media across the globe" (Universalia 2001). Some of the product-related communication strategies apply to the pre-release, others to the release and post-release

phases of the assessment, with relatively little active promotion between peak events (**Table 8**). There are, however, regional and thematic products, some of which also have their own approach. The Pachamama Report, the GEOs for small island developing regions and of course the release of GEO translations to various UN languages all attract media attention, although the attention is not necessarily global. The challenge for GEO – as well as other distributed assessment and reporting systems – was to coordinate communication efforts involving different products, audiences and regions to ensure messages are cohesive and facilitate local action that will also further global goals.

With regard to internal participation, CCs are natural allies of UNEP in communication. Given that CC status does not automatically entitle an organization to speak directly for GEO, each CC needs to be invited to play a communication role (e.g., by UNEP-HQ or the UNEP regional office; of course a CC is free to speak for its contribution to GEO). While this helps control what is being communicated, where and how, it probably constrains CCs from taking initiatives. Decentralizing not only assessment but also communication activities may be an important asset of distributed systems, but this has to be balanced with ensuring cohesion, maintaining quality control and credibility.

Besides communicating assessment results to external audiences, global distributed systems face a particular challenge with regard to internal communications within the network itself. In fact, networks can be conceptualized as organizations "characterized by voluntary, reciprocal, and horizontal patterns of communication and exchange" (Keck and Sikkink 1998). CCs are separated not only by physical distances, but also by differences in capacity, expertise, culture and work style. Making diverse groups cooperate on different aspects of what is supposed to be a cohesive plan and shared products requires smooth data, voice and, increasingly, video communications. With regard to the latter, as of late 2001 good quality videoconferencing is still expensive and in the absence of broadband connection web-camera based video is rather poor quality. High quality video is an option only for selected CCs and UNEP offices with dedicated facilities, although with technological progress and cost the situation may quickly change.

Having resolved basic access problems that still constrained some CCs during the GEO-2000 process, GEO is in principle ready to implement an Internet-based internal communication program, even if electronic communications alone do not meet all needs. Besides operating through email and attachments UNEP has commissioned the development of the GEO Support System (GEOSS), a restricted access Internet utility to facilitate data sharing and communication among network members. Although functionally adequate, GEOSS' use up to mid-2001 remained very low. Rather than concluding that such systems have little potential as an internal communication tool in IAR, I find this to support a better understanding of internal communication needs. Since the start of the Internet revolution in the early 1990's we have learned much about information system design, but there is more to learn about human choices and preferences with regard to electronic communication, particularly under increasing information overload.

### *6.5.3. Implications for effectiveness*

My observation is that IAR systems should have a communication plan / strategy, preferably from the early stage of the assessment process. In a world of increasing information overload it can no longer be assumed that messages get through. Formal communication strategies are needed with a clear understanding of audiences, purposes, detailed objectives, budgets, staff, products, activities and time lines.

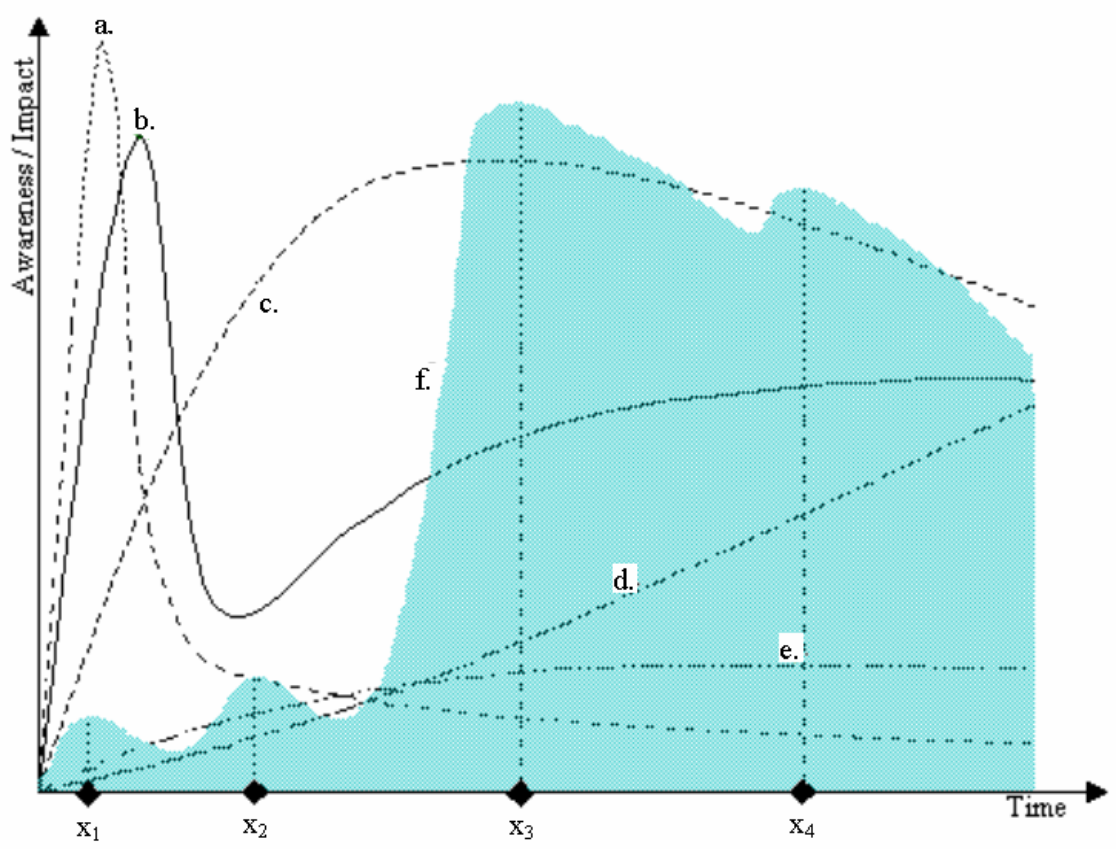
| Recommendation |
| --- |
| Overall plans for IAR should include detailed communication strategies and make provisions for their successful implementation. |

Some aspects of communication seem to be preconditions of all four criteria of effectiveness. Publishing the assessment in local languages, using plain language when communicating with non-expert audiences, and clear graphics are basics without which no report will be understood. The large number of unintelligible, esoterically written, poorly designed reports proves that these limits are not always understood. Communicating complexity simply but without loss of critical information is a real

challenge. Some tools, such as indicators and models can help simplify the communication of complexity. However, indicators and models required by experts may need to be further simplified and condensed when communicated to lay audiences. Communication through meta-models that "simplify the results of much larger models and present them in an interactive way to target groups" is a relatively new challenge, now that complex models and modeling capacities are more readily available and applied (Bakkes *et al.* 1999).

Among the effectiveness criteria identified, the most direct relationship of communication is with awareness. Denisov and Christoffersen (2001) suggested that the impact of information over time could be charted as a function of communication strategies. They identified five communications strategy 'archetypes' and accompanying impact dynamics (**Figure 11**). Although the relationships sketched out are very general and await empirical testing, I find them a useful starting point to study GEO's communication strategies and their possible influence on awareness over the life-cycle of the assessment. Because the authors worked on a very general level, the relationships are simplified and applying their logic to a real life system like GEO leads to more complex (and ambiguous) dynamics. I propose to re-label the Y-axis from the original 'impact' to 'awareness'. Awareness is an intermediate attribute that is easier to associate with the process and products of assessment. On the other hand, as earlier shown, impact is a cumulative outcome of many forces and attributing it to a particular information system can be highly problematic.

The shaded area on the graph shows possible changes in general awareness of GEO over time, as a combined effect of all associated communication strategies. In the absence of empirical data this is only a working hypothesis, but I find it useful in identifying key stages in the dynamics of impact. The four inflexion points marked from $x_{1\text{-}4}$ indicate those landmarks in the GEO process that may have major influence on awareness. Compared with the archetypes originally suggested, the awareness dynamics of GEO were more complex. As a minimum, one has to keep in mind that due to regionally different communication strategies and processes, GEO's awareness dynamics may




a. Instant publicity (news, short-run advertising)

b. Publicity with strengthening evidence (climate change debate)

c. Continuous persuasion (long-term advertising, public relations, propaganda)

d. Participatory approach (integrated assessment, post-normal science)

e. Pure research (ivory tower, minimum publicity)

f. Shaded area: GEO impact dynamics

$x_1$ GEO launch

$x_2$ regional consultations

$x_3$ launch of major report(s)

$x_4$ launch of foreign language translations

**Figure 11:** Working hypothesis of the impact of various communication strategies over time (modified after Denisov and Christoffersen 2001; not drawn to scale).

change from one region to another. They may also change from product to product and from stakeholder group to stakeholder group.

The first point on the shaded area, $x_1$ indicates the publicity surrounding the launch of the entire assessment process when goals, participants and thematic foci are introduced to expert and general audiences. This is followed by a period of lower level publicity where internal guidelines and protocols are developed, tasks are assigned, issues identified, and initial drafts developed. The second point, $x_2$ indicates regional consultations and peer reviews of drafts that create awareness at least within professional communities or broader audiences, depending on the scope of participants. This is again followed by lower levels of publicity, when drafts are corrected and finalized, followed by editorial work and production of the main reports. Point $x_3$ signals the launch of the main GEO report with global publicity. This is the communication effort in which most resources and effort are invested. Ideally, these efforts would take place not only around the GEO launch, but continuously, partly in response to unforeseen emerging issues and information needs, and partly to put certain issues on the agenda as information becomes available in the GEO assessment. With regard to the latter strategy, an increasing number of organizations pre-release some sections of major reports (e.g., the World Resources Institute (WRI) issued a preview of the 2000 World Resources Report (World Resources Institute 2000)), presumably to build expectations and to increase publicity through an extra media event. There seems to be a trade-off between concentrating resources and efforts around a major media event, and thus creating a large 'spike' on the awareness graph, versus spreading communication efforts more evenly.

| Recommendation |
| --- |
| Assessments should weigh the tradeoffs associated with having evenly distributed and diverse versus concentrated communication strategies. |

Point $x_4$ indicates the launch of subsequent foreign language translations that may have their own publicity campaigns and events, though on a more limited scale. The GEO

graph covers one GEO cycle, but as GEO is iterative, at some point after $x_4$, the curve would restart with the launch of the next GEO cycle, presumably from a higher level of awareness and assuming there is not a significant data overload effect. As pointed out by some of my interviewees, it may take well over a decade for even a well-publicized global reporting program to develop a stable constituency and awareness (A. Hammond and G. Golubev interviews). In conclusion, spacing communication efforts requires striking a balance between a small number of large events with huge publicity versus smaller, but temporally more evenly distributed efforts. GEO is probably closer to concentrating its efforts around launch events, and it could probably benefit from communication efforts that take place earlier, during the production process.

Although communication efforts do not involve creating new knowledge *per se*, they can bring information developed through the assessment to the attention of fora where the information is needed and thus make the information look more salient. Without this, information may remain buried in the assessment not realizing its potential. This may involve some repackaging, reemphasizing and most importantly recognizing the linkages between issues and pieces of the assessment.

As Jasanoff, Miller and MacCarthy (1998) in the GEA project observed, the credibility (and I would add legitimacy) of scientific knowledge is influenced, among others, by who is communicating with whom. They call attention to the role the recipient's perception of risk plays in accepting or rejecting a particular knowledge. For instance, quoting VanDeveer (1998) they point out that the acceptance of Western knowledge by Central and Eastern European countries in the 1990s as legitimate could be influenced by the perception that rejecting knowledge could risk access to other benefits, such as foreign aid or EU accession[21].

In comparison with some of the narrower science assessments, the audience of GEO is broad and diverse. The challenge is shared by other inventory style operational assessment and reporting programs, such as the European Environment Agency's (EEA) regular reporting on Europe's environment. The overall conclusion is that effectively

responding to diverse needs requires an up-front communication strategy developed as an integral part of the assessment system. The strategy needs to be cognizant of regional communication needs and contexts, policy cycles and landmark events, and respond through products, process and institutional arrangements with network partners and other ad-hoc partners as necessary. The challenge in GEO is increasingly to balance the diversity of outputs with strategic impact.

## 6.6. Data and indicators

### *6.6.1. Why data and indicators?*

Integrated assessment and reporting is a data intensive exercise. This is not to say assessment is primarily about data; in fact, it should be information demand, not data driven. However, without good quality, relevant data and indicators, the assessment loses not only valuable communication tools, but also credibility and the ability to measure progress towards sustainability goals and objectives. Data collection, processing, storage, retrieval and analysis are all influenced by governance, participation and other design elements identified in the GEA project, but beyond that there is a set of specific issues and challenges that require considering data and indicators a design elements on their own. In fact, most strategic analyses recognize the central role of data and indicators in environmental information and assessment systems (Meadows 1998, European Environmental Agency 2000; Parris, Zracket and Clark 1998; National Council for Science and the Environment 2001, etc.).

Assessment and reporting faces particular challenges and design choices associated with data and indicators, as discussed by a vast literature (e.g., van Woerden 1999, Moldan and Billharz 1998, International Institute for Sustainable Development 2001). There are particular challenges affecting multi-scale assessment systems. Some of the problems are related directly to the availability, quality and structure of data itself and the difficulties associated with aggregating data for large spatial units. Aggregation can be seriously constrained by the lack of harmonized data for component regions or the aggregating data

for spatial units – e.g., ecosystems - whose boundaries do not coincide with monitoring and data collection units. Aggregation, while often useful and necessary to provide an overview picture, may also hide important fine-scale detail. A coarse-scale biodiversity map may provide overall trends, but it may fail to identify biodiversity hot-spots that become visible only in the analysis of finer scale data.

Technical and data problems aside, organizations associated with distributed assessment systems may need to adopt common data protocols, including protocols for data quality assessment and quality control, otherwise both data aggregation and comparability across or within a given scale can be compromised. Reality is that even if protocols are available, the capacity and inclination of centers to follow them is usually not the same.

Despite large improvements in access and incremental improvements in primary measurement, it would be illusory to think data availability is no longer a major problem. Many analytic techniques, such as modeling and scenario analysis used in IAR require not only compatible data sets, but also continuous streams of high-resolution data and time series. These are rarely available, particularly when it comes to field measurements from developing countries. The reasons are often historic:

> "If we look back in the history of the British colonies, like India, Pakistan, Singapore, Malaysia and so on, their systems of data collection are still in the 1930s. Whatever the tax collector needed then was collected and they are still following the same thing."
> (S. Shresta interview)

Besides availability, however, with the increasing policy relevance and scrutiny of some integrated assessments, other issues related to data quality, organization, resolution, and privacy receive more attention. As one of my interviewees leading a global indicator program observed "data and sometimes the information that comes out of international reports is so poor that people in the countries don't even recognize it" (L. Flanders interview).

Despite the release of large swaths of scientific (e.g., satellite) data to the public domain, there is a counteracting trend of data commercialization with major implications for integrated assessment. National security concerns have always hampered data availability, and this trend can only expected to strengthen in the wake of the terror attacks on the United States.

Many companies collect data that may also have uses in public policy, but withhold it as confidential information. For example, the steel industry has vast amounts of data potentially relevant for global assessment but keeps it confidential to avoid passing on trade and technological secrets (J. Aloisi de Larderel interview). It is a different case when data collection is paid for by public agencies but has to be bought either from these agencies or private sector re-packagers and re-sellers to recover costs and create value added. I am not arguing against the commercial approach, but one must bear in mind implications and risks for assessments. The equity implications of this trend can be significant and worrying, unless special provisions are made, particularly for data that is gathered with the use of public funds to meet essential public demand. Data that have to be purchased are often accessible only to Western institutions that can afford it, and out of the reach for most developing country organizations; they may well be out of reach even for international assessment systems. In some cases, assessment may play a role in bringing data access controversies to the attention of decision-makers. For example, energy data were first not included in the statistical annex to the public GEO Latin America and Caribbean report, because they would have had to be bought – not possible because of insufficient funding - from a regional energy organization that received the information from countries free. Discussing the controversy openly in the assessment made the Caribbean Commission, a high level political body take action and request keeping the data in the public domain at no cost. As a result of this intervention, the data in question has been made available for the assessment at no cost (E. Gutierrez-Espeleta, Observatorio del Dessarrollo, interview).

Besides more general data problems, integrated assessment presents additional unique challenges.

Broad framing, focus on interacting environmental and socio-economic domain
In IAR, data are needed on a much wider range of issues than the environment. Given that most integrated assessment and reporting agencies have environmental backgrounds, their expertise and connections tend to be weaker to organizations responsible for socio-economic data. On the other hand, agencies that collect socio-economic data may not have expertise working according to the perspectives of environmental or integrated analysis. Dealing with the complexity and magnitude of the arising tasks usually exceeds the capacity of a single agency, and even the core group of organizations in the assessment network; there is a need for a partnership with primary data providers.

Dynamic cross-temporal perspective and corresponding need for time series data
Constructing time series can be particularly difficult not only because of data gaps, but also because measurement methods may change over time. Integrated assessment often requires not only data on one, but on several time series data sets, both socio-economic and environmental. One missing critical time series with regard to an issue domain may be enough to disable the entire analysis.

The boundary problem
The spatial units underlying integrated assessment and monitoring/data collection do not always coincide. Socio-economic data tend to be collected for political jurisdictions, while environmental data are collected according to ecosystems. Finding a classification system that matches both the needs of the assessment and the possibilities of data structures can be a major challenge. We certainly found this a major difficulty during the preparation of one of the first sustainable development reports for Manitoba (Manitoba Environment 1997). However, the problem is universal and mentioned by others in the context of global assessments: “Certainly, the data issue combined with the desire to match more closely with the Convention on

Biological Diversity's needs forced us in the Pilot Analysis of Global Ecosystems to use the very traditional mix of biome and land use categories, but to be honest we felt that this was one of the biggest obstacles to doing an integrated assessment." (W. Reid in Millennium Ecosystem Assessment 2001).

### Indicator construction and aggregation

Raw data are usually unsuitable for direct use in assessments: there is a need for algorithms to transform data into indicators and indices that are more useful in the analysis and decision-making and that are easier to communicate. Seemingly a technical matter, indicator construction can be complex and highly political - think about the problematique of compliance monitoring and performance measurement in the context of climate change. Even if an assessment opts not to construct new indices or indicators, it has to choose from available indicator menus. How one makes a choice is a matter for assessment design involving many technical and political challenges. As earlier noted, indicator construction or selection has implications for other design elements as well. Who leads indicator selection and how; who would take part and in what role; what is the role for expert opinion versus stakeholder input; how to associate indicators with targets; how to assign weights to particular indicators used in the construction of an index?

As the degree of discontent with the GDP-based measures of progress has grown since the early 1990's, there have been a growing number of attempts to construct highly aggregate indices that reflect socio-economic or environmental sustainability concerns (Hardi 2000). Aggregate indices are no substitutes for more detailed sets of indicators, as they both add and lose information. In fact, they are both needed: as one would be able to scale up to aggregates from more detailed measures, it should be also possible to deconstruct aggregate indices into their constituent parts. As there is a need for both aggregates and details, an assessment, particularly a coarse scale assessment should build these into its information system.

Besides those listed above, IAR poses additional challenges that warrant accepting data and indicators as a separate major assessment design element. Assessment producers may choose to partner with key primary data providers, or develop their own data warehouse for key datasets and indicators. They may disregard issues where data are lacking, or choose to make data problems explicit, or to select indicators in an internal expert process or involve a broader stakeholder community, and so on. These design choices are closely related to other design elements including participation and governance (e.g., Costanza and Tognetti 1996; Boyle, Kay and Pond 1999; Hardi and Zdan 1997). Put together, data problems can be significant enough to warrant the establishment of specific subgroups in the assessment to deal with them, as has been the case in GEO.

### *6.6.2. Data and indicator issues in GEO and distributed systems*

Data issues have been given significant attention in GEO, particularly from the earliest stages of GEO-2000 when in light of GEO-1 data needs and limitations were more clear and both the GEO methodology and the CC network started to solidify (UNEP 1997b). Data has been seen as a strategic issue and as such dealt with in most production guidelines and other strategic documents.

The basis of the strategy was the need for environmental and socio-economic time series data from the global to sub-regional scale, and the position that none of the existing databases maintained by other organizations sufficiently met this need. GEO was not mandated nor enabled to undertake primary data collection and warehousing, but it could go as far as compiling and maintaining a custom data matrix and Internet based data portal with the required support system. It was also thought the CC network and the participatory GEO process had particular strengths that could be utilized in data work.

Key elements of the data strategy included:

Data Working Group

On the organizational front, a Data Working Group, led by UNEP/GRID-Geneva was established to coordinate all data work and take responsibility for developing

products. Privileged access to different regional or thematic datasets and data providers has been recognized as a strength of the network and participatory approach. Therefore, besides Working Group members, all CCs had particular data related responsibilities.

### Data / indicator matrix

The matrix includes core socio-economic and environmental data relevant for GEO. It is organized according to the DPSIR framework with thematic sub-categories. The original matrix design calls for time series data, targets, and accompanying metadata. Originally available as an Excel file, the matrix was later migrated to the Internet and provided some of the basic datasets for the GEO data portal (UNEP/GRID-Geneva 2001). With careful planning and implementation, the portal can become an important tool that serves not only the purpose of GEO but audiences in need of key global data sets.

### Core indicator set

It has been clear from the beginning that GEO would need to derive findings from and support messages through a small set of core indicators and make a concerted effort to ensure these indicators are based on the best data available and clearly associated with targets. Some other projects, such as the Organization for Economic Cooperation and Development's 2001 outlook carried through with this approach (Organization for Economic Cooperation and Development 2001). Although at various stages of GEO, draft core indicator sets have been developed, they never became fully institutionalized or systematically implemented. The need to move the issue forward has been recognized, but strategies are yet to be worked out (D. MacDevette, pers. comm.).

### Aggregation

Despite several parallel efforts, there is no international consensus on the design of an aggregate index on the environmental dimension of sustainable development (Hardi 2000). As GEO simply cannot work without a consolidated core indicator set, it

probably cannot do without a high level index of environmental sustainability either. Given the technical and legitimacy problems with other systems, it probably cannot simply adopt an existing design, although it can and it did reference some, such as the ecological footprint. As part of the GEO-3 process, a environmental sustainability and separate vulnerability indices were developed in cooperation between UNEP and the GRID network. This is an important step in the right direction, but it is also the easier part of the task. As challenging as developing an aggregate index is, it is more difficult to get that accepted by the scientific and policy community. It remains to be seen whether UNEP can gather sufficient support and attention through the GEO process, GEO involved the CC network and from what Ravetz (1999) calls 'extended peer communities', including the media.

### Gap-filling and data quality

One of the core functions of the Data Working Group was to undertake and coordinate a data-gap filling and validation exercise. The exercise was based on the fact that many global time series data have major gaps in some regions of the world, and on the assumption that some of these data may actually exist in regional, sub-national or national sources. It was also recognized that through the CC network, GEO might have better and more direct access to these sources than many other global data processes that operate from a central location. During this exercise, a concise version of the GEO data matrix was filled out from global sources and sent to CCs to compare and correct relevant regional and thematic sections based on data from regional sources. The process also served as a data quality control mechanism as discrepancies between data from global and in principle, the same data from regional sources were identified and if possible resolved. Data quality is an elusive attribute, and can be assured (and indirectly measured) through quality control mechanisms built into the assessment process (United States Environmental Protection Agency 1994; Denisov and Christoffersen 2001).

As mentioned before, GEO is not equipped to gather primary data, but it can closely cooperate with other data providers. Some other global assessments such as the MEA

do have this mandate and strengthen scientific knowledge, in the case of the MEA on wetlands, desert ecosystems and biodiversity. It is in GEO's interest to closely cooperate with these initiatives and ensure to the degree possible that their data work is informed by GEO's general data findings and helps resolve them.

Data access

Harmonized data access is particularly important for distributed assessments where using the same data and data protocols are important for ensuring cohesion and cross-sectoral or cross-regional comparability and cross-scale aggregation. While web-access and email were problematic for many GEO CCs – and UNEP HQ itself – at the beginning of GEO, this ceased to be a constraint by GEO-3. In order to facilitate access to data, UNEP has established a GEO data portal and made it available on CD-ROM and the Internet (UNEP/GRID-Geneva 2001). By the time the portal became truly functional, GEO-3 was well under way, so its full-fledged testing will likely have to wait until the GEO-4 process. Access to the portal is open to anyone with an Internet connection, although bilateral agreements between UNEP and primary data providers restrict the use of data only to members of the GEO network and for the purpose of the GEO report.

The GEO portal could co-exist and co-evolve with, borrow from and contribute to other similar systems maintained for slightly different purposes by other organizations. New information systems on environment and development will continue to emerge; more recent examples include WRI's data portal and IISD's evolving Dashboard of Sustainability (World Resources Institute 2001; International Institute for Sustainable Development 2001).

Besides the data portal, GEO CCs have also produced regional CD-ROM databases and portals compiled partly from regional and country-level sources. The database project, led by the Observatorio del Desarrollo, the GEO CC in Costa Rica, published a CD-ROM for the Latin America and Caribbean (LAC) region. The CD-ROM is useful in the GEO context, but being the first of such collection in the region it is also

targeting other external audiences (E. Gutierrez-Espeleta, Observatorio del Dessarrollo, interview; Observatorio del Desarrollo 2001). Other similar, internet-based portals include the North African (Center for Environment and Development for the Arab Region and Europe 2002) or an Asia-Pacific SoE Data portal (UNEP Regional Resource Centre for Asia and the Pacific 2002).

Summary of strategic data issues

An important aspect of the GEO Data Working Group is identifying and analyzing general and strategic data issues encountered during the preparation of GEO. The Working Group prepared a report that was published as a GEO side-product and its key findings have been summarized in the main GEO-3 report (van Woerden *et al.* 1999; UNEP 2000). This work was to inform further data work in GEO but it was also intended for external audiences, particularly for agencies involved in or funding primary data collection.

### *6.6.3. Implications for effectiveness*

The handling of data issues has major implications for the credibility of assessment and reporting systems. Some of these are related to data and data procedures; others are associated more with the way data are used. There are many ways to risk credibility, including:

Using data from sources that are unrecognized or of questionable quality;
Relying on the wrong or outdated data where better alternatives are available;
Sloppy referencing and metadata practices;
Weak data quality assurance / quality control;
Statistical error and inference related problems;
Generalization based on non-representative data.

These risks could be particularly high for global assessments that are removed from local contexts where there is better awareness of problems with particular data. In other words the more removed the assessors are from local contexts, the more difficult to judge the

appropriateness information (F. Carden interview). It is also more difficult to recognize problems with data.

Distributed systems involving a network of contributing organizations bring both challenges and opportunities with regard to addressing data quality problems. On the one hand, there is a challenge associated with coordination, uneven capacity among network members and ensuring the systematic use of data protocols, quality assurance and control. At any given point in time, there may be several parallel and simultaneous data processes with their own terms of reference and products. Unless synchronized, these can lead to results that are competing or contradictory, confusing audiences who are looking for unequivocal messages.

A case in point is the development regional data portals in GEO. They respond to the real need for information on the regional level, however, regional data is also needed to analyze typically global issues. Thus the audience  e.g., for greenhouse gas emissions in the Asia-Pacific is not only in the region, but also includes stakeholders interested in overall global trends or regional comparisons. There is thus a need to find a link between global, regional and subregional portals. Thanks to the internet, technical barriers no longer prevent us from building multi-purpose portals, but the institutional/political barriers to integrating already developed similar purpose regional portals can be significant. In GEO the challenge is how to use the CC network contribute to the global portal developed at GRID-Geneva that could be developed to have regional sub-portals and support decentralized data input and management.

Distributed systems also have some clear advantages. Individual network nodes (GEO CCs) are more familiar with regional or thematic issues, data, and data problems than a global organization like UNEP. If capacity and resources are available, quality assurance can be decentralized and contribute to better data, improved credibility and ultimately increased effectiveness. This can involve CCs, but beyond that, involve even what I referred to as ‘external participants’ or GEO’s extended peer community. The GEO data validation exercise has built on the capacity in the network and to some degree

decentralized the data quality process. It also has had its own problems from which future exercises can learn. Major lessons are: (a) data quality checking and improvement is a major task that requires matching, long-term commitment and ongoing investment by partnerships among the right institutions; (b) data acquisition, quality assurance and quality control, although more important in particular stages of the assessment, involves a a series of steps and needs to be integrated into the entire assessment and reporting process (c) data quality protocols have to be very clearly defined, accepted and followed, otherwise results will be inconsistent; and (d) expected outcomes and follow-up have to be outlined from the beginning and progressively refined; products should include recommendations and courses of action to prevent the reproduction of the diagnosed data problems in the future (L. Pintér, unpublished interview data with GEO Data Working Group members).

| Recommendation |
| --- |
| In order to understand strategic data issues and coordinate data and indicator related activities, assessments should consider establishing and maintaining data subgroups. |

Data issues – indicators and indices in particular– have important implications for saliency and awareness associated with an assessment. If there is a real demand for integrated assessment and reporting, it is a demand for key indicators and indices. Properly designed and well supported by analysis, they can concentrate key messages in a way policy-makers and the public can easily understand and to which they can react. Having indicators is necessary but not sufficient element of saliency and awareness. Not all indicators are policy relevant and easy to communicate; those that are not, need to be made so through indicator selection in the assessment process.

Although hard to objectively measure and verify, policy relevance is one of the most common sustainable development indicator criteria[22]. As participation plays a major role in ensuring the saliency of assessments in general, it is equally important in ensuring the policy relevance of indicators; in fact, indicator selection can and should be built into the participatory strategy of the assessment process. The assessment can enhance the policy

relevance of indicators by making a link to specific policies and policy targets. Policy targets add more value and weight to the analysis, but also new complications to the assessment. In addition to forming a consensus on indicators, the assessment would need to look for commonly accepted targets or get into politically and scientifically risky target setting. To paraphrase Meadows (1998), indicators (and assessments) not only respond to information needs, they also help transform positions and agendas; they are not only retrospective, but can be tools of looking ahead and play a role in envisioning and quantifying the implications of policy options.

GEO and particularly the GEO Data Working Group has done important preparatory work that has not yet been extended to indicators and indices. There are achievements upon which to build in terms of organizational structures (the Data Working Group and the connections to major primary data providers), process (the multi-scale, participatory approach and peer review characteristic of the entire GEO process), and products (the data matrix and portal). Both my interview data and personal communication over the last 1-2 years with many UNEP-DEWA staff indicate that momentum for indicator work is building. Rather than following a shotgun approach and develop a 'complete' indicator set, GEO should consider a more evolutionary and adaptive strategy. The strategy should involve focusing on a small number of indicators associated with highest-priority policy issues plus preparatory work on a highly aggregate index or indices. A similar strategy, less the creation of a high-level aggregate, has been followed on the national level (e.g., by Environment Canada in the development of the National Environmental Indicator Series (Environment Canada 2001)). The slow-go, gradual approach to initially focus on 3-4 indicators has already been proposed for GEO-2000 and would need to be implemented (V. Vandeveerd interview). This strategy would not only make the work of the assessment easier by not having the pressure to develop a 'perfect' and all-inclusive indicator set, but it could also start with indicators that are less controversial. It could more safely pilot both the indicator selection process and indicator design steps that it could later extend or modify as necessary to additional indicators – benefits of an iterative assessment process. Work on highly aggregate indices has already started. Here, again, an evolutionary, learning approach with multiple feedback loops from the science

and policy community as well as the media would increase not only credibility but improve the chances that the index builds on the views of audiences.

| Recommendation |
| --- |
| IAR systems should create a mechanism to establish and periodically review a set of core indicators and aggregate indices. |

With regard to awareness, indicators and indices can be important in 'branding' the assessment. There are many cases where high level indices in particular became well known and expected elements of reporting systems. Some of the examples include UNDP's use of the Human Development Index (United Nations Development Programme 2001), the ecological footprint in WWF's Living Planet Report (World Wildlife Fund 2000), the World Bank's World Development Indicators (World Bank 2001) or the World Economic Forum's Environmental Performance Index (World Economic Forum 2002). To reiterate points I made under the previous section dealing with the communication design element, just as the entire assessment, a lead index of sustainability has only potential impact. GEO should incorporate not only the development of a high-level index with regard to the environmental component of sustainable development, but also make sure the index and underlying, gradually developed indicators are built into the overall GEO communication strategy.

The report by the U.S. National Research Council's Board on Sustainable Development (1999) on the sustainability transition comments that "research on a large scale, long term phenomenon that need to be seen in real time demands a symbiotic relationship between scientific investigators and the routinized gathering of data". Although they can do their share, global assessment and reporting systems are usually not mandated and enabled to offer systemic solutions to data troubles. Without going too far into the problem, let me just point out that even global data-sets are often collections of data routinely measured by many networks of independently maintained field stations. Although one can often improve data sets by using statistical techniques, gap filling, use of proxies and other methods, satisfactory and long-term resolution of the data problem is not possible without

improvements in underlying monitoring systems. These systems are organized in multi-scale, polycentric 'panarchies' under the territorial or thematic jurisdiction of multitudes of national and sub-national semi-independent agencies. Global coordination at least in the environment is provided by the large global observation systems (e.g., World Meteorological Organization 1997; Integrated Global Observing Strategy 1999). It is in the very self-interest of GEO and other similar assessment and reporting systems to maintain close links with these groups. First as an institution situated on the interface of science and policy, they are in a unique position to transmit messages about data issues from the perspective of policy audiences, but they also are familiar with underlying technical and systemic data constraints. This information could be valuable to monitoring systems under pressure to change not only from the scientific and policy, but also the technical / technological side.

| Recommendation |
| --- |
| Assessments should establish mechanisms to work closely with organizations involved with the gathering of primary data both to ensure data access but also to provide monitoring organizations feedback concerning data usefulness and validity. |

Better cooperation between monitoring groups and assessment initiatives would help, but what is needed is a rethinking of the governance of data systems in the context of global change. The U.S.-based Commission on Global Environmental Change Information Policy referred to 'global information system cooperatives' of the large data providers and data brokers that could lead to strengthened and more commonly shared data standards, protocols and policies (Parris, Zracket and Clark 1998). Others called for the recognition of a 'global environmental information commons', and the need for joint efforts to generate and make the data essential for the management of shared environmental problems available to all (National Council for Science and the Environment 2001; BC). Taking this further, the GEO Data Working Group brought up the possibility of a global convention on data, following the precedent set by the Aarhus Convention (Economic Commission for Europe 1998). Whether through a convention or some other instrument, there are systemic multi-scale problems that constrain the use of

data in assessing global, regional and sub-regional trends that can hardly be addressed without suitable intergovernmental mechanisms.

## 6.7. Feedback to research agenda setting

### *6.7.1. Why feedback to research agenda setting?*

The role of scientific advisory processes and assessment in promoting policy learning and adaptation with regard to complex environment – development issues have received considerable attention (e.g., Estrella *et al.* 2000, Bennett and Howlett 1992, Meppem and Gill 1998, Lee 1998, Siebenhüner 2001). Integrated assessment and reporting support learning in the context of a planning-action-reflection framework by closing the loop, providing information on past performance and ideas for adaptation and future directions. Often related to adaptive management, this approach "treats policies as hypotheses, and management as experiments from which managers can learn" (Holling 1978, Walters 1986).

The primary purpose of integrated assessment and reporting systems is to produce information on environment/development trends and dynamics and answers to questions of policy-makers. However, they also tend to lead to further questions, for example related to knowledge and information gaps. As discussed by many authors, the issues with which integrated assessment and reporting deals are complex and imperfectly understood even when viewed separately. Try to integrate them and uncertainty becomes their inherent characteristics (e.g., Kay and Regier 1999, van Asselt, Beusen and Hilderink 1996). The debate about the scientific treatment of uncertainties and knowledge gaps can and increasingly does spill over to the policy domain as clearly seen for instance in the treatment of the climate change issue. While assessment and reporting systems can and do resolve some knowledge gaps, uncertainties, and unknowns, they can not deal with others except identify them, elucidate questions and point our promising directions for further research. The assessment may not only point to unknown parameters related to known problems, but also discover 'unknown unknowns', problems that were not even in

the dictionary prior to the assessment. In this context, assessments function as part of an early warning system and introduce new issues and ideas into scientific and policy discourse.

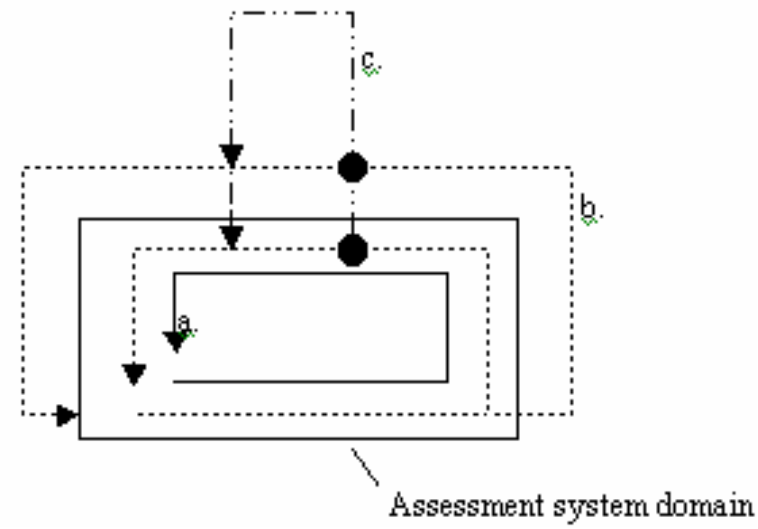
Through their contribution to the interaction of scientific and policy and by providing feedback on environment/development dynamics, assessment and reporting systems are one of many instruments influencing the direction and mechanism of social learning and scientific research. Individual scientists are able to adjust the 'short waves' of research, usually constrained by particular projects and criteria within calls for proposals. Researchers need to be enabled to carry out research on particular topics in the context of particular frameworks such as sustainable development. One may also find, however, that research agendas that are nominally set by funding agencies are actually set or at least greatly influenced by individuals, including scientists (A. Hammond interview). Individual scientists, whether on loan, seconded for 1-2 years, on consultancy, or employed in agencies write most, if not all requests for proposals. What later appear as corporate decisions can often be traced back to individual agency and advocacy. Consider the following statement about the policy decision to initiate GEO:

> "I think it is very good that somebody decided, well of course it was finally the Governing Council who decided, but there was somebody who has brought it to the attention of the Governing Council in the management of UNEP, to try to make a mixture of, if it is possible to say, of the social and natural sciences." (G. Golubev interview)

While I don't find this statement to the least surprising, it strengthens the point that in addition to understanding the institutional mechanisms involved in policy and research agenda setting, one must also understand and appreciate the role played by individuals.

Although I do not deal with the relationship of social learning and assessments in detail, I find it helpful to point to the general typology developed for organizational learning by Argyris and Schön (1996) and its application to the learning dynamics of assessment

systems Siebenhüner (2001). Their classification may also apply to analyzing the way assessment and reporting systems influence research agendas. Argyris and Schön's typology includes single loop learning (application of new knowledge without changing existing frameworks and objectives), double loop learning (deep change where frameworks, objectives and underlying assumptions are questioned) and deutero-learning (self-reflection and feedback in the context of learning mechanisms). A bare-bones diagram of the three types of learning that reflects my understanding of these mechanisms is shown in **Figure 12**.




*a. single-loop learning – self-reflection and feedback within the context of agreed-upon and existing frameworks of goals and causal beliefs*

*b. double loop – self-reflection and feedback extended to underlying goals, norms, belief structures and system orientation*

*c. deutero-learning – self-reflection and feedback on the meta-level, based on learning processes*

**Figure 12:** Schematic diagram of three types of learning based on Argyris and Schön (1996) and Siebenhüner (2001).

Assessment systems may contribute to research agendas on all these levels: to single loop learning for example by pointing to the lack of data regarding known environmental problems; to double loop learning by identifying fundamental flaws in policy measures that would require rethinking on the level of goals, objectives and paradigms; and perhaps also, to deutero-learning by calling attention to the mechanisms of in-depth policy evaluation, feedback and control. While changes that do not affect deeper layers of research design and direction may be implemented on the short term, more fundamental shifts in the direction of research programmes that affect underlying worldviews and assumptions arise on a less frequent basis. Reorienting major research programmes according to the challenges and needs of sustainability requires a major shift, as pointed out by many (e.g., Kates et al. 2001, Lubchenco 1998).

Research agendas are set on many levels and by many organizations from multilateral to national aid agencies, governmental scientific research programmes, foundations, universities, private firms and others. Global, multi-level assessment and reporting systems may exert some influence on any of these. The relevant questions are on what level, through what mechanisms, through what institutional arrangements and with what results. For integrated assessment systems, making a connection may be problematic because of paradigmatic differences between the assessment and funding organizations. If the assessment comes to the conclusion that more research is needed on putting a price on environmental amenities, but funders are still approaching environmental issues strictly from the perspective of ecology, the first – not insignificant – task may be to bridge the worldview barrier.

It is likely that a well-publicized assessment and reporting system will have some influence on the direction of research informally, almost by default. However, it is also reasonable to assume that an assessment where this issue is recognized and followed by the preparation of an integrated strategy and systematic implementation, the impact can be more significant. This is presumably a good and desirable thing - assuming the information produced by the assessment is actually credible and legitimate. Let us not forget particularly some industry sponsored research programs related to climate change

or biotechnology, that are characterized by generous publicity efforts but often-problematic credibility and legitimacy.

### *6.7.2. Feedback to research agenda setting in GEO and distributed systems*

Although GEO does refer to the scientific community as one of its audiences, the connection to scientific research agenda setting itself is still rather implicit. The research community at large is represented in GEO through the Scientific Committee on Problems of the Environment (SCOPE). The changing intensity and character of SCOPE's role in GEO indicates that although the importance of this link is recognized, it is not yet systematically integrated into the current strategy. There is, however, a preceding question – given what GEO is, could it at all aspire to advise research agenda setting? Given its role as an integrator and compiler, GEO is certainly not in the position "to tell atmospheric chemists what are the horizons of atmospheric chemistry, or geologists what are the problems with soil erosion" (R. Mnatsakanian interview). Neither will people read the GEO reports to get a comprehensive and well substantiated view of what research priorities should be. These priority setting documents would have to be purposefully crafted that could nevertheless fit UNEP's profile and draw on the experience gathered through the GEO process (P. Raskin interview). The argument is not that GEO cannot contribute to the formulation of new research priorities – probably all forward looking assessments by definition should - but that such information would have to be first carefully constructed in the GEO process, second carefully compiled and third purposefully presented at the right fora.

Some participants involved in the assessment and reporting system may carry out and fund research and they may be in the position to modify research programmes themselves. Compared with some other assessment systems, UNEP and GEO carry out little, if any primary research, and it is sometimes referred to as a compilation. As one of my interviewees put it, GEO is "a bit more loose than what rigorous scientists would feel comfortable with". There are, however, particular CCs that have significant research capacity and credibility and direct connection with supporters of primary research who could mediate. Some of these, such as RIVM provide actual examples where GEO led to

in-house research efforts or became incorporated into the agency's advisory work to funders (J. Bakkes interview). As Jan Bakkes also pointed out, some of the impact of GEO on RIVM's research came not so much through the GEO reports but through the GEO process. Some specific examples are as follows:

> modeling gridcell-to-gridcell flow of water as a dependent variable affected by demographics, industrial production and other factors; this work was started in direct response to the needs of GEO using the IMAGE model; the work focused on water was later spun off to Kassel University and lead to the WATERGAP sub-model;
>
> request by Directorate General-XII of the European Union for information on environmental knowledge gaps on the global level to aid the design of the Fifth Framework Programme; this resulted in a joint workshop where insights from GEO were communicated.

In other cases, such as the Environmental Outlook work of the Organization for Economic Cooperation and Development, a decision has been made that the assessment would not deal with issues where it lacks sufficient knowledge (e.g., tourism and land-use). Rather than ignore these knowledge gaps, however, the problem was to be noted and incorporated into the organization's research strategy to ensure adequate information is available by the next iteration of the assessment (L. Mortensen interview). The two examples, OECD and RIVM are different from GEO and UNEP in the sense that they are more in a position to define and fund new research directions that they themselves can then follow. In distributed systems, where the production of the assessment is decentralized and research needs emerge in many institutional contexts and on many scales, defining and responding to new research needs would definitely require distributed capacity development strategies and thus matching distributed research funding schemes.

During the GEO-2000 process SCOPE has coordinated a survey exercise to identify priority emerging issues in the environment whose results were published as a section of

the GEO report (UNEP 2000). The purpose of the exercise was to call the attention of the policy and scientific community to these issues, defined as issues that are entirely new, or previously known but likely to reappear in new contexts with previously unrecognized dynamics. While the exercise engaged a segment of the scientific community, increased GEO's awareness and produced some interesting results, it could have lead to more persistent follow-up based on the understanding of how research agendas are set and by whom.

Having said this, as the GEO assessment matured, some constraints uniquely affecting multi-scale integrated environmental assessment and reporting became apparent. GEO's strategy was partly to patch up the problem as much as possible given capacity and other constraints, but in some cases, the problems were apparently too large for GEO or UNEP alone and led to further action. Perhaps the best case is where UNEP chose to sponsor some internal work to address critical data problems (e.g., through the GEO data portal), but also to set up a thematic working group involving several CCs. This group addressed the data issue on the meta-level and reported to the broader policy and research community on the problems and required research directions (van Woerden *et al.* 1999).

While the establishment of the data group and the publication of a report was a necessary and useful step both internally for GEO and towards the outside world, this could be followed up more vigorously, say, with the agencies responsible for the coordination of monitoring and observation systems. The appropriate partners besides SCOPE, that has a rather general mandate, could be the three global terrestrial, climatic and oceanic observation systems (Global Terrestrial Observation System or GTOS, Global Climate Observation System or GCOS and the Global Oceanic Observation System or GOOS) and the large global science programs (the International Human Dimensions Programme or IHDP, the International Geosphere-Biosphere Programme or IGBP, the World Climate Research Programme or WCRP and DIVERSITAS). If such links are to be pursued, GEO would probably be required to do more than pass on reports, and at least designate a liaison within the GEO network or within UNEP to engage with the partner organizations on a more substantive and sustainable basis. Undoubtedly, this could add a further

coordination challenge, but it is unlikely that significant impact can be achieved without significant investment of effort in communication and relationship building.

### *6.7.3. Implications for effectiveness*

In comparison with other design elements, feedback to research agenda setting affects longer-term success and effectiveness of assessment and reporting, unless there is capacity within the assessment to directly add to primary research, as it is more the case in the IPCC and possibly the MEA. In comparison with centralized assessment systems, polycentric assessment structures may have better opportunities to provide feedback to the research establishment on multiple scales. Various members of the network have different comparative advantages, interests, and tend to have links to different funding organizations closer to their scope of work. In GEO, CCs with regional expertise could speak to research needs and coordination issues with more legitimacy and perhaps credibility in their geographic area than UNEP HQ. On the other hand, UNEP has more legitimacy communicating research needs on the political level or with global funding agencies such as the Global Environment Facility (GEF). The central agency leading a distributed system may pursue one type of strategy in terms of influencing research agendas and a collaborating organization lower in the hierarchy another. Ideally, the two are coordinated.

While I found no direct evidence in this work that an assessment's ability to influence research agenda setting would influence its credibility, I cannot exclude that positive correlation exists. Such correlation could emerge, particularly if scientific bodies known for rigorous quality control procedures were known to have taken the results of the assessment into account. On the other hand, recommendations for research policy can also be of concern for the assessment, as is the case in the IPCC where frequent discussions are held about the dangers of getting too involved in politics (Siebenhüner, pers. comm..).

It is widely supported that having high scientific credibility is one of the preconditions for an assessment to be taken seriously enough to lead to a reorientation of research

programs (e.g., A. Hammond and R. Maas interviews). There have been examples where this has occurred in GEO –the modeling work of RIVM and the Stockholm Environment Institute-Boston (SEI-B) is a case in point - so some participants or some aspects of the assessment obviously had what it takes to influence research. However, this is not the case for the assessment overall, and there were also opinions that in its current form, GEO may not be set up to, nor does it have the targeted products, mechanisms and partnerships that would lead to more substantive impacts on research agendas (L. Pintér unpublished interview data). To change this would first of all require a clear position that besides influencing future policy agendas GEO should also influence research agendas. This would be followed by the preparation of a strategy based on the understanding of how research agendas are set, by whom and for what priority areas in GEO? The strategy could then identify the institutional mechanisms, processes and products and assign responsibility either to single organizations associated with GEO or initiate alliances with external partners – ICSU, UNESCO or other international programs, national academies, science foundations and others - who are better positioned to influence research agendas. In recognition of the needs of GEO to improve the knowledge base on multiple scales and the opportunities inherent in a distributed network, the strategy could have elements focused on the global as well as lower, regional or national scales.

| Recommendation |
| --- |
| In order to increase particularly the saliency of scientific research, IAR systems should consider developing strategies aimed at communicating findings with lessons for the direction of scientific research to relevant organizations. |

I find the linkages to other criteria of effectiveness, particularly saliency and legitimacy similar in that these are preconditions, rather than consequences of an assessment aiming to modify research agendas, at least on the short term. With regard to saliency and research agenda setting, the assessment would need to properly identify the target groups for which its outputs are salient. There needs to be a match between the scale of the assessment and the issues with which it tries to deal and the level where a particular research funding agency works. UNEP is more likely to have the ear of GEF directors

than a regional CC; on the other hand, a regional CC, say the Asian Institute of Technology, may have more legitimacy in the eyes of a regional funder, such as the Asian Development Bank. A distributed system can be strategically positioned and position its members

| Recommendation |
| --- |
| Capitalizing on their presence in many regions, distributed assessment and reporting systems should consider influencing regional research agendas through their local member institutions. |

As a system characterized by presence in all regions, exposure to a broad range of stakeholders through the GEO process, future orientation, and integrated, future oriented perspective, and a mandate to make policy recommendations from the UNEP-GC, GEO has many structural elements that could make it have more substantive influence on research agendas. Influencing research agendas is more likely to succeed if it is not an add-on but integrated into the assessment process and participants include high level representatives of funders. GEO has attempted, with variable success, to engage such groups, though it could do better. The first step towards that would be a goal of clearly identifying priority research agendas to be influenced, followed by strategizing and implementation. One of the most predictable results of scientific research is the need for further research. Properly set up, boundary processes operating on the interface of science and policy may one day be set up to identify research needs that arise from both.

# 7. Conclusions

Ten years past the United Nations Conference on Environment and Development and only months before the 2002 World Summit on Sustainable Development, evidence coming from most parts of the world points to the continuing degradation of the environment and reinforces its intricate connection with human development. The debate on international environmental governance seems to be shifting away from centralized solutions towards more distributed institutional arrangements. As the environmental problem is pervasive and multi-scale, it makes sense to look for governance models that allow for local differentiation and innovation, yet connect to the global level through coordination and networking mechanism. This approach is particularly important in the case of assessment systems that need to combine local sensitivity with an ability to analyze and communicate issues of global significance. Put in this context, in the domain of global integrated assessments UNEP's Global Environment Outlook with its participatory approach offers interesting and important lessons about the feasibility and effectiveness of such systems.

GEO is a relatively recent addition to the family of assessment and reporting systems and processes aimed at providing policy relevant, science based information on environment and development. As both the GEO user survey and results of this research demonstrated, the program met many of the expectations of its key audiences. Put in another way, although there were some criticisms, most of them were framed as suggestions for correcting or improving an overall well-received program. GEO seems to have found and to a considerable degree, successfully filled a niche. The question is not whether global scale, regionally differentiated integrated assessment and reporting systems are needed, but who will control and implement them, how, and with what impacts and effectiveness? Ultimately, there is also a question of who will observe and evaluate them. While the emerging IARs need to learn from traditional science assessments, they cannot be simple replicas. Neither can they be thought of as extensions of operational SOE-style reporting. To partly borrow a description from Mike Hall of the United States Global Change

Research Program (USGCRP), the emerging generation of assessments can be best thought of as a connecting net rather than dividing 'wedge' of practices that have to evolve to eventually combine research and operational elements. As an operational system heavily dependent on the results of primary research from many sources, GEO offers interesting lessons.

While academic research could afford to deal with a single design element of assessment systems, most practice-oriented evaluations require a more comprehensive approach. There are some rather obvious and significant tradeoffs. Narrow evaluations have more analytic depth, lower levels of complexity, but may miss out on systemic interactions and insights that become apparent only when dealing with the system as a whole, as required by most formal evaluations.

My research dealt with the cross-section of four criteria of effectiveness and seven design elements. Besides the criteria of saliency, legitimacy and credibility originally in the GEA criteria-set I added awareness. The design elements included framing, governance, participation as covered in detail by GEA, and four others, including capacity development, communication, data and indicators and feedback to research agenda setting that I found particularly relevant for GEO. Although I would not consider these lists complete, they captured the most important aspects of a GEO-type initiative. There are remaining gaps. For instance, having considered indicators as design elements, I could have also added modeling and scenarios. I decided not to, partly because I had little interview data on this topic. More importantly, I wanted to avoid further proliferation of an already complex framework. Of course, this is not to diminish the importance of modeling and scenarios in GEO or integrated assessment, as they are essential tools in the analysis of both current trends and future policy options.

As expected, I found that most design elements affected more than one criterion of effectiveness and all effectiveness criteria were influenced by more than one design element. In this study, I discussed saliency, credibility, legitimacy and awareness as per the identified design elements. The following is a summary of findings categorized

according to effectiveness criteria. I take them to apply to distributed assessment and reporting systems in a more general sense, but they arise primarily from studying GEO.

In order to increase saliency, defined as the ability of the assessment to address the needs of a particular user or users the design of integrated assessment and reporting systems should aim to:

> Build on SOE reporting frameworks by analyzing relevant policies and policy options in the context of future scenarios;
> Provide policy relevant recommendations or, if that option is not available, information directly useful in constructing recommendations;
> Ensure early, systematic and substantive participation of stakeholders throughout the assessment and reporting process;
> Understand and address the information needs of critically important audiences and construct information products and processes that respond directly to relevant questions.

In order to increase credibility, defined as the technical believability of the information, the design of integrated assessment and reporting systems should aim to:

> Engage not only particular organizations but within each partner organization, insist on the direct participation of professionals who are credible both to the scientific and policy communities based on their track record in scientific research, policy analysis, and assessment;
> Introduce data and indicator mechanisms and protocols that help identify the best and most relevant data and ensure acceptable data quality;
> Apply systematic peer review and substantive involvement of independent, extended peer communities – including representatives of both the scientific and policy community - in all phases of the assessment, including the construction of policy relevant recommendations;

Ensure consistent communication of findings across all assessment products and target all relevant audiences;
Apply suitable frameworks that are true reflections of both stakeholder views and the scientific understanding of environment / development issues and interactions.

In order to increase legitimacy, defined as carrying out the assessment in a fair and politically acceptable way taking the views and values of its audience into account, the design of integrated assessment and reporting systems should aim to:

Employ participatory mechanisms and communication strategies that allow the expression and capture of diverse, even if contradictory, positions;
Assign balanced governance responsibilities to ensure overall coordination and cohesion in the assessment program but allow autonomy and devolution of powers to network members with higher legitimacy in local or thematic contexts;
Result in contribution to the capacity development needs of internal and external partners.

In order to increase awareness, defined as familiarity with the assessment and the implications of its key results for the policies and actions of a given audience, the design of integrated assessment and reporting systems should aim to:

Develop and implement communication strategies built into the assessment and reporting process and based on a clear understanding of information demands and decision contexts;
Present clearly explained and illustrated data, indicators and indices;
Include capacity development and participatory mechanisms to engage critical audiences, including science and policy opinion leaders and the mass media.

Assessment and reporting systems are instruments of social learning, but they also have to learn and adapt themselves. Besides providing feedback on the relationship of policy decisions and environmental outcomes, assessment should also contribute to social

learning by influencing research agendas. Neither GEO nor other global integrated assessment systems can afford not to play a role in influencing the direction of environmental research that will directly determine the success of their future work. While I found evidence that GEO did have some influence over research directions of particular CCs and – usually through the mediation of particular CCs – the research of other organizations, there is little awareness of this role. If influencing happened, it was more the result of the activism of particular CCs than coordinated effort. With more awareness and understanding of their interdependence, assessment systems should play a stronger role in research planning. This applies not only to science assessments that have more scientific credibility and demonstrated research expertise, but also to assessment and reporting systems that are closer to and have a better understanding of the information needs of policy audiences.

Special emphasis should be put on building a more substantive relationship between assessment systems and global monitoring systems and organizations. This is not to take away from the primacy of integrated assessment's policy orientation, but recognizing and responding to the need for repositioning global monitoring and data management systems based on both scientific knowledge and policy priorities. Assessment systems could both create organizations, such as special task groups, responsible for liaising with monitoring systems, and also produce reports that communicate data issues, needs and priorities that arose during the assessment and that cannot be resolved without systemic changes on the monitoring level.

Besides contributing to social learning, global assessment itself has to learn. This cannot happen without having strategies and mechanisms for self-reflection, internal and external evaluations. The results of the GEO user survey and this research are based on non-representative samples. Even if a statistically representative sample could be surveyed– a rather remote possibility - it would be very hard, as earlier explained, to isolate impacts on issue development from the influence of other information systems and products. Looking at the relationship of GEO's design elements and proxy criteria of effectiveness has been a useful approach. Besides providing information on pathways to

effectiveness, I find it particularly useful that results of such work can be more directly applied to the construction and design of assessments. This is not to say that this approach is superior to user studies that aim to directly measure impact through identifying and surveying samples of users; I rather view the two as complementary. A thorough evaluation of assessment and reporting systems should combine representative user surveys to identify potential direct impacts with an analysis focused on design elements and effectiveness.

Many assessments carried out during the past decade dealt with specific themes or sectors, such as acid rain, climate change or the energy sector. These exercises have already highlighted the complexity associated with assessment and reporting processes that aim to link science and policy. The emerging integrated assessment systems focused on entireties of spatial units or place as per Wilbanks and Kates (1997) and others quoted before would not replace, but coexist with, borrow from, and potentially contribute to in-depth thematic assessments. In this context, 'place' usually refers to contiguous spatial units with political, geographic and / or ecological boundaries and interacting forces of environment and development. Place-based integrated assessment will strike many practitioners as an extension of assessment and reporting they already do. The focus on place is often by default, given that institutions pursuing assessments often have jurisdiction over entireties of spatial units, such as cities, provinces or countries. It is not at all certain that 'place' also means fine scale. It certainly cannot mean small enough to avoid intra-scale heterogeneities. As anyone having been involved in community level assessment and reporting knows, the demand for further differentiation never stops: it can go down to the level of districts, neighbourhoods and below. Even within seemingly homogeneous places, there can be a universe of different value systems that create management, methodological and epistemological dilemmas very similar to those encountered on the national or international scale.

As suggested by Paul Raskin, at the other end of the spectrum, place-based cannot exclude global. There are many issues that unfold globally with distinct and crucially important regional differentiation, hence the need for a network approach to global

assessment and reporting. GEO is a good first approximation of a globally distributed system,

As assessment and reporting systems proliferate, the relationship between thematic and sectoral assessments and global overviews will become increasingly important, particularly as global and local symptoms of unsustainability intensify. Technology, particularly the development of information and communications will continue to be both a driving and enabling force, as will the 'virtualization' of many government services or, as more commonly known, 'electronic government'. This creates pressures to make information readily available in any format, for any issue or association of issues, for any spatial unit or time period. However, it will also reveal that the improvement of the delivery mechanisms alone is insufficient to meet the increasing demand for quality information without the parallel improvement of primary knowledge creation. Improving information infrastructure (not the same as information *per se)* and the increasing two-way 'traffic' between global and local scales will add an increasingly virtual character to assessment and reporting systems and processes where contributors are separated by space, focus of interest and expertise, but connected through the assessment process. While these 'virtual assessment communities' would address some of the issues of multi-scale, multi-issue assessments, they would also create problems of their own, ranging from the increasing importance (and difficulty) of process, data and output quality control, ownership of knowledge and information created, and others. GEO and other similar systems should and in a sense already do spearhead this movement. Making sure they better respond to people's information needs and lead to positive impacts in the environment are essential for achieving sustainability.

# 8. LITERATURE CITED


Andreasen, A.R. (1995) *Marketing Social Change: Changing Behaviour to Promote Health, Social Development and the Environment.* San Francisco: Jossey-Bass.

Anon. (2001) *Decision Theory.* Web Dictionary of Cybernetics and Systems. February 1, 2001. <http://pespmc1.vub.ac.be/ASC/DECISI_THEOR.html>

Allison, G.T. (1971) *Essence of Decision: Explaining the Cuban Missile Crisis.* Little, Brown and Co., quoted in S. Ney and M. Thompson (2000) *Cultural Discourses in the Global Climate Change Debate.* Conceptual Frameworks for Mitigation Assessment from the Perspective of Social Science, March 20-21, 2000, Karlsruhe, Germany.

Argyris, C. and D.A. Schön. (1996) *Organizational Learning II. Theory, Method, and Practice*. Reading, MA: Addison-Wesley.

Arnstein, S. (1969) "A ladder of citizen participation." *AIP Journal*, July, pp. 216-224.

Asselt, M. van (1994) *Global Integrated Assessment Models as Policy Support Tools.* Thesis in Philosophy of Science, Technology and Society. Enschede, the Netherlands: University of Twente.

Asselt, M. van, A.H.W. Beusen, and H.B.M. Hilderink. (1996) "Uncertainty in integrated assessment: A social scientific perspective." *Environmental Modelling and Assessment* 1: 71-90.

Attere, A.F. (2000) *Evaluation Report of Global Environment Outlook-1 and –2 Processes.* Nairobi: Evaluation and Oversight Unit, UNEP.

Bakkes, J. (1999) Letter to Marion Cheatle, UNEP-HQ. October 20, 1999.

Bakkes, J. *et al.* (1999) *A Contribution to UNEP's Strategy for Monitoring and Assessment.* RIVM Report No. 402001010. Bilthoven: RIVM.

Bennett, C. J. and M. Howlett. (1992) "The lessons of learning: Reconciling theories of policy learning and policy change." *Policy Sciences* 25: 275-294.

Bettelli, P. and L. Pintér (1999) "Beyond delusion: Science and policy dialogue on designing effective indicators for sustainable development." Report on workshop on 7-9 May, 1999 in San Rafael de Heredia, Costa Rica. *Sustainable Developments* 25(1), May 12, 1999.

Biermann, F. (2000) *Science as Power in International Environmental Negotiations: Global Environmental Assessments between North and South*. ENRP Discussion Paper 2000-17. Cambridge MA: Kennedy School of Government, Harvard University.

Boyle, M., J.J. Kay, and B. Pond. (1999) *Monitoring and Assessment as Part of an Adaptive Ecosystem Approach to Sustainability and Health.* Waterloo, ON: University of Waterloo. March 12, 2001.
<http://www.fes.uwaterloo.ca/u/jjkay/pubs/SCOPE/monitor.pdf>

Briggs, D.J., K. de Hoogh, and J. Wills. (1999) *Reporting Frequencies of State of the Environment Reports in Europe.* Technical Report No. 36. Copenhagen: European Environment Agency.

Carnegie Commission on Science, Technology and Government. (1992) *International Environmental Research and Assessment: Proposals for Better Organization and decision-Making.* New York: Carnegie Commission on Science, technology and Government.

Cash, D. and S. Moser. (2000) "Linking global and local scales: Designing dynamic assessment and management processes". *Global Environmental Change*, 10: 109-120.

Center for Environment and Development for the Arab Region and Europe. (2002) *North Africa Environmental Web Portal.* Cairo: CEDARE. January 12, 2002.
< http://isu2.cedare.org.eg/unep/>

Chandler, D. (1994) *The Transmission Model of Communication.* Aberystwyth, UK: The University of Wales. February 10, 2001.
<http://www.aber.ac.uk/media/Documents/short/trans.html>

Cheatle, M. (2001) *The GEO-3 Process.* GEO-3 North American Regional Consultation – United States. UNEP Regional Office for North America, Washington DC, May 10, 2001.

Clark, W.C. and G. Majone. (1985) "The critical appraisal of scientific enquiries with policy implications." *Science, Technology and Human Values* 10(3): 121-147.

Clark, W. *et al.* (2002) *Information as Influence: How Institutions Mediate the Impact of Scientific Assessments on International Environmental Affairs.* Cambridge, MA: MIT Press. Forthcoming.

Connolly, B. *et al.* (1998) *Institutions Working Group Theme Paper: Information and Governance.* Cambridge, MA: Harvard University. October 15, 2000.
<http://environment.harvard.edu/gea/pubs/98wgp_inst2.pdf>

Cook, T.D. and W.E. Pollard. (1977) "Guidelines: How to recognize and avoid some common problems of mis-utilization of evaluation research findings." *Evaluation* 4: 161-164.

Costanza, R. and S. Tognetti. (1996) *Integrated Adaptive Ecological and Economic Modeling and Assessment: A Basis for the Design and Evaluation of Sustainable Development Programs.* Paris: SCOPE-ICSU.

Crutzen, P.J. and E.F. Stoermer. (2000) *The Anthropocene.* IGBP Newsletter 41. August 30, 2000. <http://www.mpch-mainz.mpg.de/~air/anthropocene/>

Dale, A. (1998) *Sustainable Development: A Framework for Governance.* Ph.D. Thesis. Montreal: McGill University.

Denisov, N. and L. Christoffersen. (2001) *Impact of Environmental Information on Decision-Making Processes and the Environment.* UNEP/GRID-Arendal Occasional Paper 01-2001. Arendal, Norway: UNEP/GRID-Arendal. October 10, 2001. < http://www.grida.no/impact/papers/fullimpact.pdf >

Denisov *et al.* (2000) *Cities Environment Reports on the Internet: Understanding the CEROI Template.* Arendal: UNEP/GRID-Arendal.

Deshpande, R. (1981) "Action and enlightenment function of research." *Knowledge: Creating, Diffusion, Utilization* 2(3): 317-330.

Dürrenberger *et al.* (1997) *Focus Groups in Integrated Assessment.* Ulysses Working Paper 97-2. Darmstadt: Center for Interdisciplinary Studies in Technology, Darmstadt University of Technology.

Economic Commission for Europe. (1998) *Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters.* ECE/CEP/43. Copenhagen, Denmark: The Secretariat for the Aarhus Conference. August 30, 2000.
<http://www.mem.dk/aarhus-conference/issues/public-participation/ppartikler.htm>

Eckley, N. *et al.* (2001) *Designing Effective Assessments: The Role of Participation, Science and Governance, and Focus.* Report No. 2001-16. Cambrdige, MA: Research and Assessment Systems for Sustainability, Environment and Natural Resources Program, Harvard University. January 2, 2002.
<http://ksgnotes1.harvard.edu/BCSIA/sust.nsf/pubs/pub25>

Ellul, J. (1973) *Propaganda: The Formation of Man's Attitudes.* New York: Random House.

Environment Canada. (2001) *National Environmental Indicator Series.* Ottawa: Environment Canada. May 23, 2001 <http://www.ec.gc.ca/soer-ree/English/National/IndWelc.cfm>

Estrella, M. ed. (2000) *Learning from Change – Issues and Experiences in Participatory Monitoring and Evaluation.* London: Intermediate Technology Publications.

European Environment Agency. (2000) *Integrated Environmental Assessment in European Environment Agency Reporting.* Report of the Special Session of he European Forum for Integrated Environmental Assessment at the EEA in Copenhagen. Amsterdam: EFIEA Secretariat, Institute for Environmental Studies, Vrije Universiteit.

European Environment Agency. (1999) *Information for Improving Europe's Environment.* Copenhagen: European Environment Agency. November 1, 2001. <http://org.eea.eu.int/documents/brochure/brochurefull.pdf>

European Environment Agency. (1999a) *Environment in the European Union at the Turn of the Century.* Copenhagen: EEA. September 10, 2001. <http://org.eea.eu.int/documents/newsreleases/eu98.html#footnote1>

Ferguson, M. (2000) "EEA/EFIEA: Utilities of the science policy interface." in EEA, *Integrated Environmental Assessment in European Environment Agency Reporting.* Special Session of the European Forum for Integrated Environmental Assessment at the European Environment Agency in Copenhagen. Amsterdam: EFIEA Secretariat, Institute for Environmental Studies, Vrije Universiteit.

Fisher, R. and W. Ury. (1983) *Getting to Yes.* New York: Penguin Books.

Forrester, J. (1999) "The logistics of public participation in environmental assessment." *International Journal of Environment and Pollution* 11(3): 316-330.

Fritz, J.S. (1997) *Earthwatch Twenty Five Years on: Between Science and International Environmental Governance.* Interim Report IR-97-059/September. Laxenburg, Austria: IIASA. May 20, 2000. <http://www.iiasa.ac.at/Publications/Documents/IR-97-059.pdf>

Fritz, J.S. (1998) *Report on International Scientific Advisory Processes on the Environment and Sustainable Development.* Geneva: UN System-Wide Earthwatch. May 20, 2000. <http://www.unep.ch/earthw/sciadv.htm>

Funtowicz, S.O. *et al.* (1999) *Information Tools for Environmental Policy Under Conditions of Complexity.* Environmental Issues Series No. 9. Copenhagen: EEA.

Funtowicz, S.O. and J. Ravetz. (1993) "Science for the post-normal age." *Futures* September: 739-755.

Gallopin, G. *et al.* (1997) *Branch Points: Global Scenarios and Human Choice.* PoleStar Series Report no. 7. Stockholm: Stockholm Environment Institute. June 1 2000. <http://www.tellus.org/seib/publications/branchpt.pdf>

Global Environmental Assessment. (1997) *A Critical Evaluation of Environmental Assessments. The Climate Experience.* Cambridge, MA: Harvard University. November 10, 2001. <http://grads.iges.org/geaproject1997/>

Goldemberg, J., E. Martinot, and A. Miller. (2002) *Energy Since Rio: Achievements and Promising Strategies.* Washington, D.C.: Global Environmental Facility. January 16, 2002. < http://www.climnet.org/informal/PDD160102.pdf>

Grindle, M.S. ed. (1997) *Getting Good Government: Capacity Building in the Public Sectors of Developing Countries.* Harvard Studies in International Development. Cambridge, MA: Harvard University Press.

Gunderson, L., C.S. Holling and S. Light. (1995) "Breaking barriers and building bridges: A synthesis." in L. Gunderson, C.S. Holling and S. Light, eds. *Barriers and Bridges to the Renewal of Ecosystems and Institutions.* New York: Columbia University Press.

Guston, D. H. (2001). "Boundary organizations in environmental policy and science: An introduction". *Science, Technology & Human Values* 26(4): 399-408.

Hardi, P. (2000) *Review Paper on Selected Capital-Based Sustainable Development Indicator Frameworks.* Winnipeg, MB: IISD.

Hardi, P. and T. Zdan. (1997) *Assessing Sustainable Development: Principles in Practice.* Winnipeg: IISD. May 30, 2001.
< http://iisd.org/pdf/bellagio.pdf>

Hodge, R.A. (1995) *Assessing Progress Toward Sustainability: Development of a Systemic Framework and Reporting Structure.* Ph.D. Dissertation. Montreal: School of Urban Planning, Faculty of Engineering, McGill University.

Holling, C.S. (2001) "Understanding the complexity of economic, ecological, and social systems." *Ecosystems* 4: 390-405.

Holling, C.S., ed. (1978) *Adaptive Environmental Assessment and Management.* London: Wiley.

Integrated Global Observing Strategy. (1999) *Integrated Global Observing Strategy.* Geneva: UN-System-Wide Earthwatch. November 15, 2001.
<http://www.unep.ch/earthw/igosstr.htm>

International Institute for Sustainable Development. (2001) *Performance Measurement for Sustainable Development: Compendium of Experts, Initiatives and Publications.* Winnipeg: International Institute for Sustainable Development. September 20, 2001. <http://iisd.ca/measure/compindex.asp>

International Union for the Conservation of Nature, United Nations Environment Programme and World Wildlife Fund. (1980) *World Conservation Strategy: Living Resource Conservation for Sustainable Development.* Gland: International Union for the Conservation of Nature.

Jäger, J. (1998) "Current thinking on using scientific information in environmental policy-making." *Environmental Modeling and Assessment* 3(3): 143-153.

Jäger, J. and A. Farrell, eds. *Global and Regional Environmental Assessments: Design and Practice.* Forthcoming.

Jasanoff, S. (1990) *The Fifth Branch: Science Advisors as Policy-Makers*. Cambridge, MA: Harvard University Press.

Jasanoff, S., C. Miller and J. MacCarthy. (1998) *Impacts Working Group Theme Paper: Assessing Impacts: Framing, Communication, Credibility.* Draft for comments. Cambridge, MA: Harvard University.

Jensen, M.C. (1994) "Self-interest, altruism, incentives, and agency theory." *Applied Corporate Finance*, 7(2). April 2, 2001. <http://papers.ssrn.com/sol3/delivery.cfm?cfid=311027&cftoken=20102315&abstractid=5566>

Kates, R., *et al.* (2001) "Sustainability science." *Science* 292: 641-2.

Kay. J. and H. Regier. (2000) "Uncertainty, Complexity, And Ecological Integrity: Insights from an Ecosystem Approach ", in P. Crabbe, A. Holland, L. Ryszkowski and L. Westra (eds.), *Implementing Ecological Integrity: Restoring Regional and Global Environmental and Human Health.* Dordrecht: Kluwer, pp. 121-156.

Kay, J. *et al.* (1999) "An ecosystem approach for sustainability: Addressing the challenge of complexity." *Futures* 31(7): 721-742.

Keck, M.E. and K. Sikkink. (1998) *Activists Beyond Borders. Advocacy Networks in International Politics.* Ithaca: Cornell University Press.

Kluck, P. (1995) *An Analysis of Participation in Ten Comparative Risk Projects.* Boulder, CO: The Western Center for Environmental Decision-Making.

Knight, C.G. (n.d.) *A Framework for Integrated Regional Assessment of Global Climate Change.* University Park, PA: Center for Regional Environmental Assessment, The Pennsylvania State University. April 2, 2001. <http://www.iges.or.jp/openH/36.pdf>

Kolk, A.J. van der and J.N.M. Dekker. (1999) "Functions in integrated region-oriented environmental policy: A classification system." *Land Use Policy* 16: 107-119.

Kuhn, T. (1962) *The Structure of Scientific Revolutions.* Chicago, IL: University of Chicago Press.

Lee, K. (1993) *Compass and Gyroscope: Integrating Science and Politics for the Environment.* Washington, D.C.: Island Press.

Leviton, L.C. and E.F.X. Hughes. (1981) "Research on the utilization of evaluations. A review and synthesis." *Evaluation Review* 5(4): 525-548.

Long, M. and A. Iles. (1997) *Assessing Climate Change Impacts: Co-Evolution of Knowledge Communities and Methodologies.* Discussion Paper E-97-09. Cambridge, MA: Harvard University. November 10, 2001. <http://environment.harvard.edu/gea/pubs/e%2D97%2D09.pdf>

Losee, R.M. (1990) *The Science of Information. Measurement and Applications.* New York: Academic Press.

Lubchenco, J. (1998) "Entering the century of the environment: A new social contract for science." *Science* 279 (5350): 491-497.

MacNeill, J. (1990) "The dialogue between scientists and policymakers?", in Norwegian Research Council, *Sustainable Development, Science and Policy.* Oslo: Norwegian Research Council, pp. 505-512.

Manitoba Environment. (1997) *State of the Environment Report for Manitoba, 1997.* Winnipeg: Manitoba Environment, Government of Manitoba.

Martinos, H. and E. Humphreys. (1992) "Transnational networks: Evolution and future issues." *Ekistics* 352 (January/February): 13-20.

McLean, L. (2000) "Reflections on program evaluation, 35 years on." *Canadian Journal of Program Evaluation,* Special Issue, pp. 185-190.

Meadows, D.H. (1998) *Indicators and Information Systems for Sustainable Development.* Hartland Four Corners, VT: The Sustainability Institute.

Measurit. (2001) *Tracking Training Impact.* Shawnee Mission, KS: Measurit, Inc.

Menou, M.J., Ed. (1993) *Measuring the Impact of Information on Development.* Ottawa, ON: International Development Research Centre.

Meppem, T. and R. Gill. (1998) "Planning for sustainability as a learning concept." *Ecological Economics* 26: 121-137.

Millennium Ecosystem Assessment. (2001) *Millennium Ecosystem Assessment: Objectives, Audience, Process and Conceptual Framework.* Washington DC: Interim Millennium Ecosystem Assessment Secretariat c/o World Resources Institute. September 8, 2001.
< http://www.ma-secretariat.org/english/publications/conceptual.framework.pdf>

Millennium Ecosystem Assessment. (2001a) "Sub-Global Working Group Summary.", in MEA, *Millennium Ecosystem Assessment. Strengthening Capacity to Sustainably Manage Ecosystems for Human Well-Being.* 2nd Meeting of the MEA Assessment Panel, June 30 – July 1, 2001, London, UK. Washington, DC: Interim MEA Secretariat.

Miller, C. (1998) *Extending Assessment Communities to Developing Countries.* Global Environmental Assessment Project report no. E-98-15. Cambridge, MA: Kennedy School of Government, Harvard University.

Moldan, B. and S. Billharz., eds. (1997) *Sustainability Indicators. Report of the Project on Indicators of Sustainable Development.* Chichester, UK: John Wiley & Sons.

Moltke, K. von (2001) *On Clustering International Environmental Agreements.* Winnipeg, MB: International Institute for Sustainable Development. January 6, 2002.
<http://www.iisd.org/pdf/trade_clustering_meas.pdf>

National Council for Science and the Environment. (2001) *Draft Recommendations.* 2nd National Conference on Science, Policy and the Environment, December 6-7, Washington D.C. Washington, D.C.: NCSE. January 2, 2002.
<http://www.cnie.org/NCSEconference/2001conference/page.cfm?FID=1416#Information>

National Research Council, Board on Sustainable Development. (1999) *Our Common Journey: A Transition Toward Sustainability.* Washington DC: National Academy Press. October 2, 2000. < http://www.nap.edu/catalog/9690.html>

Ney, S. and M. Thompson. (2000) "Cultural discourses in the global climate change debate". Presented at the *Conceptual Frameworks for Mitigation and Assessment from the Perspective of Social Science* workshop, March 20-21, Karlsruhe, Germany.

Observatorio del Desarrollo. (2001) *GEO América Latina y el Caribe. Estadísticas, expertos y documentación sobre el medio ambiente.* San Jose, Costa Rica: Observatorio del Desarrollo. CD-ROM.

Organization for Economic Cooperation and Development. (n.d.) *Donor Support for Institutional Capacity Development in Environment: Lessons Learned.* Paris: OECD.

Organization for Economic Cooperation and Development. (2001) *OECD Environmental Outlook.* Paris: OECD.

Organization for Economic Cooperation and Development. (1993) *OECD Core Set of Indicators for Environmental Performance Reviews.* Environmental Monographs No. 83. Paris: OECD. June 2, 2001.
< http://www1.oecd.org/env/docs/gd93179.pdf>

Parris, T., C.A. Zracket, and W.C. Clark. (1998) *Usable Knowledge for Managing Responses to Global Environmental Change: Recommendations to Promote Collaborative Assessments and Information Systems.* ENRP Discussion Paper E-98-26. Cambridge, MA: Kennedy School of Government, Harvard University.

Patton, M.Q. (1986) *Utilization Focused Evaluation.* Newbury Park: Sage Publications.

Pintér, L. (1997) "De-mystifying sustainable development through performance measurement," in A. R. Magalhães, *Sustainable Development - Implications for World Peace.* Austin, TX: Lyndon B. Johnson School of Public Affairs and The University of Texas at Austin, pp. 61-73.

Pintér, L., K. Zahedi and D. Cressman. (2000) *Capacity Building for Integrated Environmental Assessment and Reporting. Training Manual.* Winnipeg: International Institute for Sustainable Development for the United Nations Environment Programme.

Raskin, P. *et al.* (1998) *Bending the Curve: Toward Global Sustainability.* PoleStar Series Report no. 8. Stockholm: Stockholm Environment Institute. June 1, 2000.
<http://www.tellus.org/seib/publications/bendingthecurve.pdf>

Ravetz, J. (1999) “Citizen participation for integrated assessment: New pathways in complex systems.” *International Journal of environment and Pollution* 11(3): 331-350.

Rawls, J. (1987) “The idea of an overlapping consensus.” *Oxford Journal of Legal Studies* 7(1): 1-25.

Rijksinstituut voor Volksgezondheid en Milieu (1999) *RIVM Postscripts to Abstracts of Selected Environmental Outlook Reports.* Report No. ENV/EPOC/RD(99)3. Bilthoven: RIVM.

Rijksinstituut voor Volksgezondheid en Milieu (2000) *Natuurbalans.* Bilthoven: RIVM (in Dutch). January 20, 2001. <http://www.rivm.nl/milieu/nationaal/nb2000/>

Rijksinstituut voor Volksgezondheid en Milieu. (2000a) *Environmental Balance 2000 – A review of the State of the Environment in the Netherlands.* Bilthoven: RIVM. January 20, 2001. <http://www.rivm.nl/environment/national/mb2000/>

Rijksinstituut voor Volksgezondheid en Milieu. (2000b) *National Environmental Outlook-5. 2000-2030.* Bilthoven: RIVM. January 20, 2001. <http://www.rivm.nl/environment/national/mv5/index.html>

Rotmans, J. (1998) "Methods for IA: The challenges and opportunities ahead." *Environmental Modeling and Assessment* 3(3): 155-179.

Rump, P. (1996) *State of the Environment Reporting: Source Book of Methods and Approaches.* Nairobi: UNEP/DEIA. TR.96-1.

Sabatier, P.A. and H.C. Jenkins-Smith. (1993) *Policy Change and Learning: An Advocacy Coalition Approach.* Boulder, CO: Westview Press.

Schön, D.A. and M. Rein. (1994) *Frame Reflection – Toward the Resolution of Intractable Policy Controversies.* New York: Basic Books.

Sen. A. (1999) *Commodities and Capabilities.* Oxford: Oxford University Press.

Shannon, C.E. and W. Weaver. (1949) *The Mathematical Theory of Communication.* Urbana, IL: University of Illinois Press.

Shindler, B. and K.A. Cheek. (1999) "Integrating citizens in adaptive management: A prepositional analysis." *Conservation Ecology* 3(1). December 11, 2001. <http://www.consecol.org/Journal/vol3/iss1/art9/index.html>

Siebenhüner, B. (2001) *How do scientific assessments learn? A comparative study of the IPCC and LRTAP.* Draft BCSIA Discussion Paper. Cambridge, MA: Environment and Natural Resources Program, Kennedy School of Government, Harvard University.

Simon, H.A. *et al.* (1986) "Decision Making and Problem Solving," in National Academy of Sciences, *Research Briefings1986: Report of the Research Briefing Panel on Decision Making and Problem Solving.* Washington, D.C.: National Academy Press.

Simonett, O. *et al.* (1998) *UNEP's Capacity Development in Environmental Information Management. The Environment and Natural Resources Management Information Network (ENRIN) in Central and Eastern Europe and the NIS.* GRID-Arendal Occasional Papers 2. Arendal, Norway: UNEP/GRID.

The Social Learning Group. (2001) *Learning to Manage Global Environmental Risks: A Comparative History of Social Responses to Climate Change, Ozone Depletion and Acid Rain*. Cambridge, MA: MIT Press.

Thorngate, W. (1996) *Measuring the Effects of Information on Development.* May 12, 2001. <http://www.idrc.ca/books/focus/783/thorn2.html>

Tolba, M. K. and O. A. El-Kholy, eds. (1992) *The World Environment 1972-1992.* London, UK: Chapman Hall for UNEP.

Tóth, F. (1989) *Policy Exercises.* Report RR-89-2. Laxenburg: IIASA.

United Nations. (1997) *UN General Assembly Resolution No. A/RES/S-19/2, Programme for the Further Implementation of Agenda 21.* New York: UNGASS, September 19, 1998.

United Nations. (1993) *Agenda 21: Programme of Action for Sustainable Development.* New York: United Nations.

United Nations Development Programme. (2001) *Human Development Report 2001.* New York: United Nations Development Programme.

UNEP. (1982) *The World Environment 1972-1982.* Dublin: Tycooly International.

UNEP. (1991) *The State of the World Environment.* Nairobi, Kenya: UNEP.

UNEP. (1997) *Global Environment Outlook.* Oxford: Oxford University Press. <http://www.unep.org/unep/eia/geo1/ch/toc.htm>

UNEP. (1997a) *UNEP Governing Council Decision No. Na.97-1040.* Nairobi: UNEP.

UNEP. (1997b) *Report of the Inaugural Meeting of UNEP's GEO Collaborating Centre Network. Groningen, the Netherlands, 3-5 March 1997.* Report no. UNEP/DEIA/MR.97-6. Nairobi: UNEP.

UNEP. (1997c) *Report of the GEO-2 Planning Meeting with UNEP's GEO Collaborating Centres. Beijing, China, 19-23 May 1997.* Report no. UNEP/DEIA/MR.97-07. Nairobi: UNEP.

UNEP. (1998) *United Nations System-Wide Earthwatch Reference Paper for the UNEP Environmental Observing and Assessment Strategy.* Geneva: United Nations System-Wide Earthwatch. June 2, 2001. <http://unep.ch/earthw/unepstrf.htm>

UNEP. (1998a) *First GEO-2 Drafting Meeting with GEO Collaborating Centres. Manama, Bahrain, 15-20 November 1997.* Report no. UNEP/DEIA/MR.98-1. Nairobi: UNEP.

UNEP. (1998b) *Second GEO-2 Drafting Meeting with GEO Collaborating Centres Centres. Brasilia, Brazil, 2-6 February 1998.* Report no. UNEP/DEIA/MR.98-2. Nairobi: UNEP.

UNEP. (1999) *Caribbean Environment Outlook.* Mexico City: UNEP-ROLAC.

UNEP. (1999a) *Western Indian Ocean Environment Outlook.* Mexico City: UNEP-ROLAC.

UNEP. (2000) *Global Environment Outlook 2000.* London: Earthscan Publications. May 20, 2000. <http://www.grida.no/geo2000/english/>

UNEP. (2000a) *Environmental Observing and Assessment Strategy. Extended Version.* Geneva: United Nations System-Wide Earthwatch. <http://www.unep.ch/earthw/unepstr5.htm>

UNEP. (2000b) *Environmental Observing and Assessment Strategy – Activities for Strategy Implementation.* Geneva: United Nations System-Wide Earthwatch. <http://www.unep.ch/earthw/unepstac.htm>

UNEP. (2000c) *Global Environment Outlook Prospectus.* Nairobi: UNEP.

UNEP. (2000d) *GEO-3 Start-Up Meeting. Nairobi, 15-19 November 1999.* Report no. UNEP/DEIA&EW/MR.2000-1. Nairobi: UNEP.

UNEP. (2000e) *GEO-3 First Production Meeting Report. Asian Institute of Technology, Bangkok, Thailand, 4-6 April 2000.* Nairobi: UNEP. June 20, 2001. < http://geoss.unep.org/library/filearchive/Library/2_68.doc>

UNEP. (2001a) *Brief Update on the Twenty-First Session of the Governing Council of the United Nations Environment Programme / Global Ministerial Environment Forum (GEMF), 5-9 February 2001, Nairobi.* Nairobi: UNEP. Memo circulated to GEO Collaborating Centres

UNEP. (2001b) *UNEP website.* Nairobi: UNEP. <http://www.unep.org>

UNEP Division of Environmental Policy Implementation. (n.d.) *Environmental Capacity Development by the United Nations Environment Programme.* Nairobi: UNEP.

UNEP/GRID-Arendal. (2002) *The Structure of GEO-Compatible Assessment.* Arendal, Norway: UNEP/GRID-Arendal. December 20, 2001 < http://www.ceroi.net/GEOdev/Index.htm>

UNEP/GRID-Arendal. (1998) *State of the Environment Reporting on the Internet – Cookbook.* Arendal, Norway: UNEP/GRID.

UNEP/GRID-Geneva. (2001) *GEO Data Portal.* Geneva: UNEP-GRID-Geneva. <http://geo3.grid.unep.ch/>

UNEP Regional Resource Centre for Asia and the Pacific. 2002. *Development of SoE Database.* Bangkok: UNEP-RRC.AP. <http://www.eapap.unep.org/reports/soe/soedb.cfm>

United Nations System-Wide Earthwatch. (1996) *Progress Report. Earthwatch Working Party 3. New York, 17-18 January 1996. Working Paper UNEP/EWWP3/1.* Geneva: UNEP-EW. < http://www.unep.ch/earthw/ewwp3pr1.htm>

Universalia. (2001) *Global Environment Outlook: User Profile and Impact Study.* Montreal: Universalia.

United States Environmental Protection Agency. (1994) *Guidance for the Data Quality Objectives Proces*s. Washington, D.C.: Office of Research and Development, United States Environment Protection Agency. December 20, 2000. <http://www.epa.gov/region10/www/offices/oea/epaqag4.pdf>

VanDeveer, S. (1998) *European Politics with a Scientific Face: Transition Countries, Integrated Environmental Assessment, and LRTAP Cooperation.* ENRP Discussion Paper. Cambridge, MA: Harvard University.

VanDeveer, S. and G. Dabelko. (2001) "It's capacity, stupid: International assistance and national implementation." *Global Environmental Politics* 1(2): 18-29.

Vaz, S.G. *et al.* (2001) *Reporting on Environmental Measures: Are We Being Effective?* Environment issue report no. 25. Copenhagen: EEA.

Vennix, J. (1996) *Group Model Building: Facilitating Team Learning Using System Dynamic*s. New York: John Wiley and Sons.

Vitousek, P.M. *et al.* (1997) "Human domination of Earth's ecosystems." *Science*, 277(5325): 494-499. October 27, 2001. <http://www.well.com/user/davidu/domination.html>

Walters, C. (1986) *Adaptive Management of Renewable Resources.* New York: McMillan.

Ward, B. and R. Dubos. (1972) *Only One Earth. The Care and Maintenance of a Small Planet.* New York: W. W. Norton & Company.

Watanabe, S. (1975) "Informational relativity," in A. Debons and W.J. Cameron (Eds.), *Perspectives in Information Science,* Leyden: Noordhoff, pp. 119-126.

World Commission on Environment and Development. (1987) *Our Common Future.* Oxford: Oxford University Press.

Webber, D.J. (1991-92) "The distribution and use of policy knowledge in the policy process". *Knowledge and Policy, The International Journal of Knowledge Transfer and Utilization,* 4(4).

Weiss, C.H. (1977) "Research for policy sake: The enlightenment function of social research." *Policy Analysis* 3: 531-545.

Weiss, C.H. (1973) "Where politics and evaluation meet." *Evaluation* 1(3): 37-45.

Wilbanks, T. J., and R. W. Kates. (1997) *Global Change in Local Places: How Scale Matters.* Discussion Paper No. 1. Washington, D.C.: Association of American Geographers.

Wilson, P. (1995) "Some consequences of information overload and rapid conceptual change," in J. Olaisen, E. Munc-Pedersen and P. Wilson (eds.), *Information Science – From the Development of the Discipline to Social Interaction.* Oslo: Scandinavian University Press, pp. 21-34.

World Meteorological Organization. (1997) *GHOST: Global Hierarchical Observing Strategy.* Geneva: WMO. June 7, 2001 <http://www.wmo.ch/web/gcos/pub/ghost.html>

Woerden, J. van *et al.* (1999) *Data Issues of Global Environmental Reporting. Experiences from GEO-2000.* Bilthoven: RIVM for the United Nations Environmental Programme.

World Bank. (2001) *World Development Indicators 2001.* Washington, D.C.: World Bank.

World Economic Forum. (2002) *2002 Environmental Sustainability Index.* New Haven: Yale Center for Environmental Law and Policy. February 15, 2002. <http://www.ciesin.org/indicators/ESI/ESI2002_11FEB02tot.pdf>

World Wildlife Fund. (2000) *2000 Living Planet Report.* Gland: World Wildlife Fund.

World Resources Institute. (1996) *World Directory of Country Environmental Studies.* Washington DC: WRI.

World Resources Institute. (2000) *World Resources Report 2000.* Washington, D.C: World Resources Institute.

World Resources Institute. (2001) *Earth Trends: The Environmental Information Portal.* Washington, DC: WRI. May 24, 2001. < http://earthtrends.wri.org/>

Young, O. *et al.* (1999) *Institutional Dimensions of Global Environmental Change. Science Plan.* Bonn: IHDP Report No. 9. November 12, 2000. < http://www.uni-bonn.de/ihdp/IDGECSciencePlan/>

Young, C.J. and J. Comptois. (1979) “Increasing congressional utilization of evaluation,” in F. Zweig (ed.) *Evaluation in Legislation.* Beverly Hills: Sage Publications.

# APPENDICES

# APPENDIX 1: UNEP GOVERNING COUNCIL DECISION 18/27 C

(United Nations System-Wide Earthwatch 1996)

C. New state-of-the-environment report

The Governing Council,

Recalling General Assembly resolution 2997 (XXVII) of 15 December 1972, by which the Assembly, inter alia, conferred the following functions upon the United Nations Environment Programme: to keep under review the world environmental situation; to coordinate, review and assess environmental programmes within the United Nations system; and to finance, wholly or partly, the costs of new environmental initiatives undertaken within the United Nations system,

Recalling also its decision 17/6 of 21 May 1993 on state-of-the-environment reports,

Further recalling the reconfirmation of the Programme's mandate by the United Nations Conference on Environment and Development,

Noting the call of the United Nations Conference on Environment and Development for even greater efforts to coordinate environment and development activities in the United Nations system,

Noting with appreciation previous reports of the Executive Director on the state of the environment,

Stressing the overall objective of the integration of environment and development issues and actions, at the national, regional and international levels, including within the United Nations system,

Deeply concerned that consensus cannot be reached on several essential issues in the field of environment and development within the United Nations system, which delays implementation of the principles and recommendations of Agenda 21,[6]

1. Requests the Executive Director to prepare a new, comprehensive report on the state of the world environment, which will consist of the following three parts:

(a) The present state of the global environment;

(b) The state of the global environment in the year 2015;

(c) The response: findings, conclusions and recommendations;

2. Recommends the inclusion in parts (a) and (b) of the report all essential problems of and threats to the environment, inter alia , the environmental status of the main components of the global ecosystem (waters, forests, soils and farming lands, ozone

layer, etc.), basic trends in environmental change (for example, climate change, coastal and marine degradation, desertification, deforestation and habitat loss, pollution, soil degradation, ozone depletion, etc.); and the global effects of expected development growth, population increase and main trends in consumption, production and urbanization patterns (for example, energy consumption, transportation and sanitation problems, waste disposal, land reclamation and destruction, etc.);

3. Also recommends the inclusion in part (b) of the report of the expected impact of population increase, consumption and production patterns and economic development on the environment;

4. Further recommends the inclusion in part (c) of the report recommended measures and actions that could effectively reverse unwelcome trends and challenge principal threats to the environment and also specific institutional and legal measures for the implementation of proposed actions;

5. Requests that preparation of the report, which shall be undertaken within existing resources, be based primarily on the existing data collected and prepared by the United Nations Environment Programme, in close cooperation with the United Nations Development Programme, the World Health Organization, the Food and Agriculture Organization of the United Nations, the World Bank and other United Nations agencies and bodies, and on the results of research and studies by public and private scientific and statistical institutions engaged in formulation of environmental and development assessments and forecasts;

6. Also requests the Executive Director to consult periodically with the Committee of Permanent Representatives on the preparation of the report;

7. Further requests the Executive Director to submit a first report for the consideration of the Governing Council at its nineteenth session.

## APPENDIX 2: INTERVIEWEES

Interviewee represents:

1. Global Environment Outlook Collaborating Centre
2. United Nations
3. Non-United Nations multinational
4. Corporate
5. Non-governmental organization/other
6. Academic
7. Governmental

| Organization | Contact | Address | Date | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Notes |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Tellus Institute | Paul Raskin<br>President<br>Tel: +1-617-266-5400<br>Email: praskin@tellus.org | Tellus Institute<br>Boston, MA, United States | Boston, MA, Jan. 30, 2001 | ✔ | | | | ✔ | | | |
| 2. Ministry of Public Health and the Environment (Rijksinstituut voor Volksgezondheid en Milieu, RIVM) | Leon Braat<br>Tel:<br>Email: leon.braat@rivm.nl | RIVM<br>Antonie van Leeuwenhoeklaan 9<br>Bilthoven, P.O. Box 1, 3720 Bilthoven<br>The Netherlands | Bilthoven, NL, November 14, 2000 | ✔ | | | | | | ✔ | |
| 3. World Wide Fund for Nature (WWF) | Jonathan Loh<br>Project Manager<br>Tel: +41-22-364-9111<br>Email: jloh@wwfint.org | WWF<br>Gland, Switzerland | Gland, CH, December 7 | | | | | ✔ | | | |
| 4. Bureau for Environmental Assessment, National Institute of Public Health and the Environment (Rijksinstituut voor Volksgezondheid en Milieu, RIVM) | Rob Maas<br>Head<br>Tel:<br>Email: rob.maas@rivm.nl | RIVM<br>Antonie van Leeuwenhoeklaan 9<br>Bilthoven, P.O. Box 1, 3720 Bilthoven<br>The Netherlands | Bilthoven, NL, November 14, 2000 | ✔ | | | | | | ✔ | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 5. National Institute for Health and the Environment (Rijksinstituut voor Volksgezondheid en Milieu, RIVM) | Jan Bakkes<br>Head of UNEP Collaborative Centre<br>Tel: +31-30-274-3112<br>Email: jan.bakkes@rivm.nl | RIVM<br>P.O. Box 1, 3720 BA<br>Bilthoven, The Netherlands | Bilthoven and Utrecht, NL, November 14, 2000 | ✔ | | | | | | ✔ | |
| 6. Division for Technology, Industry and Economics, United Nations Environment Programme (UNEP DTIE) | Jacqueline Aloisi de Larderel<br>Tel: +33-1-4437-1450<br>Email: j.aloisi@unep.fr<br>Secretary: Tess Cieux | UNEP-DTIE<br>Tour Mirabeau 39-43, quai André Citröen 75739<br>Paris, Cedex 15<br>France | Paris, France, December 13, 2000 | | ✔ | | ✔ | | | | |
| 7. IUCN – The World Conservation Union | Nancy MacPherson<br>Monitoring and Evaluation Coordinator<br>Tel: +41-22-999-0271<br>Email: nmm@iucn.org | IUCN<br>Rue Mauverney 28, 1196 Gland<br>Switzerland | Gland, CH, December 7, 2000 | | | | | ✔ | | | |
| 8. Environment Canada, Policy and Communications | Sabrina Barker<br>Tel: (819) 953-0912<br>Email: sabrina.barker@ec.gc.ca | Environment Canada<br>10 Wellington Street<br>Hull, Quebec K1A 0H3 Canada | Telephone, Feb. 6, 2001 | | | | | | | ✔ | |
| 9. Organization for Economic Co-operation and Development (OECD) | Lars Mortensen<br>Tel:<br>Email: lars.mortensen@oecd.org | OECD<br>15 boulevard Admiral Bruix<br>16th Arondissement, 5th Floor, Rm. 5021<br>Paris, France | Paris, France, December 13 | | | ✔ | | | | | |
| 10. Arabian Gulf University (AGU) | Osama El-Kholy<br>Tel:<br>Email: osama@agu.edu.bh | Arabian Gulf University<br>Manama, Bahrain | Manama, BH, November 20, 2000 | | | | | | ✔ | | |
| 11. United Nations System-Wide Earthwatch | Arthur Dahl<br>Coordinator<br>Tel: +41-22-917-8207<br>Email: dahla@unep.ch | International Environment House<br>13 Chemin des Anemones<br>CH-1219 Chatelaine, Geneva<br>Switzerland | Geneva, CH, December 13, 2000 | | ✔ | | | | | | Tape problem |
| 12. Bellagio Forum for Sustainable Development | Michael Hanssler<br>Executive Director<br>Tel: +33-450-990-430<br>Email: mhanssler@bfsd.org | The Bellagio Forum for Sustainable Development<br>c/o Deutsche Bundesstiftung Umwelt<br>An der Bornau 2, 49090 Osnabrück<br>Germany | Geneva, CH, December 12, 2000 | | | | | ✔ | | | Tape problem |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 13. International Development Research Centre (IDRC), Evaluation Unit | Fred Carden<br>Senior Program Specialist<br>Tel: (613) 236-6163, 2107<br>Email: fcarden@idrc.ca | 250 Albert Street<br>PO Box 8500 Ottawa K1G 3H9<br>Canada | Telephone, Feb. 19, 2001 | | | | | ✔ | | | |
| 14. United Nations Environment Programme (UNEP) | Marion Cheatle<br>Chief<br>Tel: +254-2-623520<br>Email: marion.cheatle@unep.org | State of Environment Assessment Unit<br>UNEP<br>P.O.Box 30552, Nairobi<br>KENYA | Telephone, Feb. 19, 2001 | | ✔ | | | | | | |
| 15. Millennium Ecosystem Assessment (MEA) | Walter Reid<br>Tel: +1-202-729-7794<br>Tel in Seattle: +1-206-782-7963<br>Email: waltreid@attglobal.net | Interim Millennium Assessment Secretariat<br>c/o World Resources Institute, 10G Street, N.E., Washington, D.C. 20002 USA | Telephone, Feb. 9, 2001 | | ✔ | | | | ✔ | | |
| 16. Observatorio del Dessarrollo, University of Costa Rica | Edgar Gutierrez-Espeleta<br>Director<br>Tel: +506 207 3327<br>Email: egutierr@cariari.ucr.ac.cr | Observatorio del Desarrollo<br>University of Costa Rica<br>San José, Costa Rica<br>2060 San Jose, Costa Rica | Medford, MA, February 2001 | ✔ | | | | | ✔ | | |
| 17. United Nations Intellectual History Project, The Graduate Program, The City University of New York | Richard Jolly<br>Senior Research Fellow<br>Tel: +1-212-817-1920 B<br>Tel: +1-212-308-3473<br>Email: richajolly@aol.com | United Nations Intellectual History Project<br>The Graduate Program<br>The City University of New York<br>365 Fifth Avenue<br>New York, NY 10016 USA | Telephone, Feb. 7, 2001 | | | | | | ✔ | | |
| 18. World Resources Institute (WRI) | Allen Hammond<br>Chief Information Officer and Senior Scientist<br>Tel: +1-202-729-7777; 662-2574<br>Email: allen@wri.org | WRI<br>10 G Street NE, Suite 800<br>Washington D.C., 20003 USA | Telephone, Feb 15, 2001 | ✔ | | | | ✔ | | | |
| 19. Moscow State University (MSU) | Genady Golubev<br>Professor<br>Tel: +7-095-939-3962; +7-095-336-2353; +7-095-932-8836<br>Email: ggolubev@mtu-net.ru | Faculty of Geography<br>Moscow State University<br>117899, Moscow<br>Russian Federation | Telephone, March | ✔ | | | | | ✔ | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 20. The Hague Coordination Office of the Global Programme of Action for the Protection of the Marine Environment from Land-Based Activit. | Veerle Vandeweerd<br>Coordinator<br>Tel: +31-70-311-4461/4472<br>Email: v.vandeweerd@unep.nl | Global Programme of Action for the Protection of the Marine Environment from Land-Based Activities<br>Vuurtorenweg 35-37, Scheweningen,<br>Den Haag 2500 BE<br>The Netherlands | Den Haag, NL,<br>Feb. 28, 2001 | | ✔ | | | | | | |
| 21. Bangladesh Center for Advanced Studies (BCAS) | Atiq Rahman<br>President<br>Tel: +880-2-811-3977<br>Email: atiq.r@bdcom.com and bcas@bdonline.com | Bangladesh Center for Advanced Studies<br>Dhaka<br>Bangladesh | San Miguel<br>Regla, Mexico,<br>April 4, 2001 | ✔ | | | | ✔ | | | |
| 22. UNEP/Asian Institute of Technology (AIT) | Surendra Shresta<br>Regional Coordinator for Asia and the Pacific<br>Tel: +662 516 2124/524 5365<br>Email: surendra@ait.ac.th | Outreach Building, Room No.304<br>Asian Institute of Technology<br>P.O. Box 4, Klong Luang,<br>Pathumthani 12120, THAILAND | Telephone,<br>March 16, 2001 | | ✔ | | | | | | |
| 23. Musokotwane Environment Resource Centre for Southern Africa (IMERCSA) | Munyaradzi Chenje<br>Program Officer, UNEP<br>Former Director, IMERCSA<br>Tel:<br>E-mail:<br>munyaradzi.chenje@unep.org | State of Environment Assessment Unit<br>UNEP<br>P.O.Box 30552, Nairobi<br>KENYA | Telephone,<br>February 19,<br>2001 | | | | | ✔ | | | |
| 24. United Nations Division for Sustainable Development (UN-DSD) | Lowell Flanders<br>Assistant Director<br>Tel: +1-212-963-8792<br>Email: flanders@un.org | UN Division for Sustainable Development<br>New York, NY 10017<br>United States | New York, NY,<br>Feb. 12, 2001 | | ✔ | | | | | | |
| 25. Central European University (CEU) | Ruben Mnatsakanian<br>Professor and Department Head<br>Tel: +36-1-327-3071<br>Email: mnatsaka@ceu.hu | Central European University<br>H-1051 Budapest, Nador u.9<br>Hungary | Telephone,<br>March 21, 2001 | ✔ | | | | | ✔ | | |
| 26. Peking University and State Environment Protection Agency (SEPA) | Shengji Luan, Professor<br>Tel:<br>Email: s.luan@ces.pku.edu.cn | Beijing<br>China | San Miguel<br>Regla, Mexico,<br>April 6, 2001 | ✔ | | | | | ✔ | | |
| 27. European Environment Agency (EEA) | Ronan Uhel<br>Tel:<br>Email: | European Environment Agency | San Miguel<br>Regla, Mexico,<br>April 7, 2001 | | | | | | | | |

# APPENDIX 3: INTERVIEW PROTOCOL

The interviews were conducted in the narrative style; thus the questions in the protocol were for general guidance. Phrases might be changed, questions omitted and new questions added as required by the logic, flow and direction of a specific conversation. By so doing my purpose was to maintain the integrity of the story or narrative, if one was provided, and to allow exploration in relevant directions.

Some of my interviewees were situated at GEO Collaborative Centres and familiar with the details of GEO; others were either leading comparable initiatives or had taken interest in the science or policy of global integrated assessment and reporting. Interviewees in the second group might not be familiar with GEO *per se*, so in the course of the interview I might rephrased questions in a more general form or asked only questions that were applicable.

## 1. Introduction and context setting

1.1. Personal introduction
1.2. The purpose of this research is…
1.3. I thought about you as an interviewee because …
1.4. Professional conduct: research protocols, recording the interview, confidentiality
1.5. Questions/concerns

## 2. Organizational overview related to assessment and reporting

2.1. Does your organization have a specific mandate related to assessment and reporting?

*2.1.1. How is the mandate formalized?*
*2.1.1.1. Do you have an organizational unit specifically dealing with assessment and reporting?*
*2.1.2. Is your organization formally linked to the GEO program? How? What is your role?*

## 3. Participation and network design

3.1. What are the tradeoffs associated with participation, assessment carried out by a polycentric network versus a single agency?

*3.1.1. What are the risks and opportunities associated with participation?*
*3.1.2. What specific challenges do you encounter when working on assessments in a participatory or network setting? Can you give an example from your practice?*

*3.1.3. What are your organization's rights and responsibilities in the assessment network?*
*3.1.4. How does the participatory aspect of the assessment affect its results?*
*3.1.5. Does it affect the usefulness of the information produced? How? Can you give an example?*

3.2. Is the composition of the network suitable for the task?

*3.2.1.1.If you are a member of the GEO network, is your role in the GEO network formalized? How?*
*3.2.2. Are there suitable membership criteria and are they systematically applied?*
*3.2.3. Is there an optimal network size? If there is, should this be used to draw limits to participation?*
*3.2.4. Is the current division to Collaborative and Associated Centres appropriate? Are there additional membership tiers that would be useful?*

3.3. What helps keep the network together?

*3.3.1. How do you interact with other members of the network?*
*3.3.2. Among all types of interactions what are the most important and what are less important? Why?*
*3.3.3. What are the essential factors that allows your organization to be involved?*
*3.3.4. What is your view of GEOSS and the potential for this and other types of electronic communication?*
*3.3.5. How important are capacity building activities undertaken as part of the assessment process?*

## 4. The assessment and reporting process

4.1. How is the process planned?

*4.1.1. Do you see assessment process planning as the sole responsibility of a lead agency?*
*4.1.2. Are network members sufficiently involved in process design? What is your role, if any?*
*4.1.3. In your view, is there or should there be a room for mid-course correction in the process plan? What could justify a mid-course adjustment?*

4.2. Does the process support the assessment's overall objectives?

*4.2.1. Are the steps in the assessment process properly sequenced?*
*4.2.2. What are the most critical phases to increase GEO's relevance?*

4.3. What is the role of participation?

*4.3.1. Is participation a central element of the assessment strategy or an add-on?*
*4.3.2. In what stages of your process does external stakeholder participation take place?*
*4.3.3. How does participation actually happen in various stages of the assessment? What is your role?*
*4.3.4. Who decides what external stakeholders to involve and how?*
*4.3.5. Do you involve the same set of participants in global and regional or subregional assessments?*
*4.3.6. How do you deal with assessments that require cooperation among several units of your institute?*
*4.3.6.1. How does participation affect the outcomes of the assessment? Can you give a specific example?*

## 5. Assessment and reporting tools and methods

5.1. Is the current GEO methodology appropriate given the objectives?

*5.1.1. Is the current structure (SOE assessment + policy analysis, scenario analysis, vulnerability assessment and recommendations) appropriate?*
*5.1.2. What is the component that has the potential for the most significant policy impact? How? Can you provide an example?*
*5.1.3. Is there sufficient methodological support provided through the production guidelines?*
*5.1.4. Are the guidelines sufficient to ensure consistency of assessment methods across regions and scales?*

5.2. Are GEO's conceptual and technical tools adequate?

*5.2.1. What are the GEO tools most important for your work?*
*5.2.2. Would these tools be useful for your work apart from GEO? Can you give an example?*
*5.2.3. What further tools would be necessary? How should they be developed?*
*5.2.4. Is the PSIR (or SPIR) framework appropriate for the assessment?*
*5.2.5. What long-term role do you see for a production support system like GEOSS and the global data portal?*

## 6. Links to scientific research agenda setting

6.1. How do you link your work on GEO to the scientific research community in your region?

*6.1.1. Do you see GEO as a **scientific** or as a **science based** process?*
*6.1.2. Do you identify specific networks or groups to work with on your GEO input?*
*6.1.3. What is the nature of their involvement? What is their interest in taking part?*

*6.1.3.1. Can you give an example of a specific link you maintain for GEO with the scientific research establishment in your region?*

6.2. Is there adequate feedback from GEO to the scientific research community?

*6.2.1. What are the most important specific inputs from science in your region?*
*6.2.2. Are your GEO inputs subjected to some form of peer review on the regional level?*
*6.2.3. Are there links to both natural and social science groups?*

6.3. How could the lessons from GEO influence scientific research coordination and agenda setting?

*6.3.1. Can you think of current examples of such influence?*
*6.3.2. How should this connection be made? Through what mechanisms and at what fora?*
*6.3.3. What could be your role as a GEO participant and what the role of UNEP?*

## 7. Targeting the products of assessment

7.1. Over time GEO has evolved a suite of specialized products to communicate with specific audiences. Is there a clear understanding of key audiences in GEO?

*7.1.1. What is the most important audience for GEO?*
*7.1.2. In your region what is the audience that makes the most use of GEO products?*
*7.1.3. Can GEO have multiple audiences? What are the implications of this for the GEO products?*

7.2. Should GEO continue diversify its range of products?

*7.2.1. Is the continued emphasis on the main GEO report appropriate?*
7.2.2. *Should there be more attention paid to electronic publication?*
7.2.3. *Should GEO continue to concentrate on the release of the GEO report as a major event or diversify and organize many smaller events focused on partial results or specialized outputs?*
7.2.4. *What role could CCs play in this strategy?*

# ENDNOTES

[1] The literature on visioning is voluminous. I have to mention particularly my – limited – experience with the *appreciative inquiry* method in the course of a field project in India involving MYRADA, a prominent Indian NGO and IISD in 1999 (International Institute for Sustainable Development 2000). For detailed discussion of the AI method, see Elliott (1999).

[2] I would be inclined to nest these ideas within the 'Policy Reform' scenario of the Global Scenario Group, although there are also elements from 'Great Transitions' (Gallopin *et al.* 1997; Raskin *et al.* 1998).

[3] There are several processes today working towards similar objectives that are constrained by weak political will – as compared with the magnitude and importance of the problem - inadequate resources, conceptual difficulties, and less than sufficient coordination (UNEP 2000a and 200b).

[4] In contrast with the term 'hierarchy' I take 'panarchy' to mean a nested structure of dynamic processes whose constituent components are characterized by relative internal autonomy but also cross-scale, system-wide coordination. The concept has come out of the work of Holling's Resilience Alliance.

[5] Sustainability is an inherently forward looking and multi-dimensional criterion. It has an *ex post* and an *ex ante* dimension. In an *ex post* analysis one would examine whether the actual outcomes of a past decision satisfy sustainability as defined in the context of the particular case. In an *ex ante* situation one would need to anticipate outcomes, and weigh and adjust planned courses of action or policies based on the expected – and by definition uncertain – future results.

Despite the difficulties of operationalizing a sustainability assessment, either in the *ex post* or *ex ante* context, there are serious attempts to make it work. An interesting case in point is the work of Canada's Commissioner for the Environment and Sustainable Development, sitting in the office of the Auditor General. The Commissioner has the power to request detailed audits of any departmental program or policy from the perspective of sustainability, requiring branches of government not normally accustomed to the idea of environment/development thinking to apply the lens of sustainability to their own work (Office of the Auditor General of Canada 2001). This could lead to situations where a departmental program that is seen and perhaps demonstrated as successful based on narrow evaluation criteria becomes seen less so if it fares poorer when viewed through the 'lens of sustainability'. Given the attention the Commissioner's reports receive in Parliament, the Cabinet and the media, there are compelling reasons to take the reports seriously.

[6] I find the term 'developing countries' problematic for several reasons, but will use it for lack of a better one. Also, I find that in scientific and policy parlance commonly stuck in the developed-developing dichotomy transition economies are left out, or assumed to be clustered with one or another. For the purposes of this study I find the issues faced by transition economies take them closer to the developing rather than developed end of the spectrum. In order to avoid unnecessarily cluttering language most of what I say when I talk about developing should be taken to apply also to transition countries.

[7] In the case of GEO its products and the GEO process are typically mentioned together. Many of my interviewees said that regular involvement in the GEO process was critically important. A formal, institutionalized role, for example membership in the Collaborating Centre network is a precondition of participation. It is easy to see that CCs consider the process important because it provides them continuous recognition, prestige, perhaps resources, and learning opportunity. From the perspective of UNEP and GEO it is an opportunity to increase legitimacy.

[8] Core elements of UNEP's mandate as described in the Nairobi Declaration (UNEP 1997a):

- "To analyze the state of the global environment and assess global and regional environmental trends, provide policy advice, early warning information on environmental threats, and to catalyze and promote international cooperation and action, based on the best scientific and technical capabilities available;
- To further the development of its international environmental law aiming at sustainable development, including the development of coherent interlinkages among existing international environmental conventions;

- To advance the implementation of agreed international norms and policies, to monitor and foster compliance with environmental principles and international agreements and stimulate cooperative action to respond to emerging environmental challenges;
- To strengthen its role in the coordination of environmental activities in the United Nations system in the field of the environment, as well as its role as an Implementing Agency of the Global Environment Facility, based on its comparative advantage and scientific and technical expertise;
- To promote greater awareness and facilitate effective cooperation among all sectors of society and actors involved in the implementation of the international environmental agenda, and to serve as an effective link between the scientific community and policy makers at the national and international levels;
- To provide policy and advisory services in key areas of institution building to Governments and other relevant institutions."

[9] In contrast with what could be seen as a narrower mandate, recent MEA releases refer to the MEA producing the first report card on the global environment. Besides being inaccurate, this may lead the assessment to very broad tasks further and sooner than envisioned by some of its original supporters.

[10] Interview data reflecting this view was collected in February 2001, before a significant expansion of the GEO team was known.

[11] UNEP is supported only by voluntary – therefore volatile and not particularly predictable - contributions of member states versus other UN organizations, such as FAO or WHO that receive funding from the general UN budget. In personal communication a senior government official from an industrialized country pointed out to me that had there been a *requirement* for higher level funding for UNEP would pose no financial or political difficulties whatsoever. The mere fact that the contribution to UNEP is *voluntary*, however, means that requests for higher contribution are easily rejected on the grounds that increasing or decreasing support for the donor country is inconsequential.

[12] For the first time, a detailed example has been included in the GEO-3 production guidelines, although the template has probably changed many times as the assessment progressed (UNEP 2000d). Also, the example did not cover all sections of the assessment.

[13] The development of an electronic template or 'tutorial' was initiated by UNEP-ROLAC, UNEP-GRID-Arendal and IISD in late 2001. As this is work in progress at the time of finalizing my thesis, no details have been published, but an outline is already available (UNEP-GRID-Arendal 2002).

[14] Consider this example from my interview with Nancy MacPherson:

> "I wonder if you ever met Diana Lee Smith, she was on one of the first pilot assessments. At the time she was working in East Africa, she is a Kenyan, for UNHCS and on this multi-regional project with squatter-women, women in squatter-settlements developing of indicators of sustainability through their eyes. And once she said one of my big achievements in this is getting squatter women accepted as a recognized UN expert group. Because, you know, the UN has this label as UN expert group, and she had this fight with superiors in the UN for two years because they wouldn't accept squatter women as experts in anything. And Diana said they are experts in their life, they have learned to survive in circumstances that you would be dead in two days. So they are experts, and it was a wonderful, wonderful debate … It was a very good example of how you can legitimize and give structure and systematic value to user views and they can be accepted in a system. I don't know how well they did, a little test would be whether Habitat projects use those indicators as telling them anything about squatter women, or whether they revert back to their own 'real' experts and data."

[15] "But we have just gone through two years of trying to define within IUCN, with IUCN's membership, which is thousands of organizations out there what is IUCN's program, where our accountability starts, and where our line of influence but no accountability begins, and I think that's probably if your are looking at assessments, like doing anything else, that's absolutely critical. So you are dealing with a group of people, a group of institutions, a group of users to be really clear about who they are, what they need, who is going to use assessment for what purpose, and who is going to deliver those results. Without clarity on those questions chaos reins and then it tends to be driven by people who have specific agendas, who got resources, who make it their lives' job, and then you get a skewed picture of use, because it's no longer then users driving things, it's people whose job it is to produce data." (Nancy MacPherson interview)

[16] Identifying and building on strengths is part of the underlying philosophy of appreciative inquiry. The link between appreciative inquiry and capacity building in assessment and reporting has not been elaborated, though I believe there is significant potential. Some exploratory, community scale work has been done in India in the joint IISD-Myrada project (Elliott 1999; Pintér 2000).

[17]One of the more interesting and relevant packages is Publikit, produced in collaboration by Publikit SA in Norway and UNEP-GRID/Arendal < http://www.publikit.net/>. The beta version of Publikit was used in producing Southern Africa's most recent SOE report, available on the internet < http://www.ngo.grida.no/soesa/nsoer/index.htm>.

[18] Participant evaluation is built into both the program of individual training events and training materials. Conducting a comprehensive evaluation, including the review of the written feedback from participants of eight sessions on file, has been suggested in a comprehensive capacity development proposal prepared by UNEP and the GEO Capacity Working Group. As of July 2001 the proposal is awaiting funder approval.

[19] The ethical dilemmas associated with the responsibility of the scientist are the subject of an entire field of social science research and many pieces of art. A personal favourite applicable here is Arthur Miller's "*In the Matter of J. Robert Oppenheimer*".

[20] As noted in this publication, the GEO training manual, Michael Keating developed the communication section, with my contributions.

[21] Of course, these trade-offs don't always pay off. To stay with the CEE example, in my interview R. Uhel brought up that according to the results of a recent EU workshop only around 15% of the questions new accession countries need to – and do - respond to in the field of air, waste and water are policy relevant. The policy relevance of the rest is questionable. Needless to say, transforming CEE data collection and analytic system to match those known to be problematic in the West *in the first place* is not only costly, it is also questionable from the ethical perspective. Legitimacy and credibility of the communicator and the uncritically derived legitimacy and credibility of what is being communicated can thus lead to problems and over time actually *erode* both attributes.

[22] For a more comprehensive review of SD indicator selection criteria see e.g., Murcott (1997), Costanza and Tognetti (1996) or Hart, Mazzota and Kellman (1999). Although selection criteria are widely used in SD indicator projects, they are rather soft, qualitative and hard to use in a rigorous evaluation of indicators (although it is also hard to think about any other approach now that is better yet still practical). They are also diverse: as one would be hard pressed to find two identical indicator sets, it would be equally difficult to find agreement on indicator selection criteria. To me personally one of the main conclusions of Maureen Hart's project to come to agreement on a core set of criteria for SD indicators was that criteria *alone*, without knowing context and indicator selection process are insufficient to determine whether an indicator is suitable or not.