ULYSSES

Urban **L**ifest**Y**les, **S**u**S**tainability and Integrated **E**nvironmental A**S**sessment

ULYSSES Working Paper

Jerome R. Ravetz

Integrated Environmental Assessment Forum: developing guidelines for "good practice"

ULYSSES WP-97-1

This paper was originally prepared under a contract from the EC Joint Research Centre, No. 11972-96-06-F1E1 ISP GB. We are grateful to Silivo Funtowicz for permission to reproduce the paper as a ULYSSES Working Paper.

Darmstadt University of Technology ·
Center for Interdisciplinary Studies in Technology

The ULYSSES project

The European research project ULYSSES aims to bridge the gap between environmental science and democratic policy making in the climate domain. For this we study judgements of informed citizens on climate policy and make these judgements available to policy makers. As a support in this process the citizens will be given access to state-of-the-art computer models on environmental change.

Between 1996 and 1999 ULYSSES will conduct and analyse group discussions with citizens in Barcelona, Venice, Athens, Zurich, Frankfurt, Manchester, and Stockholm.

Objective

Policy makers dealing with complex environmental problems need knowledge from environmental science as well as from the social sciences to back up their decisions. Integrated Assessment (IA) research aims at providing useful overviews of relevant problems and elements of possible solutions for this purpose. In this context ULYSSES - short for **U**rban **L**ifest**Y**les, **S**u**S**tainability, and Integrated **E**nvironmental A**S**sessment - develops procedures for including public participation in IA.

Science can provide neither unique descriptions nor unique solutions for truly complex environmental problems. On the other hand, citizens are used to deal with a variety of conflicting yet legitimate interpretations in democracy which is basically government by public debate. In this situation assessments can improve both in quality and in political relevance by combining expert knowledge with public participation.

Methodology

In order to develop a procedure for public participation in Integrated Assessment, ULYSSES designs a discursive process which is based on the focus group method. In a kind of microcosm of social learning small groups of citizens share a moderated discussion on climate risks and options for climate policy. These **IA-Focus Groups** meet approximately five times. They debate on climate policy, have access to relevant information to support their debates, and express their resulting judgement. The range of arguments and judgements expressed by the citizens will then be condensed and made available to interested decision makers.

One of the basic tasks of ULYSSES is to design an interface between these IA-Focus Groups and computer models. Our hypothesis is that citizens can arrive at reasonable and informed judgements on environmental policy if they have the opportunity to share an in-depth debate and if they are provided with relevant information in a suitable format. For this information the IA-Focus Groups organised by ULYSSES have access to state-of-the-art computer models relevant for decision support on environmental issues. These models include IMAGE (the Integrated Model to Assess the Greenhouse Effect), TARGETS (the Tool for Analysing Regional and Global Environment and Health Targets for Sustainability), PoleStar, and NAIADE (the Novel Approach to Imprecise Assessment and Decision Environments).

ULYSSES tests this IA-Focus Group procedure in the domain of urban lifestyles and their connection to climate change. In order to experience the cultural diversity within Europe in our experiments we will conduct IA-Focus Groups in urban regions throughout Europe: Barcelona, Venice, Athens, Zurich, Frankfurt (Rhine/Main), Manchester, and Stockholm.

See also: ***http://www.zit.tu-darmstadt.de/ulysses/***

The ULYSSES working papers

The ULYSSES working papers present work being done in the context of the ULYSSES network. The responsibility for the contents lies with the individual authors.

The ULYSSES research team

Researchers at 10 institutions throughout Europe contribute to ULYSSES:

coordinator

Prof. Carlo Jaeger: TUD, Darmstadt University of Technology, Germany

partners

Dr. Silvio Funtowicz: JRC, Joint Research Centre - Commission of the European Communities, Italy

Prof. Brian Wynne: CSEC, Centre for the Study of Environmental Change, Lancaster University, United Kingdom

Prof. Salvador Giner: IESA, Instituto de Estudios Sociales Avancados de Cataluña, Spain

Åsa Gerger: SEI, Stockholm Environment Institute, Sweden

Prof. Maria Giaoutzi: NTUA, National Technical University of Athens, Greece

Prof. Ferenc Toth: PIK, Potsdam Institute for Climate Impact Research, Germany

Dr. Jill Jäger: IIASA, International Institute for Applied Systems Analysis, Austria

Dr. Jerry Ravetz: RMC, Research Methods Consultancy, United Kingdom

Dr. Gregor Dürrenberger: EAWAG, Swiss Federal Institute for Environmental Science and Technology, Switzerland

ULYSSES is supported by the European Commission, DG XII, RTD Programme ‘Environment and Climate’, area ‘Human Dimensions of Environmental Change’ (Contract No. ENV4-CT96-0212). The **EC scientific officer** responsible for ULYSSES from the side of the European Commission is Dr. Angela Liberatore.

ULYSSES also profits from the collaboration with ZIT, the Center for Interdisciplinary Studies in Technology, at Darmstadt University of Technology.

For further information on ULYSSES please contact the **project assistant**
Dr. Bernd Kasemir (EAWAG, Human Ecology Group, Überlandstraße 133, CH- 8600 Dübendorf, Switzerland. e-mail: kasemir@eawag.ch)

The **ULYSSES** working papers can be ordered at:

ULYSSES
c/o ZIT - Center for Interdisciplinary Studies in Technology
Darmstadt University of Technology, Darmstadt / Germany
Fax.: ++49-6151-16 67 52

Integrated Environmental Assessment Forum, developing guidelines for "good practice"

Jerome R. Ravetz

Introduction 1

1. The Problem - Why "good practice" in Integrated Environmental Assessment? 4

1.1 Decision-making under inadequate certainty. 5

1.2 Quality assessment in its policy context. 6

1.3 Quality assurance: products or processes? 8

2. Traditional approaches to the "good practice" problem. 10

2.1 The Quantitative Path and its Pitfalls. 10

2.2 Uncertainties in Environmental Economics: sophistication or sophistry? 12

2.3 Computers: the perils of GIGO. 14

3. Irremediable Uncertainty and its Implications 19

3.1 A Cautionary Tale. 19

3.2 Coping with uncertainty. 20

3.3 Implications of uncertainty for "good practice". 22

4. Escaping from "scientific immaturity" - examples from history and design. 25

4.1. "Immaturity" 25

4.2 Engineering and Design. 26

4.3 History. 27

4.4 History and the Maturing of a Discipline. 29

4.5 History and the evolution of design and of good practice. 30

5. Living in Interesting, or Post-Normal Times 32

5.1 Professional work as distinct from research 32

5.2 Post-Normal Science and the inadequacy of bare numbers. 33

5.3 "Pedigree" as the crucial aspect of a policy number. 35

5.4 New social frameworks for policy-relevant science. 36

5.5 The Structures of Post-Normal Science. 37

Conclusion: the maturing of a new conception of science through the practice of Integrated Environmental Assessment. 39

“IA has emerged during the last decade, but methods and tools for integrated assessment are still relatively immature. So far, IA has been merely an intuitive process, without concrete rules and matters of good practice. What lacks in this IA process is a unifying theory which tells us how to combine the various disciplinary pieces of the IA puzzle. This implies that there are no theoretically-based minimum standards that indicate a certain degree of quality. As a consequence, IA often lacks credibility both in disciplinary science and in the policy community. Another weakness is that current IA efforts do not include a broad participation of all disciplines, especially social scientists are missing.”

(Rotmans and v. Asselt 1996)

“What kind of role should scientific consultancy play in the design and implementation of climate protection measures? On the one hand, an appropriate public discussion on climate change and climate protection measures is not possible without scientific contributions. ... On the other hand, the natural and environmental sciences involved have had to deal with a problem that is related to the subject of research itself: the complexity of and uncertainties about the climate system. ... The ‘certainty of uncertainty’ makes scientific consultancy in the climate change debate more difficult, especially when uncertainty extends from the purely technical level (How precise are simulations?) through the methodological level (What are appropriate instruments and techniques of research?) even as far as the epistemological level (In which terms and tradition of research should a study be designed?)”

(Schuele 1996)

“The success of an IEA in providing informed response strategies for complex issues with a substantial environmental component will depend on the quality of the assessment process and the adoption of good practices during its development. ... Several key aspects of good practice identified in this study are presented here – the treatment of uncertainty, the incorporation of plurality of approaches and how to address fairness and equity. The development of good practice will evolve over time and is probably the major challenge for the successful implementation of IEA in the future.”

(Bailey et. al. 1996)

Introduction

This report has been prepared as a contribution to the work of the Integrated Environmental Assessment Research Forum, sponsored by DGXII of the European Commission. It is part of a response to a need, felt with increasing strength among practitioners, for some critical perspective on this technique. Integrated Environmental Assessment is, at this point, the best scientific response that we have to the unprecedented problems of climate change. There are many indications that our industrial civilisation has disturbed the processes which govern the global climate. But none of them are quite conclusive; and even if we accept that climate change is coming, its patterns and its consequences are even harder to forecast. Yet we want to do everything we can to prevent unwelcome changes and to mitigate their likely consequences. And if significant changes do occur, they will require adjustments in our economies which could take years to prepare properly.

Planning a response to such eventualities is a challenge of a different character from those which science and technology have previously faced. During the centuries of triumphant scientific discovery and industrial innovation there seemed little need to explore possible adverse consequences of progress. The conquest of ignorance, and the associated conquest of Nature, seemed to carry with them an assurance of their beneficence. From discovery, to application, to diffusion, and then to an increase in human welfare, the path seemed straightforward. For those trained in science, technology or engineering, there seemed no need for any perspective other than that of simple progress. Those who warned about possible ill effects of this great engine of innovation, or about our irremediable ignorance of all its possible consequences, were left on the margins.

When these new problems of the global environment first manifested, about thirty years ago, it was only natural for those who were trained in that older conception of the tasks of science to attempt to solve them in that tried and true fashion. They had available several fields of applied mathematics, developed in the post–war period to cope with problems of planning and decision. For a time it appeared as if these new techniques, allied with computers of ever increasing power, would enable us to foresee the relevant aspects of our future, and then plan the best responses.

In the event, the early optimism was not sustained (Shackley 1996). Problems of “modelling” continued to resist solution, in spite of all the advances in data collection and computer methodology. The issue has come to a head over the attempt to find the value of a single quantity, which (for immediate policy purposes) could stand proxy for the future of the environment: global mean temperature some half–century in the future. Now, there is a growing consensus (though by no means one commanding unanimity among experts) that there is a significant increase. But as to its magnitude, and its global and regional consequences, as yet we are in the realm of deep uncertainty.

In recent years the awareness has grown that there is no single master key to the complex problems of ourselves in the biosphere. For their effective management, and also for scientific understanding, it is necessary for many aspects of the total picture to be comprehended at once. In the policy realm, neither automatic market mechanisms nor state command–and–control will suffice. More recent studies have exhibited a more synthetic, systemic, approach to the problems, making some attempt to include factors from the natural world, together with technology, economy, lifestyles and even values. The name "Integrated Environmental Assessment" has been assigned to a family of methods, all using computer models strongly, with a focus of attention varying between the behaviour of the biosphere and likely responses of policy or markets.

This newest development in policy–related science has had one very healthy feature from the outset: self–criticism, both private and public. The naive euphoria of the early days of global modelling is no more; now we all know that this is a difficult and demanding craft, requiring much effort and (up to now) guaranteeing little in the way of certainty. Whereas in some other mathematical sciences of human behaviour, the management of uncertainty has been neglected to an extreme degree, here it has been recognised as central to the effort. The risks of irrelevance from policy, and even vacuity of content, are well recognised as ever–present.

The articulation of these concerns began a while ago (Parson 1995), and has been gaining strength ever since. Through the period of composition of this report, I have become aware of several excellent reports, each one analysing the methodological problems of Integrated Environmental Assessment from its own perspective, achieved either last year (Parson, Toth) or within this current year (Bailey *et al.*, Kandlikar and Risbey, Morgan *et al.* Rotmans and van Asselt, Shackley, and van der Sluijs). When, in the near future, all of these are published, we will have a unique corpus of commentary on a field of practice, so that anyone interested in engaging in it will have a public source of insights on its promise and its pitfalls. It is a very welcome sign of the genuine maturing of science in this domain, that such critical reflection is encouraged and then integrated into the self–understanding of the field. Should this practice be emulated in other policy–relevant disciplines, and then perhaps even in the traditional sciences as well, the practice of natural science of any sort could become once again a genuinely humane discipline.

In the context of this very welcome ongoing activity of assessment, my own task in discussing “good practice” was made considerably easier. I am relieved of much of the burden of showing that there is a problem. I can take it as read, that Integrated Environmental Assessment is not producing the sort of reliable results that one might have hoped, and that a host of methodological problems are still awaiting solution. Then I can consider the more general problem, what would be involved in the establishment of “good practice”? And in the light of that I can offer general reflections on the sorts of criteria that would be relevant to such a task.

My message is essentially a simple one. The mathematical techniques themselves, as they have been taken over from other fields, do not contain in their formalisms either the means of criticism or the recognition of the possibility of failure. In their abstract, quantitative form, they seem to provide hard facts. Their results might not be “wrong” in the traditional scientific sense; but worse, they might be incapable of being tested as right or wrong. And they might also be irrelevant to policy or even empty of content. In this field, as in any other nascent and critical endeavour, the pitfalls are everywhere dense. There has not been sufficient history of both failure and success, for a matured craft experience to guarantee reliable work. In preparing for “good practice”, our task is to enrich those techniques with means of evaluation and criticism, and also to recognise them not as providers of hard facts, but as instruments of policy processes.

The task for the development of “good practice” will involve an assimilation of the detailed critical assessments I have mentioned, and then an articulation of criteria in the terms of some perspective along the lines of that just above. This effort will take place as part of a growing dialogue among those who devise the methods, along with all their users. There will probably never be a simple code of “good practice” for Integrated Environmental Assessment ; by the time such a state could be achieved, the approach with this name might well be obsolete! But we can imagine a continuous dialogue, leading to greater self–awareness among all those involved with these methods. Should that process bear fruit, then Integrated Environmental Assessment will rightly be judged as a success. The present report, taken in conjunction with the other studies, and the ongoing dialogues among practitioners and users, is intended as a contribution to that process. It should be seen as an exploration of certain aspects of the problem, on which the author has a special competence and concern. It is neither exhaustive, definitive, or conclusive. If it stimulates discussion and debate, it will have performed its function.

1. The Problem - Why "good practice" in Integrated Environmental Assessment?

A topic like this, which involves reflecting on a field of practice rather than doing it, can be looked at in two ways. One view is to consider it an irrelevance, or at best an excursion from the real work of the disciplined study. Such a dismissal is the common attitude among practitioners in most science-based disciplines. The other view to is to say that in the absence of such reflections, the discipline is in danger of losing contact with its supposed tasks, and eventually also in danger of losing its clientele. Then such a reflection is considered as a "added-value" activity; although it must be undertaken outside the ordinary routine of practice, it eventually brings benefits to all practitioners and to their clients as well.

In the case of Integrated Environmental Assessment, there is now a growing consensus that critical reflection is not a luxury, but an urgent necessity. We are increasingly confronted with complex problems deriving from our material interaction with the natural environment, whose solution will require a synthesis of science, technology, politics and lifestyles. Although highly abstruse scientific studies are employed in the attempted resolution of such problems, and can indeed make a real contribution to that end, they are not, and cannot be, sufficient on their own. The task, then, is to integrate such techniques with the array of other approaches. This is more easily said than done. It can seem simple only if the "integration" consists of feeding technical parameters into a purely macro–economic analysis; and even then that appearance of simplicity is very likely to be deceptive. Hence "integration" becomes a special task on its own, one for which our inherited knowledge offers only the most rudimentary guidance. The quotations from colleagues with which this report begins, are evidence that the tasks facing us in establishing "good practice" are far from trivial.

We are thus faced with a new sort of scientific problem: we want to define and establish "good practice" in this important but still immature field. But who are "we"? How should we go about this task? Who will guarantee that we have been following "good practice" in our own work? And what are the credentials of any possible assessors of our work, to say that it has been successful? All these questions show that to a considerable extent, this must be a sort of "bootstraps" operation, where we have no choice but to make the road as we travel along it. For people whose experience of science has been in the normal path of education and research, such a prospect is new and perhaps somewhat frightening. But it is also a great challenge; if we succeed at it, it can show the way to a new conception of scientific practice, one that is appropriate for its leading tasks in the times to come.

The problem we face is twofold. On the one hand, there seems to be a consensus that Integrated Environmental Assessment is, as a technical discipline, “immature”. The task is somehow to bring it towards maturity and effectiveness in performing its functions. However, there is another difficulty, somewhat more subtle, but whose significance is equally great. This is, that Integrated Environmental Assessment is a “science” only in a special sense. For as a practice it does not come under any of the traditional headings for science. It is not research of the ordinary sort, producing results intended to become knowledge; nor is it R&D, and capable of being tested fairly directly through use. Nor it producing devices intended to pay their way by satisfying some perceived need, and being tested by their success on some sort of market. It might be considered a “decision aid”; and tested by its utility in the decision process. But we know that the vast majority of decisions are taken, for better or for worse, in the absence of such assistance. In spite of that gap between intention and practice, “decision aid” seems to be the best description of this family of techniques; and accepting that, we must ask what difference is made to the governance of the practice, that it has this novel character. This question relates particularly closely to the issue of “good practice”, or, more generally quality assurance. In the case of research, that is seen to by the peer community; in R&D, by managers; but here, when the intended users are diffused among decision-making communities, it is not so clear who is to be involved in this exercise. I will show that in this case, as indeed in that of the problems of immaturity, Post-Normal Science (Funtowicz and Ravetz, 1992, 1993) can provide some insights, for the resolution of these two problems. And these problems are the key to the achievement of “good practice” in Integrated Environmental Assessment.

1.1 Decision-making under inadequate certainty.

It is a commonplace now, that policy decisions must be made under conditions of uncertainty, sometimes even of ignorance. At last we can move beyond the hitherto dominant tradition in the theory of decisions, which assumed that good decisions must have a basis in quantified information, either the facts themselves or at least their probabilities. But the point can be sharpened: it is not merely that there is some uncertainty; but rather that whatever certainty we may have, is inadequate for determining the decision or even for guaranteeing its correctness. For some, this situation is equivalent to the rule of irrationality. For others, particularly those confronted by real policy problems, it is a challenge which must be met somehow, simply because it will not go away.

There are various ways of expressing this new situation. We may summarise it in saying that when engaged on a science–related issue, decision-makers find, typically, facts uncertain, values in dispute, stakes high, and decisions urgent. Another approach is to recall the traditional opposition of the “hard”, objective facts of natural science, against the “soft”, subjective considerations of personal values in politics. How great would be the benefit if the latter could be driven out of the policy process, so that rationality (and hence fairness) could reign supreme. Now, however, we find that in many policy issues, the scientific inputs are irremediably “soft”; and the decisions to be taken are “hard” in several respects. The science is inconclusive and debatable; and the decisions involve contested interests and have grave consequences.

Formerly, theorists distinguished between “risk”, which could be quantified, and “uncertainty”, which could not. Now we have various kinds and degrees of uncertainty, and ignorance as well. And even ignorance reveals a host of varieties within itself. In particular, we have “policy-critical ignorance”, as when there are no parameters, no theories, hardly any guesses, to supply a necessary foundation for policy decisions. Such was the case in the early days of the BSE/neo-CJD crisis. And there is another sort of ignorance, which we may call “ignorance of ignorance”, which is the most insidious and dangerous sort. This occurs either when we falsely believe that we, or someone in authority, possesses the necessary knowledge; or when we are simply unaware that there is something out there getting ready to impact on us. I have argued elsewhere that this “ignorance of ignorance” is to some extent a social construction. Up to the time of Descartes, everyone who knew some philosophy was aware of the category; but with the rise of triumphalist modern science, it was suppressed and forgotten (J.R. Ravetz 1993). All this may seem at first like recondite philosophy; but as we view the increasing unpredictability of the disturbed environment around us, then “policy-critical ignorance of ignorance” may well become one of the more crucial aspects of our knowledge of the external world.

1.2 Quality assessment in its policy context.

Although such detailed considerations about ignorance are not yet broadly diffused, the public is fully aware of the occurrence of nasty surprises in the environment, and of disagreement and confusion among supposed experts about their nature and cause. They must also know that the current crop of such problems is hardly likely to be the last. So there is no point in anyone still upholding a dogma of scientific infallibility. The task is rather to do the best that one can with the tools available as decision aids, and to hope that that best is good enough. But what is “good enough”? What are the criteria, and who are the assessors? It is not appropriate to require that decisions be “correct”, for in many cases the systemic processes have such long lead and delay times, that it could be decades or more before the effects of present decisions can be discerned. In these novel circumstances, we are shorn of the assurances of quality that traditional science and technology have enjoyed. Those we adopt will necessarily be looser, less secure, and played out among different communities. The sort of dialogue which has previously prevailed among scientists and their audiences cannot continue unchanged in these new circumstances.

Certainly, the era of “academic science”, when researchers could commune with each other and effectively ignore the outside world, is long since gone. Now the criteria of value of projects to be chosen, and increasingly the criteria of adequacy of results to be accepted, depend on agencies outside the research process itself. Either planners of science in a State bureaucracy, or private sponsors and contractors of research, have an increasing, sometimes decisive influence on the process of assessing quality and also establishing its framework. Such extended peer communities have a sort of legitimacy. If they are not actually competent in the work itself, then they are at least advised by some who are. And in their own work of governance, they are answerable to others who can (in principle) apply sanctions if they fail. But now there are more widely extended audiences, whose influence can be crucial, and whose legitimacy derives not from their status but their concern. Sometimes they affect the research process directly, as in the case of sufferers from particular diseases, notably AIDS. But they can also act in a diffuse and perhaps uncontrollable manner, as when masses of consumers suddenly exercise their functions as citizens through boycotting a particular suspect product. This might be Shell petroleum, in a notorious case of political agitation; or it might be British meat, as in the scandal of BSE/neo–CJD.

It is now accepted that the solution of climate change problems lies in the modification of our technology; and much of this will depend on an alteration of lifestyles. Taming the motor car is the most obvious example. But to induce people to make a real, present sacrifice (however small) on behalf of some uncertain benefit for future generations will require persuasion of a new sort. It is in such propaganda that Integrated Environmental Assessments are invoked. It can be tempting to pretend that their conclusions are hard scientific facts like Newton’s Laws of Motion; but such a pretence would be punctured very quickly. And with it would go the credibility of those who are advancing the argument. There have been enough well–known cases of predictions gone wrong, both official reassurances of safety, and prophecies of imminent ecological doom, to make environmental scientists cautious in their assertions.

The task of environmental scientists, then, is somehow to convince a public, whose trust or even whose interest cannot be assumed at all times, that on the basis of confessedly uncertain arguments, they should accept policies that bring a known loss for the sake of a speculative benefit. This would appear to be an uphill struggle, and so it is. In the absence of scientific rigor, what source of strength is there for the scientists' arguments? It is only in their integrity; they are judged as witnesses giving testimony, rather than as oracles delivering truth. The evidence for their integrity is their transparency. Paradoxically, the open admission of uncertainty, provided it is supported by a display of real competence, strengthens their case more than any claim of certainty.

Good practice, being done and being seen to be done, is then the bulwark of the environmental scientists' position. Only in that way can scientists provide real assurance to the public that their management of uncertainties is the best possible. And when they are confronted by critics who, for whatever reason, deny their recommendations, it will be on the strength of their testimony that their conclusions will prevail.

1.3 Quality assurance: products or processes?

Quite suddenly the issue of Quality has become big business. What had previously been a minor function in most manufacturing processes, and something taken for granted in almost all others, has become a central feature of our economies. Doubtless the painful lesson of Japanese competition was important in convincing people that quality counts. With this came the discovery that in many spheres, informal self–regulation of practice falls short, to a lesser or greater degree, from what is necessary for the assurance of quality to a proper standard. So a variety of bureaucracies and initiatives have come into being, for implementing quality assurance on the broadest possible scale, extending into all aspects of the functioning of all organisations.

This effort displays many of the features of systems behaviour, in a conveniently obvious way. First, there are its recursive properties, recalled by the classical motto "Quis custodet custodies ipsos?" ("Who guards the guardians?"). Also, the complementarity of what is formally specified and what is informally understood is present in all its practices. Too much specification, and regulations become ridiculous and counterproductive; too little, and they become easily evaded and lax. Another complementarity, the choice between product and process as the focus of concern in quality assurance, brings us to our present concerns.

Where there is a tangible product, made in accordance with explicit specifications, then the testing of its quality is convenient and effective. But otherwise, tests on the product may become so indirect and so dependent on proxy indicators, that the inherently more difficult path of testing, of the process, may actually be more practical. This is the case in scientific research, where the products are results whose best testing resides in their eventual use in subsequent research. Scientific research has another important special feature, in that only immediate or close colleagues of the researcher are fully competent to test the process. For the craft work of research becomes very specialised around each particular sort of problem and its associated techniques. Referees from more remote areas of research can provide only a general assessment of quality.

This property, of effective, largely informal self–assessment, is a distinguishing, perhaps unique feature of research science. It both depends on, and also fosters, a special sort of ethics and morality. For it is extremely vulnerable to all the sorts of abuse that are commonplace, sometimes even dominant, in many other spheres of social activity. The fragility of science as a social enterprise is attested by the relative rarity of the geographical and cultural contexts in which scientific excellence can flourish. In matured scientific fields, quality assessment is learned by students through a very traditional means of precept and example; it is a sort of knowledge that can be only partly specified. But in less matured fields, the creation of appropriate criteria of quality, and then their diffusion and enforcement, requires special efforts. Indeed, the striving for maturity, which is equivalent to reliable quality of work, has dominated many of the fields of intellectual enquiry for generations; and it has also inspired significant philosophical exploration.

In the case of Integrated Environmental Assessment, quality assurance of the products can go only part of the way. Since they do not predict in any great detail, they cannot be tested in the ordinary way. One can, of course, compare models in “benchmark” exercises, to see which agree among themselves or with some externally known phenomenon. But in general, the quality of the model is assured only by the quality of its production. And here, as so many are keenly aware, that quality, as assessed through “good practice”, is in a far from satisfactory state. How is “good practice” to be assured?

2. Traditional approaches to the "good practice" problem.

We must begin somewhere, but where? Let us review what has previously been tried, for the achievement of good practice in similar situations. Integrated Environmental Assessment is not the first decision aid based largely on mathematical methods. In fact, its pedigree goes back many centuries, or even millennia. As we infer from prehistoric monuments, "applied astronomy" has long functioned as a way of marking the divisions of the year for production and for ritual. More recently it was elaborated into a sophisticated science of prediction and counsel, both for individuals and for empires. We no longer believe in the cosmos which made astrology plausible, although from outside our particular culture the model of man as "marginal–utility homo economicus" might appear equally implausible. Once astrology was discarded, the exemplar science was Newtonian astronomy. Its combination of precise observation, abstruse mathematics, and astonishing prediction made it appear nearly divine in its certainties. For some centuries now, various behavioural sciences seeking maturity have attempted to emulate, sometimes even to imitate, the mathematical sciences; if we can predict the motions of the faraway planets, then why not those of human society right here? Perhaps now, as we confront the problems that call for Integrated Environmental Assessment, we are at last at the end of that particular road of mathematical fantasy.

2.1 The Quantitative Path and its Pitfalls.

It is no accident that the pedigree of the mathematical decision aids extends back continuously through astrology to prehistoric astronomy. For in all cultures, and especially in ours, mathematics has had a fascinating effect. Somehow its union of the general (numbers, sizes, shapes) with the particular (as counted measured or perceived), has seemed to hold the key to truth and wisdom. From Pythagoras onwards, a great strand in European intellectual culture has sought for certainty and even meaning in the mathematical patterns of the phenomena of experience. More recently, in early modern times, there was added a faith in mathematics as the embodiment of objective, impersonal certainty. The combination of these two (somewhat contradictory) sorts of expectation has provided mathematics with an overwhelming plausibility. Every time a new applicable mathematical technique is created, the burden of proof is on any critics who might query its validity or effectiveness. We should therefore not be surprised at the enthusiasm with which one mathematical gadget after another (theory of games, catastrophe, chaos, fractals) is received, in spite of the recent memory of its disappointing predecessors.

The programme of uncritical application of quantities to qualitative phenomena could be successful only so long as the systems being considered were relatively simple. The subject matters of, say, physics and chemistry could for a long time be considered in isolation from their context, natural and social. Indeed, great power could be created thereby, and the human aspects of its use and control could be left to some “hidden hand”, either of the market or of some diffused societal benevolence. But in this century we have discovered that there is an ethical responsibility involved even in the most fundamental physical science; and now we are learning that some natural systems are too rich to be captured by numbers alone.

This latter lesson may be strange, and indeed unwelcome to many. Many theorists in decision–making for economic and political affairs tend to believe that as quantification weakens, rationality is lost. Their model for true knowledge and correct decisions is derived from the sciences which in some languages are called “exact”: those using quantitative data and formal arguments. It will doubtless seem very paradoxical to be told that the strength of these “exact” sciences derives from their control of *inexactness*. For every measurement involves inexactness; and every computation does so as well. The “natural numbers” are well designed for counting and calculating, but for representing measurements, they are imperfect and misleading. Those who do sophisticated calculations, as on computers, know full well how much judgement, skill and artifice are required for obtaining reliable results (Davis and Hersh 1986). But this knowledge is not widely diffused, as it is frequently in the common (short-term) interest of both providers and purchasers to hope that any given set of calculations is as secure as the addition of 2 and 2.

If we are to achieve good practice in Integrated Environmental Assessment, we must sooner or later face the fact that our common accepted skills in the manipulation and expression of quantities, particularly in the representation of inexactness, are defective. Since this aspect of quantification has received so little attention in our reflections on science, it should not be surprising that it should have been neglected in teaching and practice. In particular fields of research, practitioners have developed craft skills in this area; but these are usually left implicit, passed on during training, and not appreciated as crucial for the practice. Then when other disciplines seek to achieve that same state of maturity, or objectivity, through quantification, they copy the outward form of exactness while remaining ignorant of its true content. There cannot be many practitioners of these would–be quantitative disciplines who have ever heard of the maxim of C.F. Gauss, one of the greatest of mathematicians: “Lack of mathematical culture is revealed nowhere so conspicuously as in meaningless precision in numerical calculation” (J.R. Ravetz 1995, p. 158).

This principle can be illustrated by a simple joke about schoolchildren's use of numbers. We can imagine a child being asked how long ago the dinosaurs were on the earth, and responding "sixty–five million and three years ago". Asked how he got this number, he replied that it was three years since he saw "Jurassic Park", where he learned of the sixty–five million. Like a good pupil, he had done the sum: 65,000,000 + 3 = 65,000,003. That may seem ridiculous; but should he have been told that the correct sum is: 65,000,000 + 3 = 65,000,000? If he wrote that in a test, surely it would have been marked wrong by the teacher. Of course there are ways around the paradox; we might use "65E+6" for the larger number, and develop rules for neglecting much smaller terms in an addition. But then we are abandoning the calculation with integer numbers, which is something that "every schoolboy knows", and no schoolboy is warned against.

With Silvio Funtowicz I have discussed such problems in our book, *Uncertainty and Quality in Science for Policy* (Funtowicz and Ravetz 1990). We argue the point at length that blunders and misinterpretations in numerical arguments are bound to persist until two things are accomplished: first, that there is an awareness of the existence of a problem; and second, that tools are deployed to solve it. We make a contribution towards the latter, with our "NUSAP" system of notation; and we hope that greater understanding of the methodological problems of Integrated Environmental Assessment will contribute to the necessary background awareness.

2.2 Uncertainties in Environmental Economics: sophistication or sophistry?

In the meantime, however, it is possible for sophisticated practitioners to perform sophistical manipulations with uncertainties, and to influence policy debates with numbers whose precise appearance belies their uncertain foundations. I shall analyse an example of this practice at considerable length, partly because of the importance of the result (a key policy indicator valued at 2%) in policy debates, but also because in this case the reasoning is so conveniently transparent. The paper, by W.D. Nordhaus (1991) is on the economics of the greenhouse effect; it calculates the damage resulting from a modest degree of global warming. (Silvio Funtowicz and I dealt with this, as an important example, in our paper "The Worth of a Songbird" (Funtowicz and Ravetz 1994b)). Unlike many others who fling numbers about in such debates, this author is clear and ostensibly quite scrupulous in his management of uncertainty. Thus the paper is liberally sprinkled with caveats, including such choice lines as, "We now move from the *terra infirma* of climate change to the *terra incognita* of the social and economic impacts of climate change" (p. 930). He devotes the last two of the five points of the conclusion of his paper to a discussion of the various severe uncertainties in his analysis. The recommendation he makes for his model is carefully phrased around its weaknesses: "Notwithstanding these simplifications, the approach laid out here may help to clarify the questions and help identify the scientific, economic and policy issues that must underpin any rational decision." There is nothing here about quantitative prediction, or even of policy entailments; "to clarify the questions" and "to identify the issues" (p. 937) are surely suitably modest goals.

Further, the author displays a commendable awareness of the different sorts of uncertainties that are involved in any quantitative statement. Most estimates and model outputs in Nordhaus's paper are given as a set of three numbers, or as a central estimate with a $\pm$ for Spread; thus the precision of his quantities is well expressed. The accuracy of his estimates is described in various ways; in one crucial place it is conveyed by the descriptive term "an informed hunch" (p. 936). (This latter attribute qualifies his magnification of the leading estimate (flow of damages to the U.S. economy) from $^1/_4$ % of global output, to a possible 2%). The deeper uncertainties are listed in a paragraph of "important oversimplifications", concluding with a reminder that the analysis "ignores the issues of uncertainty, in which risk aversion and the possibility of learning may modify the stringency and timing of control strategies" (pp. 936-7). This displays the "border with ignorance" of the analysis quite clearly and explicitly. One could hardly ask for more in the way of sophisticated awareness of uncertainty and of the need for its expression and management.

But when we come to scrutinise the actual quantitative reasoning, a different impression emerges. We will not discuss his use of higher mathematics for the analysis of the functional relationships among very inexact quantities, for in the absence of that practice how many disciplines would survive? But we can observe a marked lack of practical skill, or prudence, in the handling of quantitative information. His crucial Table 6 (reproduced in Table 1 below) has eleven entries, of three different sorts. Five entries have numerical values (in billions of dollars) around -1 or -2, expressed to three significant digits. One entry (impact on farms) is a large interval, [-10.6 to +9.7]. And five entries are unquantified, with indicators such as "?" or "small". The sum of these eleven entries is given as -6.23, which seems to be the sum of the five precise terms, added to the average value for the "farms" entry (-.45), although that addition done correctly would be -8.55. The percentage of national income is then calculated to two significant digits, as -0.26%. If the variability in the dominant term were included in the calculation, this result, the one relevant to policy, would be the interval [-0.4% to +0.4%]. In his discussion, the author says that he "might raise" the -$^1/_4$% to -1%, by an adjustment that is "purely ad hoc"; and he then gives his "hunch" that it is less than -2%. (He tactfully ignores the possibility that his calculation yields an equally likely *benefit* to the economy). To raise the percentage term to the 2% limit would require the sum of the non-quantified terms to be greater than the sum of the quantified and averaged terms by a large factor, around 3. All the precision in the quantified terms is then completely lost, first in the huge uncertainty interval of the leading term, and then in the uncertainties in his adjustments factors, where the "hunch" terms are even bigger than the "ad hoc" ones.

What then is the point of all the calculations with those entries? The hyper-precision in the expression of the key number -0.26%, which could equally well be anywhere in the range $\pm{}^{1}/_{4}$%, shows that the final "2%" is one of those "magic numbers" designed to produce confidence in the existence of a hard core of objective fact somewhere deep inside the mass of intuitive fuzz. By the time that the author has admitted the manifold oversimplifications and uncertainties in his analysis, and has shown how strong are the *ad hoc* adjustments and hunches which are needed to bring his numbers back into the realm of plausibility, we might ask whether the statistical exercises are totally redundant except for rhetorical purposes.

I cite this example for several reasons. First, the author is a recognised scholar who has made important contributions to the climate change debate; second, that he shows a keen awareness of the problems of managing uncertainties; and third, that in spite of this, he has encountered some pitfalls which on retrospective analysis seem brutally obvious. How much more severe would be the criticisms of arguments in which the uncertainties are not even discussed, and where conclusions are presented in bare numbers as if engraved on stone? We should recall that in any "exact" natural science, a quantity presented without at least an error–bar is not considered as meaningful. Would–be quantitative social sciences that ignore uncertainties are really in the state of imitating the precision of numerical outputs, while remaining ignorant of their foundation in the controlled accuracy of quantitative inputs. This is a recipe for vacuity.

2.3 Computers: the perils of GIGO.

A new variant on the mathematical approach is through computers. They can accomplish complicated calculations on vast arrays of data, and so offer a real hope that the complexities of environmental systems can be captured, in part at least, through "models" in which their quantitative relationships are simulated. During their early development and diffusion, computers seemed to be the material embodiment of all the perfection of mathematics. Operating "faster than thought", and supposedly error-free on account of the simplicity of their basic operations, they seemed to realise all the virtues of the mathematical way, and that in a practical, powerful form.

The effectiveness of computers in a vast variety of tasks is of course not to be disputed; but our present concern is with decision aids. And here it has been found that all their speed and accuracy cannot fully compensate for deficiencies on the empirical side of the process. If the input data are inadequate or uncertain, then only so much can be done by way of remedy. And it turns out to be all too easy for such low quality to be overlooked in a calculation. Computers can actually become instruments for perpetuating our ignorance of ignorance, by fostering the illusion of knowledge through their highly technical operations and highly precise outputs. Indeed, a special acronym has been coined to describe this phenomenon. It is “GIGO”, from the American “(if you put) Garbage In, (then you’ll only get) Garbage Out”.

The pitfalls of computers can be characterised in terms of GIGO. We may define a “GIGO–science” as one where the uncertainties in inputs must be suppressed, lest the outputs become indeterminate. And a clue to the presence of GIGO is that precision of numerical outputs goes up, as accuracy of quantitative inputs goes down. Where do we find GIGO? This can be an embarrassing question to answer; indeed, some might say that the real question is, where do we *not* find GIGO? Our inherited ideas of mathematics and Scientific Method make it very difficult for us to conceive of GIGO as a widespread or significant phenomenon. But if there were any doubt on this score, one would only need to study the history of the American “Strategic Defence Initiative”, otherwise known as “Star Wars”. In this effort, some billions of dollars per year were devoted to the creation of a radically new technology, involving performance standards that were impossible, socially, physically, perhaps even logically. Until it was finally wound down (although a change in Administration could yet lead to its resuscitation), it survived on a combination of murky sectarian politics, imaginative TV computer-reconstruction of intended action, and old fashioned boondoggling. It did have some real results. It is said that some inside the Soviet government took it seriously enough to realise that they could not win another arms race, and so gave Gorbachev a chance. Also, the scandal was so egregious that computer professionals actually organised themselves and went public. It was only then that the myth of “unerring” computers was shattered, and the importance of programming, as opposed to binary arithmetic, was revealed. In that connection, the maximum size of error-free first-draft software was first publicised. This was some ten lines of code; “Star Wars” required code of the order of millions of lines, for complex interactive systems which would need to function faultlessly on its first real trial. So in a nearly post-modern fashion, “Star Wars” contributed to human understanding by showing that GIGO exists, but also that it can pay, even paying very well, at least for a while (J.R. Ravetz 1990a).

The secret of the eventual exposure and demise of “Star Wars” was debate and publicity, starting with revelations of the highly dubious character of the experimental data on which all its plausibility rested (X-rays emitted in an H-bomb explosion). Thus the precious insight of Karl Popper about the life of science, that it depends on criticism rather than on a truth-generating Method, is confirmed; without debate comes stagnation and corruption, in science and technology as much as anywhere else. In that context, we can be confident that Integrated Environmental Assessment will be protected from GIGO. There is healthy debate and critical reflection among the community of those who work on the models used in Integrated Environmental Assessment. For there is ample evidence of their concern with the management and communication of uncertainties, and with good craftsmanship in general. Whatever other problems may be present, at least we are protected from vacuity in this field.

The field of modelling is maturing rapidly. In its earlier years it went through euphoria and crisis, as methods which seemed assured of success led to controversy and failure. Although the propaganda message of *Limits to Growth* was not immediately impaired by the many scientific criticisms it attracted (Cole 1978), its value as a policy tool became more and more compromised. At the end of that decade of the 1970's another model, even more ambitious in its way, suffered even worse embarrassment; this was the IIASA energy model, which was discovered to have deep flaws in its very structure (Keepin and Wynne 1984), and which was therefore incapable of supporting the policy conclusions about energy sources that had been confidently drawn from it.

However, there has still persisted a sense that models are not merely the best available means of peering into our environmental future; but that they really should be able to provide some sort of positive information that is useful for policy. In this respect they could be performing a function that I have elsewhere called a "folk–science"(J.R. Ravetz 1995), or even "elite folk–science" (J.R. Ravetz 1994). This function is the provision of basic assurance and some practical guidance, without too strong a requirement on effectiveness. This interpretation can be argued for the case of Global Climate Models, particularly when allied with neo-classical economics, along the following lines. If our technical–industrial system is indeed perturbing the environment, perhaps with serious consequences, and we are now discovering our deep ignorance of its trajectory, then we are in a state of some impotence in the face of what we have wrought. In that case our marvellous natural science will have turned out to be an essential part of the problem, and is therefore less likely to provide an instant solution. Also, in relation to our own habitat, we will have been rather less like masters and conquerors, and rather more like sorcerers' apprentices. This disturbing perspective is not necessarily true, by any means; but then it is not necessarily false either; and neither verdict is capable of immediate proof. To shield us from its consequences for our assessment of ourselves and the project of our civilisation, we apply Science yet again, this time in the development of environmental/economic models from which correct policies can be deduced. Even if the current generation of models do not predict so well, and theoretical considerations weigh against their ever doing so, we can at least say that we are doing our best, and call for a redoubling of our efforts (and resources) as the goal recedes.

What, then, are the genuine practical functions of climate models? The public might be surprised to learn how few of the leading practitioners believe that they can predict. Indeed, some of them make pronouncements that may seem paradoxical or shocking: thus Morgan and Henrion say “Every model is definitely false” (Morgan and Henrion 1990, in van der Sluijs 1996, Introduction). Or, expressed in a more practical way: “The key point to remember is that without thorough and systematic modelling and analysis of the uncertainties of the problem, we can not be sure that the results of a model, especially a very large and complex one, mean anything at all.” (*ibid*., Preface).

Given the high costs (of every sort) involved in such an effort, and its consequent customary lack of complete implementation, we are left with the uneasy feeling that by such a criterion we have a very severe quality–assurance problem indeed.

As to what usefulness, if any, is possessed by climate change models as currently used, we have the discussion in van der Sluijs’s report. There we find controversy on a number of fronts, but a consensus that models are not “truth machines”. Much of their use seems to be internal, as in “revealing our assumptions” (Toth 1995), for “furthering of a science–policy dialogue” (ref: Hellström 1996), or “an interpretative and instructive value” (ref: Rotmans and de Vries, 1996). Their uses (and abuses) as rhetorical devices, in dialogues, debates or adversarial contests, has been tabulated as well (ref: Mermet and Hordijk 1989). (All the above quotes are from van der Sluijs 1996, Section 5).

All this would make for very depressing reading, if our only source of insight on the future environment were the models themselves. But we shall see that a more comprehensive integration, in which the models themselves are but one component, can offer them a genuinely scientific role. “Good practice” will then refer to a broader set of activities, so that useful tools can be constructed, whose strength is greater than that of the sum of their parts.

Table 6.
Impact estimates for different sectors, for doubling of CO_2, U.S. (positive number indicates gain, negative number loss).

Sectors	Billions (1981 $)
Severely impacted sectors	
Farms	
Impacts of greenhouse warming and CO_2 fertilisation	–10.6 to +9.7
Forestry, fisheries, other	Small + or –
Moderately impacted sectors	
Construction	+
Water transportation	?
Energy and utilities	
Energy (electric, gas, oil)	
Energy demand	–1.65
Non–electric space heating	1.16
Water and sanitary	–?
Real Estate	
Land–rent component	
Estimate of damage from sea–level rise	
Loss of land	–1.55
Protection of sheltered areas	–0.90
Protection of open coasts	–2.84
Hotels, lodging, recreation	?
Total	
Central estimate	
Billions, 1981 level of national income	–6.23
Percentage of national income	–0.26

Sources for Table 6: Underlying data on impacts are summarised in EPA (1988). Translation into national–income accounts by author. Details are available on request.

W.D. Nordhaus (1991) p. 932.

Table 1.

3. Irremediable Uncertainty and its Implications

Although the management of uncertainty within computer models is not sufficient for the full assurance of quality in Integrated Environmental Assessments, it is quite necessary. Indeed, that is already the main focus of skilled craft work in the field, and even when the practice of Integrated Environmental Assessment is broadened, the management of quantitative uncertainty will continue to be central. Unless there is quality assurance at the quantitative end of the decision process, then at the qualitative end there is no protection against all sorts of subjectivity and caprice.

3.1 A Cautionary Tale.

I have already indicated the sorts of things that can go wrong in the manipulation of uncertainty, even in quite simple calculations concerning the environment. An equally instructive example is that of the IIASA Energy Systems Program, which was criticised sharply in 1983 by Bill Keepin. The classic account of this salutary event, which deserves to be more widely remembered within the modelling community, was produced by Brian Wynne (1984). In brief, Keepin discovered that in various crucial respects the computer model was vacuous: in one phase, outputs were linear multiples of inputs; the solution space was hypersensitive to uncertainties; the key policy conclusion (favouring breeder reactors) was extremely brittle towards assumptions about future comparative prices of fuels; and the claimed iteration of different sub-models had not been done because it could not be done.

The treatment of Keepin's discovery of these features was conditioned by his status as a young outsider. As Wynne describes it, "The gist of these rejoinders was that Keepin did not understand the 'craft' nature of systems analysis (as distinct from 'scientific') [I blush - JRR], and that he had naively analysed the models according to scientific standards. Furthermore, it was argued, Keepin had naively exaggerated the centrality of the models to the scenario-construction exercise. It was asserted that 'everyone competent already knew' what Keepin had pedantically 'revealed'. No significant technical rejoinders have yet been made to Keepin's analysis." (p.279)

Wynne later explains the logic of this defence. "Keepin's 'misdemeanour' was essentially that he rendered explicit (and thus made available to a wider audience) the implicit 'half-knowledge' of the IIASA team; i.e., the uncomfortably messy nature of the real 'analytical' process. .... [H]e had attacked a ‘straw man’, created in the public rhetoric, that did not exist in the knowledge of the mature scientists. Had he been a mature scientist and not a kindergarten one, he would not have taken this public image so seriously and wasted any energy over it. ... Yet the public rhetoric was not of Keepin’s creation, it was created by the modellers. His innocence was not of modelling *per se*, but of the tacit conventions of cryptic discourse that not only maintain credibility with and exclude external audiences, but also thereby sustain a particular process of policy analysis and policy itself. ...” (p.315)

Brian Wynne was definitely not accusing the IIASA team of a premeditated Machiavellian plot to sell their defective policy goods to an unsuspecting public. His explanations rest on an analysis of the implicit assumptions behind the team’s defence; and it is possible that had there been no Keepin incident these assumptions would have remained as safely hidden from themselves as from the public. And certainly, since then anyone aware of that incident would ensure that his or her model should never be vulnerable to such destructive criticism by an untrained outsider. Yet the IIASA incident, painful as it is, serves as a sort of laboratory experiment, a pure case, for the real dilemmas that confront scientists whose models turn out to be incapable of meeting expectations, either of the clients or themselves.

3.2 Coping with uncertainty.

Jeroen van der Sluijs (1996, Section 4) provides a deep and comprehensive survey of the sources of uncertainty in Integrated Assessment Models, and how they might be managed. Going through the whole cause–effect chain which provides linkages for the models, he shows how each step is fraught with uncertainties. For the climate models themselves, he reviews the following aspects: incomplete understanding of the modelled system; chaotic behaviour; multiple equilibria/non–smooth behaviour; and linkages with other anthropogenic environmental changes. The associated uncertainties in these aspects can at least be sketched, although not quantified. But when we consider the downstream effects of possible climate changes, the uncertainties begin to increase beyond control. For the models depend on a variety of assumptions, necessarily cast in quantitative form, for which little justification is available. Thus we have assumptions about: demands for goods and services; demographic trends; choice of technologies; fluxes of materials into the environment (e.g. waste gases from energy cycles); vulnerability of different societies; consequences; and “valued environmental components”, that is indicators that are seen as salient. Those models which incorporate economic considerations will depend on the concepts of general equilibrium theory, in spite of the empirical evidence for far–from–equilibrium states of real economies, and the theoretical possibility of multiple equilibria. Valuations depend on the reduction to a single *numeraire*, usually money in some currency. The non–market values are empirically defined through “willingness to pay” studies. But these raise problems in equity, both practical and theoretical. For a measure relating to ability to pay has been accused of discriminating against the world’s poor; but a measure based on “willingness to accept” harm (at a price) can produce values so high as to be effectively non–negotiable. Finally, all the measures depend on the assumption of a “social discount rate”, which over long time–spans has paradoxical properties, leading either to an extremely rigorous policy of conservation and caution, or to an apparent willingness to sacrifice our grandchildren’s welfare to our present pleasures.

In the face of such uncertainties, which might seem overwhelming to an outsider, what are the practitioners to do? A response at the semi-official level has been provided by the IPCC: “... it is generally believed that it is only through such models that we can gain a scientific understanding (and hence a reliable predictive capability) of climate and climate change" and "This faith in the fundamental soundness of the modelling approach does not deny the presence of significant errors in current models nor the utility of models known to be incomplete, but does provide confidence that these errors can and will be reduced through continuing modelling research.” (Gates *et al.*., 1996) The sentence bears close examination; what is it that “does provide confidence”? Is it the "faith" in their soundness, or the “belief” in their uniqueness? In previous times, scientists criticised those who sought such foundations for confidence, which were characteristic of religion rather than of science.

My own critical comments are quoted by van der Sluijs with general approval, so they are worth reproducing here.

“...the lack of quality control and good scientific practice might be a result of three sorts of causes:

1. Some people have effective “good practice” but their standards are not diffused among all practitioners. This is quite common in science, where a ‘leading’ lab can get genuine results that few others can emulate.

2. The discussion of ‘good practice’ is of the sort I call ‘lamp-posting’, from the old story about a man who was seen by a neighbour in the early hours of the morning, crawling on the ground near the lamp-post. Asked what he was doing, he replied that he was looking for his keys. ‘Did you drop them there by the lamp-post?’, the neighbour asked. ‘No, it was near my front door’. ‘Then why are you looking at the lamp-post?’ ‘Because at least it is light here, so if they were here I would find them.’ (I saw this recently as a joke about a drunken man; but it is also quoted as an original Sufi story!) Translated into practical terms, this means that the researchers concentrate on the soluble problems, even if the insoluble ones are more important. (Silvio

Funtowicz has enriched the ‘lamp-posting’ story with the comment, ‘and the light is fading anyway’.

3. Finally, there is the possibility that discussions of quality are only a game. This might be played for political advantage within the field (who can demolish the other’s research more effectively?), or to comply with external requirements.”

van der Sluijs comments, “To Ravetz number 1 we can add the cost–problem. We saw that despite the availability of adequate tools for uncertainty assessment such as Monte Carlo Simulation, these are yet hardly being used in IAMs, due to their resource consuming character. We found support for Ravetz number 2 in the evaluation report of the Dutch NRP, where it was observed that the program focused on strengthening of existing areas of excellence in Dutch climate research, rather than on the key questions that should be answered ... . Ravetz number 3 is supported by the ‘backlash phenomenon’ which is most prominent in the US, where industry organisations such as the Western Fuels Association grant research to actively undermine the scientific credibility of the IPCC, while on the other hand the IPCC does its best to enhance their credibility by maximising representativeness and maximising the process legitimacy of the consensus building process ... .” (van der Sluijs 1996, Section 6.4)

3.3 Implications of uncertainty for “good practice”.

It is both fair and prudent to start with the assumption that the scientists involved are wanting to do a good job; if they may seem inconsistent or less than fully aware on problems of uncertainties, this is because those problems are strictly insoluble in terms of the practice that they command. Worse, if the problems of uncertainty are so severe as to compromise the possible achievement of the stated goals of the practice, then the utility of the whole exercise is called into question. The definition of good practice will then depend wholly on the internal quality of the work, its craftsmanship, and will be deprived of that other dimension, namely its performance of its promised functions.

These functions are a varied lot; in the early days of modelling, it was believed that they could predict, sufficiently well to inform policy. Now there is a general retreat; and the claims made are correspondingly more modest. Some, including researchers, are quite definite about the severe limitations of Integrated Environmental Assessments for policy purposes. van der Sluijs reminds us that in the last resort any model on a global scale is calibrated to the results of the “General Circulation Models”. Concerning these, the verdict of Ann Henderson–Sellers is severe:

“Today’s climate models are essentially useless for virtually all forms of policy advice related to climate change. They are useful for some forms of short–term forecasting and medium range climate advice (e.g. El Nino projections, ...) but for long–term advice related to the enhanced greenhouse effect the value is minimal at best. The key conclusions of the models are driven by the assumptions and the various structures and devices used to simplify the calculations to make the models compatible with today’s technology. This massive problem is an important feature of the difficulty in linking the science and the policy.” (Henderson–Sellers 1996, in van der Sluijs 1996, Section 4.4).

Others state other interesting functions for the models, but none of them carry the same conviction. For a leading example, we have:

“First, integrated assessment can help (indeed is necessary) to answer the broadest bounding question, how important is climate change. Second, IA can help assess potential responses to climate change, either with a benefit–cost framing that compares costs of responses to the impacts they prevent, or with a cost–effectiveness framing that assesses relative effectiveness and cost of different response measures to meet a specified target. Third, IA can provide a framework in which to structure present knowledge, providing several benefits. Perhaps the most important contribution is structuring of uncertainty and sensitivity: how well quantities and relationships are known, and how strongly valued outputs depend on them. Finally, integrated assessment can serve the longer–term goal of capacity building.” (Parson, 1994, in van der Sluijs 1996, Section.5).

And then all such assignments of functions encounter the difficulties in their successful realisation, and worse. For, as Risbey *et al.* (1996) observe, in its present state modelling presents a “Catch-22”. “On the one hand, the evaluation of forecasts is difficult to do for IA models, and reduces to an evaluation of insights. Yet, as we noted earlier, the generation of insights from IA models depends to some degree on their ability to provide reliable quantitative information." (Section 6.1). It seems, then, that there is no secure middle ground for Integrated Environmental Assessment lying between the patently false claims of power for prediction and policy, and the true claims that are so modest as to be platitudinous.

Integrated Environmental Assessment models are not alone in being in the condition of a mathematical tool still looking for a function. I have already mentioned the rise and fall of mathematical gadgets, some quite vacuous in practical terms, which commanded great prestige and popularity in their time. What may make the difference for Integrated Environmental Assessment, is that the ultimate function is still there, in our coming to terms with our place in the global environment. The task is to develop the instruments so that they are more appropriate to the possible functions they can perform. Honest craftsmanship is essential for this process, but it also involves another, challenging dimension. As I shall discuss soon, many practitioners recognise the "immaturity" of Integrated Environmental Assessment, as witnessed by the insuperable difficulties of controlling uncertainties. Bringing the field to a state of "maturity" will require the exercise of both imagination and strength. The ideas of maturity inherited from the prestigious natural sciences will need adapting if they are to be appropriate to these new tasks. Maturity in Integrated Environmental Assessment will not be achieved by making uncertainty appear negligible, and values irrelevant.

In the following chapters I will sketch ideas of a positive way forward, in which the self-conscious interaction of practitioners, clients and affected publics with each other and with their tools, can eventually achieve a proper management of uncertainty and incorporation of values. It can also achieve a new conception of good practice in the social activity of a maturing science. In this I will build on the accomplishments the recent set of impressive and important studies of the methodology of Integrated Environmental Assessment that I mentioned above. These do not “debunk” the practice in a post-modern fashion, but they provide detailed criticisms and prescriptions that point the way forward to the improvement of the work. Their work is the foundation of good practice in Integrated Environmental Assessment in the future.

4. Escaping from “scientific immaturity” - examples from history and design.

4.1. "Immaturity"

There seems to be a strong consensus among commentators on Integrated Environmental Assessment, that the field is in some sense “immature” (Rotmans and van Asselt 1996). It would be very surprising if it were otherwise, given the scant historical experience its practitioners have had to go on. The problems of immaturity can be severe; inside the academic sector colleagues in such “Mickey Mouse” fields are treated with a mixture of amusement and contempt by those from more established subjects. The whole research enterprise can be dominated, or distorted, by the quest for the chalice of Maturity. It is only rarely that a leader in such a field can admit that the immature state is likely to persist, and then articulate a strategy for coping with it (J.R. Ravetz 1995, pp. 378–9).

In the new policy–related fields, we are least spared the miseries of what has been called “physics–envy” among academics afflicted with disciplinary immaturity. Some practitioners believe that their fields have Arrived, thanks to a mixture of numerical data and algebraic arguments; but others view their pretensions as a cover for “GIGO”. Now that we can confront the immaturity of Integrated Environmental Assessment without becoming neurotic about it, we can work through the problems of developing “good practice” without needing the security of an imminent accession to maturity.

First, we might ask, what would be involved in Integrated Environmental Assessment being “mature”? If our paradigm is of traditional natural science, that would involve the whole complex system of the biosphere being in principle reducible to equations, so that global or local future states could be reliably calculated from past history and present activities. The negative lessons of “chaos theory” should have finished off that particular fantasy. Even without the reflexive element deriving from human activity, a system as complex as any ecosystem will have a sufficiency of feedback and nonlinearity, so that small disturbances or changes in initial conditions (to say nothing of errors in models) can produce large and unpredictable changes in outputs.

But if the ideal of maturity, or of good practice, is not to come solely from traditional natural science, then what should be its source? We have available a variety of approaches, some of them from the world of practice, as engineering and design, and others from within the world of scholarship, as history. Let us review these in turn, to see what they can offer.

4.2 Engineering and Design.

We can start our consideration of engineering and design, by imagining two models of how “science” is deployed in practice. The traditional one starts with a scientific field, defined by its own sorts of experiments and theories, which is then realised to have similarities to some field of technological practice. Processes that have been tested in the lab are then reproduced under field or industrial conditions (with appropriate modifications); and in this sense the science is thus “applied”. Complementary to this is the model which starts with jobs to be done (or functions to be performed); devices or processes are designed to that end, and scientific information is used (with appropriate modifications) to ensure that the designs work. The first model is fostered by academics, either anxious to see engineering accepted as a “science”, or anxious to justify support for “pure” research on the grounds of its possible eventual utility. The second is adopted by professionals in engineering or design professions, who start with a problem to be solved, not a fact to be applied. Obviously, both models are correct some of the time, in some fields; and neither is correct all the time in all fields. In any given context of application, the task is to see which is appropriate.

It is only natural for those trained in academic science to want to construct models that are “basic”, or “general-purpose”; that way lies the advancement of knowledge, and also legitimisation within a subject-speciality community. And in some circumstances, that path will be appropriate. However, when our models are necessarily so incomplete and imperfect, it is hardly likely that any given one will be able to achieve excellence in all possible applications in all possible circumstances. Hence it becomes appropriate for our models to be *designed*, rather like a building, a tool, or software, with a particular function in mind. And a design that is perfect for one function, may be most inappropriate for another; among cars a Jeep and a Bentley each have their own excellence, but their particular performance criteria have little overlap.

We can appreciate the difference between the two sorts of problem-solving by considering how they employ criteria of quality. Research science has been called “the art of the soluble”; and this implies that the investigation of an insoluble problem (one where no attempted solution can satisfy the accepted criteria of quality) is a waste of time. But when we are confronted with the challenges of global climate change, or even some local planning issue, we do not have the luxury of deferring our research until we are sure of success. When we consider the criteria of quality for Integrated Environmental Assessments, we must keep in mind that standards can actually be set too high, and become counterproductive. The use of these instruments is reminiscent of a wartime situation, where we must do our best under the extreme circumstances. And that “best” will also be conditioned by time-frame, cost, and conditions of use. Devices that are too beautifully designed may turn out to be counter-productive; the Germans discovered this, to their cost, during the last war. A multiplicity of independently designed models of equipment, each quite excellent in its way, led to confusion in use and nightmares of spares shortages. (Wheeler 1996). By contrast, the legend is that British had the principle of designing for “third best”, since the very best design would come into use after we had lost the war, and the second best would be there after the enemy’s countermeasures were in place.

Further, when designing around particular criteria, we discover that some sorts of information are more critical than others, as performance is more demanding in some respects than in others. In some contexts, engineers must do full computer modelling of their imagined structures; in others, “back of envelope” reckoning is quite enough. To know which to do when, is an important part of the craft. Hence in any design exercise, the relevant items of scientific information need not all be of equal strength; and to require that it all be the best possible, is recipe for waste and inefficiency. The relevance of this for Integrated Environmental Assessment is that in the construction of a particular sort of model, its intended use should be kept in mind. And then both the requirements on its inputs, and the expectations on its outputs, will be appropriate to its intended use.

4.3 History.

Passing now to history, it might seem odd now, to imagine it as a practical, or policy–related discipline. Although history can be very educational or entertaining, no one now looks to the history of the past to teach us how to manage the future. But this relegation of history is itself a very recent thing; many times in the evolution of our culture, and most recently in the nineteenth century, people did believe that history could provide the lessons we need. In this sense it was an “élite folk science” of the sort I have mentioned previously. It was related closely to “classics”, those stories whereby the young were given a picture of the world which provided edification and guidance. The perennial “two cultures” battle was then between this form of learning taken from the secular sphere, and the corresponding material from the dominant religion. All that is behind us now; “modernisation” means that those sources of authority have lost their strength, and élite folk–sciences must present themselves as natural-scientific.

What is widely seen as a crisis in modernity (manifesting partly as “post–modernism”) is that in its turn Science has now lost much of its symbolic power. This is the background to the concern with “good practice” in Integrated Environmental Assessment; in an earlier, optimistic period only a few decades ago, it would have been obvious that any enquiry using quantitative data and mathematical methods must achieve success. In this context, we can ask whether there are some strengths in the historical approach, not as an automatic provider of wisdom, but in its style of argument and in its capability of providing information that is useful in the policy process.

An awareness of history provides three principal benefits for good practice in Integrated Environmental Assessment: one concerning the world out there; another concerning our study of it; and the third concerning the formation of our policies about it. We now know that “the environment” as we interact with it, is not a simple system to be shaped at our will through science and technology. It is complex, exhibiting feedback, growth and surprise; and to the extent that it includes ourselves, it is reflexive, including awareness and purposes that may be contrary to those we think best. Moreover, the “environment” as we conceive it, is a product of an interaction in which our prior ideas condition what we do and what we see. And these prior ideas come from the past, frequently in the sorts of assumptions that we adopt unselfconsciously during our years of maturing. So one very important function of history is to enhance our awareness of ourselves, to see how the “common–sense” we deploy is actually a cultural product.

There is another important lesson to be learned from an appreciation of history: and that is that “science” need not be a matter of numbers and formulae. Indeed, this prejudice is a very parochial one, partly a matter of language; in German, “Wissenschaft” refers to any disciplined study. And, mainly in Germany, the historical or “philological” disciplines actually achieved maturity, in their institutional form at least, well before natural–science. Starting with the study of the classics in the late eighteenth century, German scholars established the institutional apparatus of research and teaching, with codified methods, long before this was accomplished for the natural sciences (Merz 1904). Of course there were frequent debates about the relative merits of the two sorts of “Wissenschaft”; but it was never in doubt that History could indeed be a Science.

The lesson of history as a science is that non–quantitative reasoning, and non–experimental evidence, can be just as rigorous, demanding, and fruitful as any other sort. Of course, this is known in the “historical” natural sciences, including geology and palaeontology; and even in the “field” sciences including taxonomy. In none of these cases, nor in history itself, are quantitative arguments avoided; quantities are invoked where they are relevant and useful. One traditional criticism of the historical sciences was that their conclusions lack precision and rigour; only in a mathematical natural science can we really say exactly what is out there. This may well be the case, where it works; and historically it has worked well, for the simplest natural systems. But the whole point of Integrated Environmental Assessment is our discovery of the need to cope with those systems which are not simple, for which the attempt to reduce reality to a linear equation is futile.

4.4 History and the Maturing of a Discipline.

In the current state of immaturity of Integrated Environmental Assessment, we find that debates on methodology are inescapable. Indeed, the politicisation of methodology is a characteristic feature of the whole endeavour. Each side will interpret data and make inferences according to its own agenda, and then criticise those of opponents. In terms of NUSAP, the Numeral in a quantitative statement may be less significant, for policy purposes, than the Pedigree. Recent disputes over statements on global climate change reflect this change of focus (Feder 1996). Now, most people witnessing such debates, and many of the scientists themselves, find them confusing and dismaying. Thomas Kuhn had a very deep insight when he described the achievement of “maturity” of a scientific field in terms of the stifling of debate on fundamentals (J.R. Ravetz 1990b). And certainly, students of science are given very little preparation, during their formal training, for coping with methodological disputes. How many science examinations, even at final University level, have questions beginning with “Compare and critically evaluate...”? In this sense, as Kuhn saw, science teaching has become dogmatic, and its graduates are ill–prepared for coping with the scientific problems that are coming to be recognised as among the leading ones for our age. Indeed, if we consider a reflective awareness to be important for professional practice on the policy problems of a complex, reactive natural environment, then in spite of their technical competence, Science graduates may be less well prepared in crucial aspects than those from the Arts. Perhaps it was one of the many implicit ironies in Kuhn’s account of science, that this essentially dogmatic and authoritarian state of affairs is described as “maturity”.

The example of historical scholarship demonstrates that Kuhn’s vision of maturation was unnecessarily simple. He contrasted those fields which wallow in interminable debates on fundamentals (either eighteenth–century physics, or contemporary “behavioural science”), to those which just cut out the arguing and began to solve puzzles successfully. He never mentioned the historical or philological disciplines, which combine a full critical self–awareness with a genuinely positive, progressive content. Now that we are attempting to make scientific practice more rich and self–aware, the example of history can be very important. Good practice in Integrated Environmental Assessment can indeed comprehend complexity in the world out there, and foster healthy debate within the discipline. The price to be paid is the loss of the simplicity promised by traditional science, but we now know that that is an illusion anyway.

4.5 History and the evolution of design and of good practice.

History provides other insights, in describing and explaining changes in design and in practice, through an interaction of external stimuli with internal possibilities. A few of us are old enough to remember the earlier, inconvenient shapes of such familiar devices as the automobile, the aeroplane and computers, before they matured and stabilised their form and performance. Computers themselves bear the marks of their very special history. There were crucial moments of choice, as when the earliest digital machines replaced analogue computers, because of the demand for large-scale calculations for ballistics. Then again, as computers became smaller and more powerful, and the cumbersome mainframes began to give way to PC's, the dominant architecture still followed a numerical pattern (IBM) rather than the graphical (Apple) because of existing commitments and styles. Even now, the design of the PC reflects its history as an instrument whose users are expected to be skilled, so that its diffusion, first to offices in the vain pursuit of productivity, and now to households, has lagged behind expectations. The point has been well put in a recent article in *Wired*:

"Truth be told, the PC is radically different in design from any other technology or appliance used by the general consuming public. Indeed if Intel and Microsoft had designed our kitchens, we'd probably all be using a $3,000 multipurpose 'Kitchen Processors' rather than the low-cost, push-button, dedicated appliances we now have. These Kitchen Processors, of course, would need to be configured and launched to, say, heat coils for a toasting application, and then reconfigured and relaunched to heat incandescent filament for a lighting application. As the basis for a truly ubiquitous Internet, today's PC simply won't cut it."

The point of the article is that change is on the way, in the form of dedicated chips which will not merely perform many special functions, becoming totally user-friendly and hence invisible, and which will communicate via The Net (D. Kline 1996).

Similarly, we can imagine the present generation of large-scale climate models, designed around the big (and perhaps ultimately insoluble) problem of predicting climate change, gradually going the way of the old mainframes. Their replacements would utilise new powers of computation and representation, and be crafted more directly around the new functions of enhancing clarity and facilitating discussion. Such a "miniaturisation" might involve tasks which are less challenging, and perhaps less impressive in various ways. But this has been the path of development of most computing over the years; and throughout the most powerful computers have still retained plenty of work.

With such an historical perspective, the problems of “good practice” take a new and hopeful shape. There is no need to achieve a full consensus on present techniques with all their difficulties, as if (uniquely in the world of science and technology) they can be expected to remain unchanging. Rather, discussions of “good practice” can be (in part) explicitly oriented around the expectations of the functions of models, and the performance of computers, of the near future. Good practice will then be in part an exploration and dialogue, in which new users participate with providers. In this way, the evolution of practice will go hand in hand with the maturing of the field of Integrated Environmental Assessment.

5. Living in Interesting, or Post-Normal Times

5.1 Professional work as distinct from research

The burden of my argument in the previous chapter can be cast in the terms of recent theoretical studies by Silvio Funtowicz and myself. There has been a natural tendency for practitioners of Integrated Environmental Assessment to consider it as an “applied science”, one in which scientific knowledge and techniques are deployed in situations where, typically, both "systems uncertainties" and "decision stakes" are low. Then routine puzzle–solving is adequate for ordinary problems; in policy–related research, this is the analogue of Kuhn’s “normal science”. Although this approach has had obvious imperfections, neither the inherited philosophy nor the inherited ideology of “science” has provided materials for coping with the new circumstances of Integrated Environmental Assessment. To some extent, and particularly in connection with design and engineering, I have here been advocating a style which we have called “professional consultancy”. In this, either or both of the cognitive and values aspects of the task are problematic. There may be marked uncertainties, as on a large engineering project, or significant stakes, as in performing novel surgical procedures. Then, craft skills of judgement become important, as well as moral qualities of decisiveness and integrity.

We have already gone through the problems of uncertainty, which indicate one key difference. Research science has been called “the art of the soluble”; and this implies that the investigation of an insoluble problem is a waste of time. But when we are confronted with the challenges of global climate change, or even some local planning issue, we do not have the luxury of deferring our research until we are sure of success. When we consider the criteria of quality for Integrated Environmental Assessments, we must keep in mind that standards can actually be set too high, and become counterproductive. The use of these instruments is reminiscent of a wartime situation, where we must do our best under the extreme circumstances. And, as we have seen, that “best” will also be conditioned by time-frame, cost, and conditions of use. Any good decision-process must include an estimation of “error-costs”; and these will necessarily be even more uncertain than the projections of the planned actions.

For science involved in the policy process, the management of uncertainty cannot be accomplished in isolation from the incorporation of decision stakes. The professional must handle these related tasks explicitly, depending on the circumstances of each case. The difference from research science is not absolute; for any interpretation of data relies on statistical tests and indicators, themselves shaped by particular chosen parameters (such as “confidence limit”). It is a communal decision of a field, which may be adopted by individuals quite unselfconsciously, how to shape those tests. There will always be a risk of admitting false hypotheses and a risk of rejecting true ones; there must be a choice determining the balance of those risks, and that choice will reflect the underlying values of the field. Further, there are the problems of managing “outliner” data, in which craft skills of scientists, explicitly guided by considerations of value and costs, are deployed. However, such problems have generally been ignored by philosophers and publicists of science, and in matured fields they have been managed by craft skills, and so they are usually not at the centre of attention of research practice. As a result, the understanding of the role of values in shaping the research process among the public and among the scientists themselves is frequently even more defective than their understanding of the management of scientific uncertainty. Only when there is a great public debate on some health or environmental issue, does it become clear that experimental data are both selected and shaped by the values underlying the problem-definition in whose terms they are produced.

5.2 Post-Normal Science and the inadequacy of bare numbers.

If I were in a position to advocate that Integrated Environmental Assessment come to recognise itself as a variety of “professional consultancy”, then the message of this report would be simpler and more comforting than I believe it can be. To put the matter succinctly, while I believe that professional “good practice” *must* be achieved in Integrated Environmental Assessment, and also that it *can* be achieved, I maintain that it alone is not sufficient for the challenges which confront us all, and practitioners of Integrated Environmental Assessment in particular. What is still to be mastered is a “good practice” for “Post-Normal Science”, for that is the character of Integrated Environmental Assessment in the present period.

We define “Post–Normal Science” as the appropriate problem–solving strategy for policy issues when either systems uncertainties or decision stakes (or both) are high. One important feature of our analysis is that even if uncertainties are low, then so long as stakes are high, there will not be a straightforward resolution of the issue. This is most easily seen in the forensic context, when a contestant with much at stake (and sufficient resources to deploy) will utilise every possible uncertainty in the opposing case to fend off defeat. Sometimes this practice is carried to lengths which seem abusive of “due process”; but the point here is not whether it is right or wrong, but that such extreme cases highlight a universal feature of such situations.

As we are reminded fairly frequently nowadays, the scientific inferences concerning risks and the environment are fraught with ineradicable uncertainties. Inferences from high–dose toxicants on test animals to chronic–dose exposures by humans cannot be exact. Epidemiological studies can provide information only on questions which are reflected in the previously defined structure of sampling and testing, and not always then. Predicting any sort of future is beyond even the economists; how then can ecologists provide more certainty than they?

Indeed, we may say that in such scientific studies, there has been a great deal of misplaced attention, as if the quantitative outcome of research is the crucial, or even dominant aspect of the result. Traditionally, in statements based on natural science, it has been assumed that uncertainties are negligible, policy implications unambiguous, and history irrelevant. The teaching of science, much of its internal discourse, and most of its popularisation, still employs these assumptions. But in the policy areas involving Integrated Environmental Assessment, we find that the bare number may be the least significant part of the communication; its interpretation in terms of “danger” is crucial and possibly contested; and the history of previous statements by the proposing institution may well be decisive in its acceptance or rejection.

We can express the point effectively in terms of the NUSAP notational system that Silvio Funtowicz and I developed some years ago. This is a deepening and an expansion of standard practice in the matured scientific disciplines, where it is well known that “exactness” is actually the result of skilled management of inexactness. There a quantity without an error–bar is meaningless, and “insignificant digits” are avoided. With NUSAP we systematise these principles of good craft practice, and provide a standard notational scheme for any expressed quantity. First we have **N**umeral and **U**nit, and then **S**pread to convey “error” or imprecision. Following this we have **A**ssessment, which conveys a “quality–grade” of the information in some relevant aspect. An analogue for this in scientific practice is seen in some sophisticated measurements, where experimenters know from experience of deeper uncertainties (as in calibration, or the basic theory of the equipment) which cannot be conveyed in the “error–bar” of the given experiment, and so cite another ± term as a sort of **A**ssessment. Finally we have the **P**edigree, which is a coded expression (usually in matrix form) of an evaluative account of the history and intended function of the reported quantity. These three levels of uncertainty, technical, methodological and epistemological, were mentioned by Ralf Schuele in his discussion of the role of scientific consultancy in the design and implementation of climate control measures (Schuele 1996).

5.3 “Pedigree” as the crucial aspect of a policy number.

This last category is the most difficult for people to understand; but the progress of debates on environmental issues now makes it perfectly clear. Take, for instance, the current debate over the latest report on Global Warming and its possible causes (Feder 1996). The debate centred on the particular nuanced expression of anthropogenic influence on climate change. Here is a case where the “hard”, or “front–end” categories of the NUSAP scheme were not even in play; the issue was not to estimate the rate of warming, but (given that the rate is now significant) to indicate the degree of human influence. In NUSAP terms, we would say that the **N**umeral category had a very coarse metric, really “yes/no”, so that **U**nit was simply “significant change” and **S**pread was irrelevant. The real import of the statement was carried by the contents of **A**ssessment, which consisted of nuanced qualifications about the significance of anthropogenic causes. As published in the agreed “Summary for Policy Makers”, the conclusion read: “the balance of evidence suggests a discernible human influence on global climate”. But this **A**ssessment was subsequently inserted into the conclusion of one of the technical chapters of the report, to make it harmonise with the consensus. And then there were complaints that this text had not been cleared through all the review procedures. To characterise this complaint, we might think of **P**edigree as consisting of three phases: one of scientific production (to what extent does the work depend on models, or on statistical time–series, or on unverifiable input parameters, etc.); another on scientific quality–assurance (formal and high–quality peer–review, degree of consensus, openness of dialogue, etc.); and finally a phase for the various procedures and channels for drafting and approval of the text. The complaints focused on this last aspect of the **P**edigree. The argument was that since this last part of the process was defective, the particular nuanced statement in **A**ssessment in the technical chapter was not guaranteed of support by its relevant community, and hence was not justified as a consensus report.

It is not our present concern, to adjudicate on that debate. It is important for us to appreciate that the most policy–relevant aspect of the information as first produced was the **A**ssessment; and that the subsequent debate focused on the **P**edigree. This is a particularly clear and striking example of the phenomenon, where policy decisions now depend less on the “hard” aspects of the quantitative information than on the “soft”. If we think of a NUSAP quantitative expression as progressing in its categories from the “Yang” to the “Yin” aspects, then we are now witnessing a change in the conception of effective scientific information, from that where the “Yin” has been totally suppressed, to that where it becomes crucial. The debate also illustrates one of the principles of Post-Normal Science: however low may be the systems uncertainties, there will always be an issue where the decision stakes are so high as to make a policy-relevant scientific statement contested.

The presence of judgements, and the relevance of **A**ssessment and **P**edigree, are not at all unique to Post–Normal Science. Indeed, we have argued that value–loading pervades all the practice of science, including (as we have seen) even the most “normal” of puzzle–solving research. In “normal” practice, such value–commitments are adopted unselfconsciously by practitioners; in “professional consultancy” the values become explicit; and in Post–Normal Science, they may become themselves the subject of the crucial debates. As we say, in Post–Normal Science we must cope with “politicised methodology”, in an atmosphere that can be emotional and partisan.

I must make it clear that the focus on the uncertainties and value-conflicts in Post-Normal Science is not a counsel of despair. In our analysis, we have shown that the situation for any problem-solving strategy is not static, but is constantly changing in concert with the total problem itself. Thus, in the early stages of a policy debate, there will be few facts and many values; but then as research is stimulated and the problem becomes more focused, some research studies can have a critical importance; and it can happen that eventually the resolution of the problem can be assigned to the appropriate professionals and researchers. van der Sluijs (1996, Section 7) has provided an illuminating analysis of the way in which the Pedigree of an Integrated Assessment Model might be used as a guiding principle for a discussion of quality among experts. That is, one does not expect them simply to sit down and agree on a Pedigree; but the organised discussion which leads towards consensus on a Pedigree can be of great assistance in focusing debates which might otherwise wander inconclusively over all aspects of quality.

5.4 New social frameworks for policy-relevant science.

This last aspect of policy debates, quite visible in the case of global climate change, entails modifications in the appropriate conception of good practice in the social conduct of science. The old “ethos” of academic science (as crystallised by Robert Merton just before its replacement as the dominant form of practice) cannot be sustained. In particular, it is no longer possible to demand what he called “disinterestedness”, where institutional loyalties and professional commitments must give way to superior argument in the common quest for truth (J.R. Ravetz 1995. p. 312). Now, truth (or even full consensus) is likely to be unattainable in practice, and the parties in debate are anyway concerned less with knowledge than with policy. Their loyalties are then likely to be more with a cause or an institution, than to the peer–community of specialists in a discipline. It is all too easy for those with a loose institutional affiliation to condemn the others as unscientific, partisan or corrupt, as the case may be. Whether such accusations are sometimes correct is not the issue here; what counts is that policy issues involving science are substantively different from debates over scientific results. They involve different interests, and different codes of etiquette and ethics. This may be a source of grave difficulty, of distress and of regret; but it is real, and it is here to stay.

This new situation affects Integrated Environmental Assessment when its results, which are uncertain at best, and whose uncertainties may be very difficult to express with clarity (and certainty!), are adopted, transformed and perhaps distorted by one or another side in a policy debate. Individual scientists may find themselves torn between professional integrity and loyalty (or obligation) to an employer or client. Measures designed to make scientific research more “relevant” to policy may actually force scientists to be subservient to their bosses (Macilwain 1996). And scientists may find that the criteria of a media–led policy debate and those of scientific practice are just too much in conflict for any single individual to reconcile.

5.5 The Structures of Post-Normal Science.

All this is very new in the professional experience of many working scientists, and it can appear as if it is the dissolution and destruction of a world, or rather of the only world in which rationality and good practice can survive. But we can try to make some sense of this new world of Post–Normal Science in terms of the “dimensions” in which problems are cast and resolved. For this there are two sorts of structures: one is on a vertical hierarchy (as in the traditional *scala Naturae*) with the “lower”, material dimensions, for which physical science is appropriate, leading to the “higher” dimensions of consciousness and culture (Funtowicz and Ravetz 1994). This could be modified somewhat, to fit with the practice of current debates, so that the more “material” aspects are at the centre, encased in the social and cultural dimensions or layers. An example is the “generic model framework for urban sustainability” (Fig. 1) (J. Ravetz 1996). This segregates the different sorts of policy considerations, and assessment methodologies, while displaying their structural connections.

Such diagrams remind us that a truly Integrated Environmental Assessment, in the spirit of Post-Normal Science, must take account of values, including those held by citizens as well as those expressed by consumers. In the last resort, policies for managing environmental change (or “sustainability”) will be effective only if they have the moral support of the great mass of people. It would be an illusion to hope that conflict can be entirely removed from such deep issues; for that would then amount to a rule by experts, and then the real issues of values and group interests would intrude covertly, beyond accountability and control. Policies conceived in a narrow way by a morally committed minority always breed corruption and failure (as did “Prohibition” in the U.S.A. in the 1920’s); and analogous policies designed by a technocratic elite will be similarly ignored or subverted.

It will be important to take the great policy debates as much as possible out of adversarial processes, and to foster a style of sharing experience, so that even if one side disagrees strongly with the other's solution, they can still sympathise with their problem. There is now an increasing set of forums and of techniques whereby even deeply contested issues can be debated in a spirit of mutuality and of a desire to reach a common acceptable solution. It could be that Integrated Environmental Assessment, appropriately adapted, could become an essential instrument in such debates, providing critical information and guidance on structuring the issues. Others have already imagined a situation where everyone could play “what-if”? on a sufficiently rich model of a policy problem under discussion. Then extreme positions could be more easily modified, and (given enough variables in play) negotiations could lead to “win-win” agreements. To imagine such a model for the whole global environment is quite unrealistic; but given the rate of progress in information technology, it is reasonable to suppose that within a few years, there could be packages designed for just such an educational policy interaction among citizens on local and regional issues. This could be an important outcome of the technical work now being undertaken in connection with the ULYSSES project. It could be an appropriate goal for the development work on tools which will accompany the maturing of Integrated Environmental Assessment.

Conclusion: the maturing of a new conception of science through the practice of Integrated Environmental Assessment.

It is not for this report to provide detailed prescriptions for particular measures to enhance and maintain good practice in Integrated Environmental Assessment. That is for an ongoing dialogue among the skilled and committed practitioners in the field. On the technical side, there are already some impressive efforts in defining the elements of good practice, in the management of uncertainties, and in the use of “triangulation” for gaining a plurality of independent perspectives on any set of phenomena (Bailey *et al.* 1966).

My contribution to that process is best made along the lines of clarifying the framework in which it will be conducted. I have attempted to make clear and explicit some issues that have probably been in the consciousness of many, but which have not hitherto been brought into the open in a convenient and accessible form. It can take some time for the consensus among the most aware leading practitioners to diffuse out to all those involved as producers or users of the techniques. Until that process of diffusion is complete, there is a danger that evaluations of Integrated Environmental Assessment will oscillate between complacency and despair, and that extremely polarised views on policy issues will become accepted as representative. My recommendations are along the lines of identifying particular areas where this emerging consensus would best be enhanced and diffused.

We can start with an agreement on fundamentals that is now, I believe, nearly universal among practitioners. This is that Integrated Environmental Assessments cannot provide a “truth machine” for policy. Whatever could have been accomplished by mathematical methods in the past for the solution of problems in science, technology or policy, the present problems will not yield to that approach taken on its own. This assumption provides a relief from a burden on the one hand, and a new sort of challenge on the other. For there is now no need for practitioners to claim that they can provide The Answers; they will have their colleague communities behind them when they refuse to supply policy–makers with “magic numbers” whereby decisions can be given a spurious scientific rationality.

On the other hand, this admission of incompleteness raises difficult questions. Some important quantitative models do attempt an “integration” over a wide range of factors, focusing on economic variables while adding those for the biosphere and technology. Are we saying that these too are incomplete and must be supplemented? But how is this supplementation to take place? The form of all models is deterministic; and although a skilled person may interpret them in a probabilistic way, it is not immediately obvious how non-quantitative considerations, or reviews of assumptions, are to be brought into the reasoning. The enhancement of such models, which are apparently fully integrated, will call for skills which are not merely technical, but which must also involve a degree of political awareness.

It may be that one of the biggest challenges we face in the establishment of good practice in Integrated Environmental Assessment, is an inbuilt cultural expectation that because the techniques look like those of matured natural science, there is something wrong with us if they do not yield the same certainties. This is not only a matter concerning users; practitioners too may be confused. Although there is a strong current of popular scepticism about mathematics applied to social affairs, as in economics or in statistics in general, we do not yet possess techniques enabling the uncertainties in their conclusions to be assimilated into debates. It is otherwise with what we might call “moral” considerations; when we find strong partisan disagreement, and such a conflict of certainties that in effect they cancel each other out, we may see a practical problem for the governance of society, but we are not dismayed. A new sort of understanding will be required for comprehending the uncertainty of scientific, quantitative conclusions in spite of the apparent certainty of their form.

The implications of such developments for good practice in Integrated Environmental Assessment are actually quite encouraging. Given that policy debates will always tend to be polarised in form, and that proponents will utilise whatever symbols they have to hand, there will at least be a common understanding of the powers and limits of science and scientists in these contexts. Good practice will eventually come to include an “uncertainty and quality audit” on scientific information, analogous to the environmental, safety and quality audits in the commercial and administrative contexts, which responsible persons are required to provide to the regulatory authorities or to various concerned publics. To provide a safety audit which recognises the possibility of accidents is not to expect or condone such events, but rather to have protection against them. Similarly, to expose the sorts of uncertainty and degrees of quality in information is not to pass off substandard products, but rather to enable users to make their own reasoned judgements of its strengths for their purposes. I recommend that the construction of such audits for information in Integrated Environmental Assessments should be taken up by the Integrated Environmental Assessment Research Forum as an urgent and fruitful task.

In the last resort, when we give up the belief that somehow an objective science will be there to solve the problems of our social existence, then we must fall back on more robust and resilient approaches. These cannot provide the certainty which once seemed to be offered by science; but that is lost anyway. Also, the policy issues at stake when Integrated Environmental

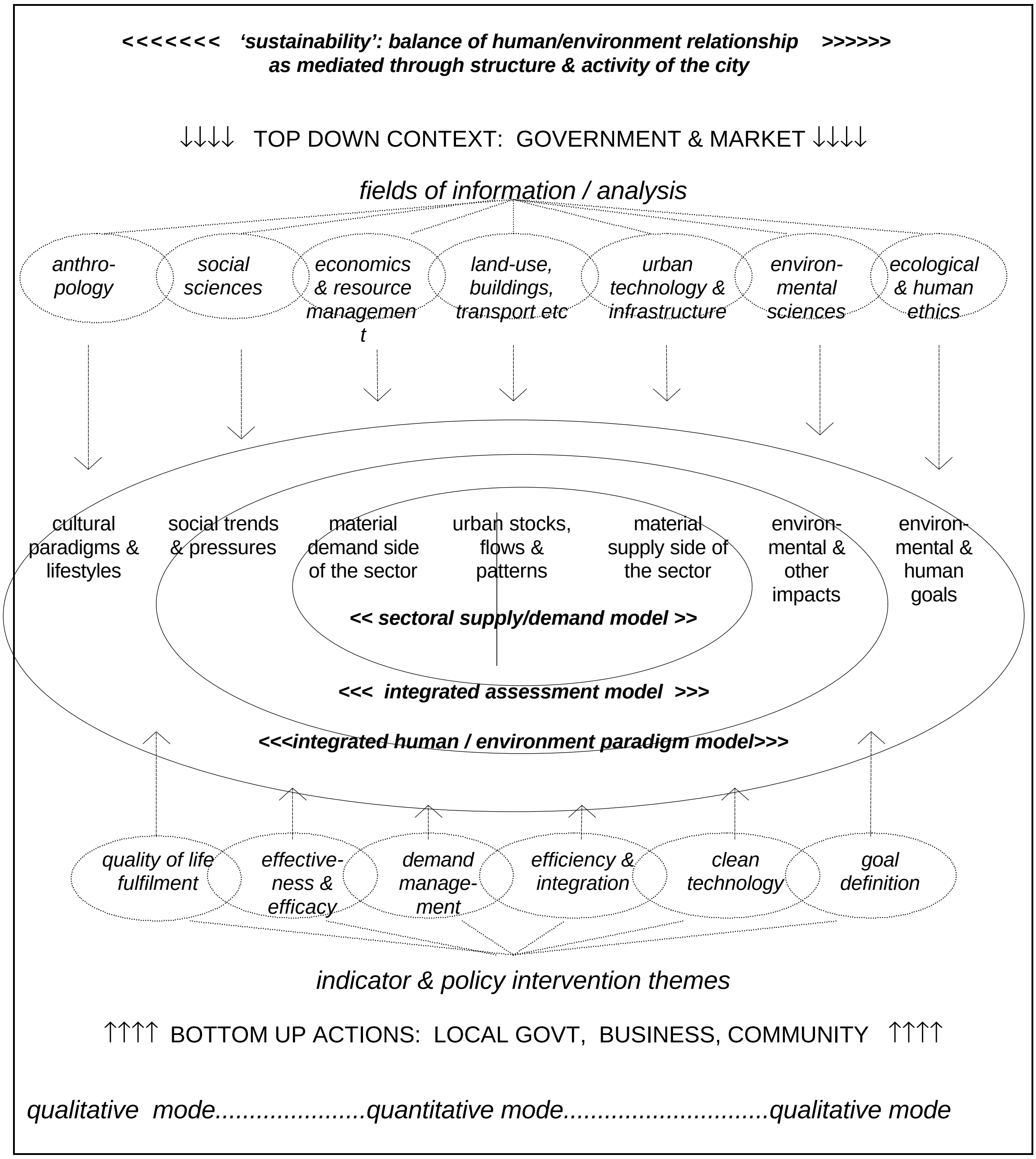
Assessment is invoked are in a sense total. In an important sense, achieving a sustainable industrial and social order is a matter of social morality. No amount of technical innovation and state regulation will avail unless eventually there are radical changes in lifestyle. And those involve an alteration in our conception of ourselves as inhabitants of this planet. Such a change may well require generations for its accomplishment, and will certainly encounter confusions and contradictions en route. But that is always so in social change, and we will manage the problems better if we are prepared to see them as challenges.

At the present stage of that deep transformation, the principal analytical tool is what we call Integrated Environmental Assessment. We no longer believe that somehow the facts of Nature can directly entail the right prescriptions for human conduct. Hence good practice in a fully developed Integrated Environmental Assessment will necessarily be forged in the context of debate and struggles over the definition of goals, and between the ideal and the practical. The enrichment of the techniques of Integrated Environmental Assessment, beyond the mathematical and into those involved in dialogue and mutual education among citizens, will take place only partly in conferences like the one at Toulouse in October 1996; there will also be forums involving participants who will bring their commitment and passion to the table. The name “Integrated Environmental Assessment” is new, and it may eventually be replaced by another, corresponding to another form of practice, technical and social. I have already discussed how our tools will inevitably evolve, becoming better designed around their evolving functions, and their practice becoming correspondingly more matured.

For a long time we have believed that in science, truth and falsity have hard edges. For every scientific problem there has been believed to be a single true answer at the back of the big book. Now the admission of radical uncertainty in post–normal science enables a great enrichment of our perspective. All our theories and models are like imagines of the complex reality out there, offering partial, overlapping and complementary insights. We can allow that two apparently contradictory scientific statements might contain the seeds of a dialogue whereby they could both be enriched. The dogmatism of Kuhn's "normal science" can now be allowed to wither away, being replaced by a practice in which the humanity of all participants can flourish.

We are here involved in attempting to foster a change in human consciousness comparable to that accompanying the previous transition from agrarian to industrial societies. This is a many–sided effort, in which poetry and science will each have their essential parts to play. The process will not be tidy, nor will it be safe, either for the participants or for humanity. Given all the good will in the world, there are enormous problems still to be solved. Although the mystification of mathematics may have been alleviated, it is still an esoteric learning. Computer displays which attempt to be transparent may unintentionally produce a false sense of simplicity and clarity. It will be a task involving all the talents, to establish a dialogue across the barriers, intellectual and social, of exclusive expertise. In the ULYSSES project, we are just now beginning the series of experiments in such communication among citizens, preparing to make those necessary mistakes from which we intend to learn. Good practice in a matured Integrated Environmental Assessment will involve a total commitment to learning, from those on all sides of the debates; and in this it can point the way to a new conception of science, one appropriate to the post–normal age.

GENERIC MODEL FRAMEWORK FOR URBAN SUSTAINABILITY



F i g u r e 1

References

Bailey, P. *et al*., 1996: *Methods for Integrated Environmental Assessment: Research Directions for the European Union* (Stockholm, Stockholm Environment Institute

Davis, P.J. and Hersh, R., 1986: *Descartes' Dream* (Brighton, Harvester), 159–164.

Cole, S., 1978: "The global futures debate 1965-1976", in Freeman, C. and Jahoda, M. (eds.), *World Futures: The Great Debate* (London, Martin Robertson), 51-75.

EPA 1988: The potential effects of climate change on the United States. Draft Report to Congress. U.S. Environmental Protection Agency, Washington, DC.

Feder, T., 1996: "Attacks on IPCC Report Heat Controversy Over Global Warming", *Physics Today*, August 1966, 55–57

Funtowicz, S.O. and Ravetz, J.R., 1990: *Uncertainty and Quality in Science for Policy*, (Kluwer, Dordrecht).

Funtowicz, S.O. and Ravetz, J.R., 1992: "Three Types of Risk Assessment and the Emergence of Post-Normal Science", in S. Krimsky and D. Golding (eds*), Social Theories of Risk*, Westport CT, Greenwood, 251-273.

Funtowicz, S.O. and Ravetz, J.R., 1993: "Science for the Post-Normal Age", *Futures*, 25/7 (September), 735-755.

Gates, W.L. *et al*., 1996: "Climate Models – Evaluation", in J.T. Houghton *et al.* (eds) *Climate Change 1995, The Science of Climate Change* (Contribution of WGI to the Second Assessment Report of the Inter Governmental Panel on Climate Change), (Cambridge, Cambridge University Press.

Hellström, T., 1996: "The Science–Policy dialogue in transformation: model–uncertainty and environmental policy", in: *Science and Public Policy* 23/2, 1996, 91–97.

Henderson–Sellers, A., 1996: "Bridging the Climate Gaps", in: T.W. Giambelluca and A. Henderson–Sellers, *Climate Change: Developing Southern Hemisphere Perspectives*, (New York, John Wiley & Sons)

Kandlikar, M. and Risbey, J., 1996: *Developing Insights from Integrated Assessment Modelling* (draft)

Keepin, B. and Wynne, B., 1984: "Technical Analysis of IIASA energy scenarios", *Nature*, 312, 691–695.

Kline, D., 1996: "The Embedded Internet", *Wired* (U.S.) 4/10 (October), 98–106; 102–3.

Macilwain, C., 1996: "US panel backs new approach to risk", Nature, 381 (20 June), 638.

Mermet, L. and Hordijk, L., 1989: "On Getting Simulation Models Used in International Negotiations: a Debriefing Exercise", in: F. Mautner–Markhof (ed), *Processes of International Negotiations*, Westview Press, Boulder CO, 205–237.

Merz, J.T., 1904: *A History of European Thought in the Nineteenth Century*, cited in J.R. Ravetz 1995, p.168.

Morgan, M.G. and Henrion, M., 1990: *Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis* (Cambridge, Cambridge University Press).

Nordhaus, W.D., 1991: "To slow or not to slow: the economics of the greenhouse effect", *The Economic Journal*, 101, 920–937.

Parson, E., 1995: "Integrated assessment and environmental policy making", *Energy Policy*, 23/4-5, pp. 463-475.

Ravetz, J., 1996: "Generic Models for Urban Sustainability", *Proceedings of the 1996 International Sustainable Development Research Conference*
x Bradford, U.K., E.R.P. (ISBN 1 872 677 05 3)

Ravetz, J.R., 1990: The Merger of Knowledge with Power: Essays in Critical Science, (London, Cassell).

Ravetz, J.R., 1990a: "Hardware and Fantasy in Military Technology", in J.R. Ravetz, 1990, 81-94.

Ravetz, J.R., 1990b: "Ideological Commitments in the Philosophy of Science", in J.R. Ravetz 1990, 180–198

Ravetz, J.R., 1993: "The Sin of Science: Ignorance of Ignorance*", Knowledge (Creation-Diffusion-Utilisation*), 15/12, December 1993, 157-165.

Ravetz, J.R., 1994: “Economics as an elite folk–science: the suppression of uncertainty”, Journal of Post–Keynesian Economics, 17/2, 165–183.

Ravetz, J.R., 1995: *Scientific Knowledge and its Social Problems* (with a new introduction) (New Brunswick, NJ, Transaction Press).

Risbey, J. et al., 1996, *Assessing Integrated Assessments*, (draft)

Rotmans, J., and van Asselt, M. B. A., 1996: "Global Change and Sustainable Development", in Rotmans and De Vries 1996.

Rotmans, J., and de Vries, H.J.M., 1996: *Insights, no Answers: Global Change and Sustainable Development* (Cambridge, Cambridge University Press) (in press)

Schuele, R.,1996: *Advancing Tools for Integrated Assessment* (presented at the Annual Meeting of the American Sociological Association, New York, August 1996)

Shackley, S., 1996: *Trust in Models? The mediating and transforming role of computer models in environmental discourse* (draft)

Toth, F.L. 1995: "Practice and Progress in Integrated Assessments of Climate Change, A Workshop Overview", *Energy Policy*, **23**, 4/5, 253–267.

van der Sluijs, J., 1996: *Integrated Assessment Models and the Management of Uncertainty*, IIASA-Workingpaper no. WP-96-119, Laxenburg, October 1996

Wheeler, R., 1996: "Who Do You Think You Are Kidding Mr. Hitler?", *History Today*, May 1996, 56. (Review of *Why the Allies Won* by R. Overy (London, Jonathan Cape, 1995)

Wynne, B., 1984: "The institutional context of science, models, and policy: The IIASA energy study", *Policy Sciences*, **17**/3, 277-320.