

---

## **2.1 Uncertainty, assumptions, and value commitments in the knowledge base of complex environmental problems**

---

Jeroen van der Sluijs<sup>1</sup>

### **1. Introduction**

Present day complex environmental problems exhibit a number of characteristics that make them hard to tackle with normal scientific procedures. Funtowicz and Ravetz (1993) have called this class of problems *post-normal*. Typically, political decisions on these problems will need to be made before conclusive scientific evidence is available. Decision stakes are high: the impacts of "wrong" decisions based on the available limited knowledge can be huge. Actors disagree on the values that should guide the decision-making. The available knowledge bases are typically characterized by imperfect understanding (and imperfect reduction into models) of the complex systems involved. Models, scenarios, and assumptions dominate assessment of these problems, and many (hidden) value loadings reside in problem frames, indicators chosen, and assumptions made.

The science involved in such issue-driven risk assessments of complex environmental problems differs significantly from the science of curiosity-driven laboratory practice. Risk assessment regarding for instance anthropogenic climate change, GMOs, or endocrine disruptors, involve uncertainties of many sorts, not all of which can be tamed. To give an example from the climate issue, we cannot perform a statistically satisfying series of experiments to test the effect of higher atmospheric greenhouse gas concentrations, because there is only one Earth available, and even the one available is poorly monitored, and other factors that influence climate are - in contrast to the situation in a laboratory - largely beyond our control. The problem is further compounded by the value-laden context of risk assessment combined with the presence of ineradicable uncertainties and indeterminacies.

The available knowledgebase consists at best of a mixture of knowledge, assumptions and ignorance. It comprises bits and pieces of knowledge that differ in status, covering the entire spectrum from well-established knowledge to judgments, educated guesses and tentative assumptions.

In such a situation the classic mode of scientific analysis in the form of puzzle-solving within an unquestioned framework, or 'paradigm' is unfeasible. However successful this approach is in disciplinary research, it meets its limits when society is confronted with the need to resolve transdisciplinary policy issues regarding trans-national and trans-generational environmental risk on which yet no unquestioned frameworks exist.

---

<sup>1</sup> Copernicus Institute for Sustainable Development and Innovation, Utrecht University  
j.p.vandersluijs@chem.uu.nl

The multidisciplinary nature of complex environmental problems poses additional requirements with regard to the systematic analysis, documentation and communication of uncertainty, in order to remedy the well known problem that when quantitative information is produced in one disciplinary context and used in another, important caveats tend to be ignored, uncertainties compressed and numbers used at face value (Wynne, 1992; Van der Sluijs et al., 1998). Knowledge utilization for environmental management requires a full and public awareness of the various sorts of uncertainty and underlying assumptions. Knowledge needs to be robust both technically and socially (Nowotny, 1999; Nowotny et al, 2001).

Knowledge Quality Assessment can help to achieve a better awareness of the limits of science in relation to the task of knowledge producers to provide a scientific basis for policy debate. A widening in focus from "reducing uncertainties" to "coping with untameable uncertainties and complexities" is needed. This can avoid misunderstandings and undue expectations of the role and competence of science in complex environmental problems.

## **2. Models cannot remedy ignorance**

Since the eighties of the twentieth century computer models are increasingly being used in complex environmental assessments and foresight: they enable analysts to simulate reality and run several scenarios, thereby integrating knowledge from different disciplines. Applied systems analysis has become the dominant method in environmental assessment. The assumption ladenness of the models themselves, the use of models, the degree to which they can be validated or evaluated, and the transparency of models have been criticized over the years.

To give some examples, Hornberger and Spear (1981, cited in Saltelli, 2002) argued that "*.... . . most simulation models will be complex, with many parameters, state-variables and non linear relations. Under the best circumstances, such models have many degrees of freedom and, with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values..*"

Oreskes et al, 1994 highlighted the assumption ladenness of models and concluded that "*Like a novel, a model may be convincing – it may "ring true" if it is consistent with our experience of the natural world. But just as we may wonder how much the characters in a novel are drawn from real life and how much is artifice, we might ask the same of a model: How much is based on observation and measurement of accessible phenomena, how much is based on informed judgment, and how much is convenience? Fundamentally, the reason for modeling is a lack of full access, either in time or space, to the phenomena of interest. In areas where public policy and public safety are at stake, the burden is on the modeler to demonstrate the degree of correspondence between the model and the material world it seeks to represent and to delineate the limits of that correspondence.*"

Beven's (2002) concept of equifinality (the phenomenon that models may be non-unique in their accuracy of both reproduction observations and prediction) and Beck's (2002) closely related notion that almost all models suffer from a lack of identifiability (many combinations of values for the model's parameters may permit the model to fit the

observed data more or less equally well) further emphasise the problematic nature of models and model predictions as a source of knowledge for decision making.

Yearley (1996) argued that values and value-laden assumptions enter into the formulation of environmental issues before the 'facts' are even established by science. Yearley gives examples regarding carbon dioxide, ozone destroying chemicals, and biodiversity and concludes that though, at first sight, science might be thought to be clearly universal and thus incontestably applicable to global problems, in practice its universality can be deconstructed and undermined.

A similar argument is made by Stirling (2000, 2001), who stresses the critical dependence of final results of risk assessment studies to the starting assumptions made. He makes a case for sensitivity analysis to explicitly link alternative framing assumptions with the results which they yield and proposes a method of multi-criteria mapping for this purpose: "*Rather than being expressed as discrete scalar numbers, then, risk assessment results might be expressed as ranges of values, with the ends of the ranges reflecting extremities in the framing assumptions associated with different stakeholders in the appraisal process.*"

In a critique on (macro economic) integrated assessment models of climate change, Schneider (1997) stresses the importance of highlighting value-laden assumptions in order to ensure that integrated assessment models of climate change (as well as other environmental problems) "*enlighten more than they conceal.*"

Van der Sluijs (1997) argued that the building of environmental assessment models inevitably involves subjective choices and value-laden assumptions. Lack of transparency with regard to these assumptions and uncertainties and lack of reflection on how knowledge that is conditioned on these models and its assumptions differs from well-established knowledge, lead to misunderstandings in the science policy interface on the nature of this type of knowledge. There is a tendency to treat this knowledge as if it is not different from well-established knowledge. The history has many examples of scandals and loss of trust in the scientific basis for policies based on lack of understanding of the nature of knowledge stemming from model based assessment and foresight. A classic example is the scandal of the IIASA energy scenarios in the eighties: In a critical review of the models used for these scenarios Keepin and Wynne (1984) concluded that "*Despite the appearance of analytical rigour, IIASA's widely acclaimed global energy projections are highly unstable and based on informal guesswork. This results from inadequate peer review and quality control, raising questions about political bias in scientific analysis.*" They made a strong case for the need to conduct a rigorous analysis of assumptions in forecasting tools used in the energy field, and the need to test the robustness and sensitivity of results. They argue strongly for an open and accessible documentation and rigorous peer review. The case led to a crisis within the institute and has triggered institutional learning towards more attention for uncertainty and quality control to regain credibility for their work with peer communities and the public.

More recently the Netherlands National Institute for Public Health and the Environment (RIVM) encountered a similar scandal: Early 1999 De Kwaadsteniet, a senior statistician, accused the institute of "lies and deceit" in their State of the Environment Reports and Environmental Outlooks. In a Dutch quality newspaper (Trouw) he criticised RIVM for

basing their studies on the ‘virtual reality’ of poorly validated computer models while RIVM presents these results as point values with unwarranted significant digits and without elaborating the uncertainties. It triggered a vehement public debate on the credibility and reliability of environmental numbers and models. The case got front page and prime time coverage in the mass media over a period of several months and led to debate in the Netherlands parliament (Van Asselt, 2000, Van der Sluijs, 2002, Petersen, 2005). The case also triggered a learning process within the RIVM and led to the development of a guidance for uncertainty assessment and communication for the institute (Van der Sluijs et al., 2003, Petersen et al., 2003, Janssen et al., 2003, Van der Sluijs et al., 2004, Janssen et al., 2005), which we will discuss later on in this chapter.

### **3. Emerging insights on the nature of uncertainty**

In recent years, an increasing body of conceptual and theoretical work in the field of coping with uncertainties in the assessment of complex risks has been accomplished. Key insights from the field include (Van der Sluijs, 1997, 2002):

- Uncertainty is partly socially constructed and its assessment always involves subjective judgement;
- More research does not necessarily reduce uncertainty, it often reveals unforeseen complexities and irreducible uncertainty;
- Uncertainty is more than statistical error or inexactness of numbers: it is increasingly understood as a multi-dimensional concept involving quantitative and qualitative dimensions. Uncertainty can manifest itself at different locations in risk assessments (for example: context, system boundaries, indicator choice, model structure, parameters, and data). In problems that are characterized by high systems uncertainties, knowledge gaps, and high decision stakes, unquantifiable dimensions of uncertainty may well dominate the quantifiable dimensions;
- Most of present day uncertainty methodologies and practices focus only on quantifiable uncertainty. Methods to address qualitative dimensions of uncertainty and issues of problem framing and value loadings are absent or in its early stage of development. Uncertainty in for instance model structure, model assumptions, and model context require more attention;
- A new class of reflective methods to cope with uncertainty, starting from the post-normal science position is now maturing.

In the following these insights will be discussed in more detail

#### ***3.1 Social construction of uncertainty***

One can distinguish between a positivist and a constructivist view of scientific knowledge. In the positivist position, scientific knowledge is seen as objective. Facts are real. Facts and values can and should be separated. There is only one reality, it is the task of science to discover and understand that reality. In the social constructivist position, scientific knowledge and scientific authority is not objective, and is not removed from society. It exists within society, is socially constructed, and therefore, subject to social

context in which it is created and used. Facts and values cannot be separated. Often more than one interpretation of reality is tenable (compare to the concepts of equifinality of models and lack of model identifiability discussed in section 2).

The positivist interpretation has a tendency to lead people to expect that science can deliver certainty for every problem that can be phrased scientifically and that uncertainty can be remedied by better science and more research. It is however not congruent with the nature of the class of complex problems on which I focus in this paper: certainty on these problems is unachievable for practical reasons (Weinberg, 1972) and epistemological reasons (Funtowicz and Ravetz, 1993). In recent empirical research into uncertainty communication (forthcoming) our group at Utrecht University found that amongst policy makers a positivist view of science strongly dominates, while amongst scientists involved in environmental health risks there is much more awareness of the social constructed nature of knowledge on these problems. In the science-society interface, this difference in understanding of scientific knowledge hampers effective communication of uncertain scientific knowledge on complex policy problems.

A closely related notion comes from the field of Sociology of Scientific Knowledge. There exists a relation between social distance (from the knowledge producers) and the perceived level of uncertainty in scientific knowledge claims. MacKenzie (1990) developed the idea of the 'certainty trough' (Figure 1).

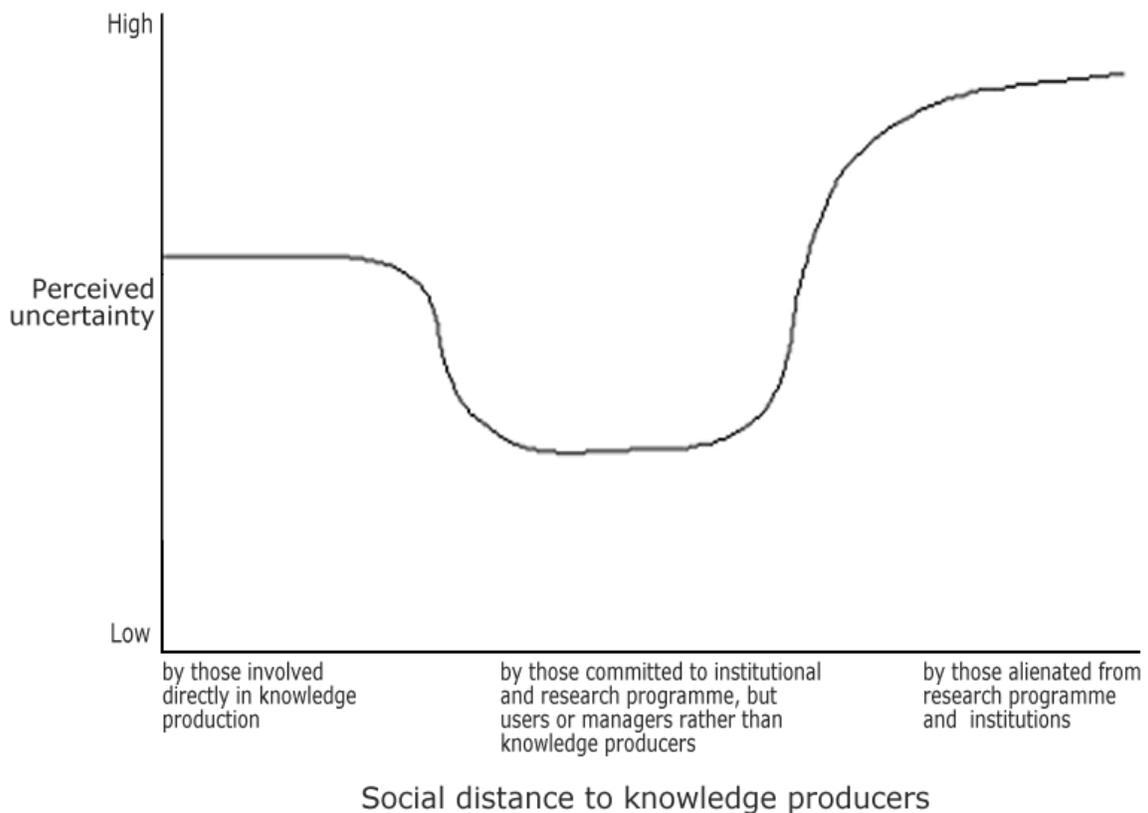


Figure 1. The certainty trough (MacKenzie, 1990)

The perceived uncertainty of knowledge claims is smallest a little bit away from the actual site of knowledge production. Scientists from one discipline may attribute less uncertainty to knowledge from another discipline than the practitioners of that other discipline would themselves attribute to it. If applied to the climate change case: policy-makers may attribute less uncertainty to claims by the official assessment reports by the Intergovernmental Panel on Climate Change (IPCC) than do the authors of the IPCC reports themselves. Those who feel alienated from the IPCC process, often referred to as climate sceptics, will attribute the highest uncertainty to the knowledge claims conveyed by the IPCC assessment reports.

Wynne (1992) noticed that the discussions about uncertainty seemed to rely implicitly on the naive notion that inadequate control of environmental risks is due only to inadequate scientific knowledge. Wynne criticized this idea, and added the concept of indeterminacy as a category of uncertainty. Indeterminacy refers to the open-endedness (both social and scientific) in the processes of environmental damage caused by human intervention. Indeterminacy introduces the idea that contingent social behaviour also has to be included in the analytical and prescriptive framework. It also acknowledges the fact that many of the intellectual commitments which constitute our knowledge are not fully determined by empirical observations. The latter implies that scientific knowledge depends not only on its degree of fit with nature, but also on its correspondence with the social world and on its success in building and negotiating trust and credibility for the science. We emphasize that this notion implies that virtually every scientific claim that comes under fire when it figures in a societal controversy with high decision stakes will turn out to be indeterminate and uncertain rather than foundational.

Other authors also have questioned the possibility of obtaining an objective definition of uncertainty. They argue instead that representation and perception of uncertainty are important factors, and that uncertainty is partly constructed as the product of implicit negotiation processes between scientists, policy-makers and the public (Collingridge and Reeve, 1986; Jasanoff, 1990; Van Eijndhoven and Groenewegen, 1991; Shackley and Skodvin, 1995, Van der Sluijs et al., 1998).

### ***3.2 To what extent can uncertainty be tamed?***

To explore the question to what extent uncertainty in the assessment of complex environmental problems can be tamed, I will analyze the phenomenon of uncertainty in the science policy and science society interfaces using a “monster” metaphor (see Van der Sluijs 2005 for a more detailed analysis). I borrowed this metaphor from the work of Dutch philosopher of technology Martijntje Smits (2002, 2004) on the cultural domestication of technologies. I use her monster-concept to explore the way in which the scientific community responds to the monstrous uncertainties that they face in the production of the knowledgebase of complex environmental problems. Building on work of Mary Douglas (1966) on purity and danger in traditional cultures, Smits has developed a theory in which she explains the co-existence of public discomfort and fascination with certain new technologies such as GMOs by using the metaphor of a monster. The idea is that we are accustomed to order the world in terms of binary categories such as humans versus animals, organisms versus machines. Douglas has shown that such categories differ from culture to culture. They are the result of a social learning process. For the

most part they are shared collectively, therefore they could be called cultural categories. They shape a symbolical reconstruction of worldly phenomena. A special case of confusion appears when at the same moment a phenomenon fits into two categories that were considered to be mutually excluding. Smits calls this ambiguous phenomenon, this unnatural being, a monster.

When we apply Smits' monster concept to the production of a knowledgebase for policy making on complex environmental problems we can make a number of interesting observations. The categories that we thought to be mutually exclusive and that now tend to get increasingly mixed up to create monsters in the science-policy interface include: knowledge versus ignorance, objective versus subjective, facts versus values, prediction versus speculation, science versus policy.

Smits distinguishes four styles of 'monster-treatment' with different degrees of tolerance towards the abnormal. These styles are: monster-exorcism, monster-adaptation, monster-embrace and finally monster-assimilation. In the following I will explore each of these strategies for coping with the monster of uncertainty.

**Monster-Exorcists** want to expel the monster. Uncertainty causes discomfort and does not fit within symbolical order where science is seen as the producer of authoritative objective knowledge. They call for more objective research that should aim at "reducing uncertainties". There is a strong belief in "objective science": the puzzle can be solved. This view can be found in for instance the first scientific assessment of climate change by the IPCC: "*We are confident that the uncertainties can be reduced by further research.*" (Houghton et al., 1990). The borders between facts and values, knowledge and ignorance, science and policy are seen as real and inflexible and often the categories are also seen as norms (as in the notion that is a good thing to keep science and policy, facts and values, objective and subjective separated). Yet monster-theory predicts these attempts will prove to be vain in the long run: for each head of the uncertainty monster that science chops off, several new monster heads tend to pop up due to unforeseen complexities.

It is interesting to see that for that reason, the IGBP (International Geosphere Biosphere Programme), one of the largest international research programmes on global change, concluded during their third Scientific Advisory Council Meeting in January 1993, that it might not be feasible to reduce uncertainties (Williamson, 1994). Williamson also notes that the increasing complexity of global models inevitably decreases the precision of their products and "*full predictability of the earth system is almost certainly unattainable.*" He makes a case for the replacement of the research objective to reduce uncertainties by a pragmatic research goal to "*provide reliable estimates of probability within defined limits, so that risks can be assessed and appropriate actions taken, rather than single value 'predictions' with spurious exactitude.*"

Another strategy to cope with unwelcome uncertainty that may fit in the monster exorcism style and that needs mention here for the sake of completeness is the strategic hiding of uncertainty. One can think here of keeping the uncertainties in knowledge claims deliberately under the table because they do not fit a political agenda, to avoid that one's knowledge claims are disputed, or because one is afraid that prevalence of uncertainty will be judged as poor science by the outside world.

The **Monster Adaptation** style attempts to fit the uncertainty monster back in the categories, a process that can be characterized as purification. In terms of taming the uncertainty monster this can be seen as attempts to quantify uncertainties. Monster adapters feel uncomfortable with anything that does not fit in a spreadsheet. They need numbers, for otherwise they cannot do their calculations that they deem to be the basis of rational decision support. Where there is no objective ground for quantification, monster adapters tend to use subjective probability and Bayesian approaches to quantify uncertainties in terms of the degrees of belief that experts assign to their knowledge claims. By normalizing the post normal along these lines, the classic paradigms of Decision Support striving for optimisation of expected utility as rational risk management strategy can be maintained. Monster adaptation by purification is further evident in the tendency to build system models based on “objective science” and then to externalise the subjective parts and uncertainties into ranges of scenario’s – grouped into story-lines representing different value orientations - that are used to feed these “objective” models. The model is then seen as grounded in science and belongs to the domain of the scientists. Policymakers and stakeholders are welcome to contribute their insights, but in the scenario’s that feed into the models, not in the science of the models itself.

The limitations of the purification, or monster adaptation approaches are obvious in the lack of convergence of outcomes of models build by different modelling groups but addressing the same issue. For example, the IPCC Special Report on Emission Scenarios (Nakicenovic et al, 2000) clearly shows that different models fed with the same scenarios produce very different results. This reflects the significance of model structure uncertainties (see also Refsgaard et al., in press).

Further, several authors have shown that current models are not so objective as they claim to be: they contain many value-laden assumptions and stem from value laden problem framings (see for overviews for instance Van der Sluijs, 1997, 2002).

The third response, **Monster Embracement** welcomes the uncertainty and can perhaps be associated with fascination about the unfathomable complexity of our living planet Gaia. It creates the possibility to be filled with wonder and respect, something that was taken away by the engineering worldview in which science is able to understand and control nature, reflected in notions seeing the biosphere as something that can be managed. The schools of thought of Holism and attempts to integrate science and spirituality in Inclusive Science can also be considered as embracing uncertainty, because it emphasizes the limits of the positivist *reductionist* schools of thought for which they provide alternatives.

Another type of response to monstrous uncertainty that can be seen as monster embracement is denial of the reality of environmental risks by pointing to all those uncertainties. Many actors have a stake in a given risk, be it food safety, global warming or electromagnetic fields from mobile phones, and therefore have agendas of their own and are not always averse to manipulating scientific uncertainty for various reasons. Just like science is often strategically used (selective and biased use of sources by quoting only those studies that fit in one’s policy agenda) by different actors in policy debates, scientific uncertainties are often magnified and distorted (Hellström, 1996). The uncertainty question can be (and is) actively used as a strategy to undermine the role of

assessment as a shared source of information, and to achieve the postponement of measures, or, by other actors, to achieve a ban on a new technology.

A variety of techniques is used to deliberately raise doubts about the realness of environmental risks, such as distortion and magnification of uncertainties (making mountains out of molehills), and even acts “at variance with good scientific practice” (as at some point the Danish Committee on Scientific Dishonesty qualified some of the things Lomborg did in his book *The sceptical environmentalist*). An interesting phenomenon is the 'manufacturing of uncertainty' as a corporate strategy by industry groups (Michaels, 2005). The unpleasant way in which such games are played and the mixture of valid and ungrounded criticisms that it produces, is the price that has to be paid for the key-advantage for quality control of the identification of weak spots in the knowledge base by those who have a strong incentive and drive to find these weak spots.

The fourth strategy, **Monster Assimilation** refers to not only adapting the monster, but also changing the cultural categories by which it is judged. As Smits (2002, 2004) argues, in contrast to the other styles monster-assimilation makes use of the insight that cultural categories are flexible and constructible. In other styles, the uncertainty monster is somehow judged in terms of existing cultural categories and these cultural categories are in turn considered more or less as facts.

Rethinking the categories that got mixed up in the monsters is at the core of the monster assimilation strategy. Post-normal science and other forms of reflexive science are clear instances of attempts to assimilate the uncertainty monster and give it a central and explicit place in a deliberative management of environmental risks. Because scientific consensus about the truth of complex environmental risks is unlikely to be achieved given the post-normal situation (facts uncertain, values in dispute, high decision stakes), we will have to drop our demand for a single certain truth and strive instead for transparency of the various positions and learn to live with ambiguity and pluralism in risk assessment.

Each of the four styles can be observed in the historic learning process and current practice of coping with the uncertainty monster in the science policy interface on complex environmental problems. We might see this ongoing process of learning to cope with complex systems as a dialectic process where one strategy tends to dominate the field until it fails followed by a rise of one of the other strategies. As evident in the Interfaces between Science and Society conference and the other contributions to this book, we now seem to find ourselves in a phase with a growing focus on monster assimilation.

### ***3.3 The multiple dimensions of uncertainty***

Uncertainty can be seen as a multi-dimensional concept involving quantitative (technical: inexactness) and qualitative (methodological: unreliability, epistemological: ignorance and societal: limited social robustness) dimensions and it can manifest itself at different locations in risk assessments (for instance, context, problem frames, indicator choice, model assumptions, model structure, model parameters, and data). Regarding the technical dimensions, *inexactness* refers to the numerical precision in data that can be hampered by limited resolving power in measuring methods or equipment or is caused by

“ordinary” spread in repeated measurements, but can also arise from variability in the “real world”. For instance, if a toxic gas escapes from a failure of a chemical installation, the way in which it spreads in the environment depends on the weather at the moment of the accident, and inherent natural variability of the weather limits our ability to assess what will happen in case of such an accident. A next dimension of uncertainty is produced by the limitations of the methods we use. *Unreliability* refers to the level of confidence in knowledge, in terms of quality, soundness, scientific status of the method etc. If a number is obtained by the method of educated guessing, it is less reliable than if it was based on a large sample of direct measurements. Another issue here is that the outcomes of many risk assessment studies are conditioned on a large number of assumptions made in the analysis, whereas the validity of these assumptions is usually poorly known. The third dimension, epistemological uncertainty, stems from limits to our capacity to know and understand: *ignorance* relates to all cases of “we don’t know what we don’t know”, but can take an active or passive form of that. Active ignorance is for instance at stake when a risk assessor is aware of the aspects in which his knowledge is limited, but has no insight in the relative importance of those aspects that he – therefore – has no good way to include these in the analysis. Passive ignorance is at stake when one is not aware of what one does not know.

### ***3.4 Quantitative uncertainty methods and their limitations***

Until recently, the field of uncertainty analysis mainly evolved around mathematical methods such as sensitivity analysis and Monte Carlo techniques. These tools address quantitative dimensions of uncertainty using sophisticated algorithms (Saltelli et al., 2000, 2004). Although these quantitative techniques are essential in any uncertainty analysis, they can only account for what can be quantified and thus provide only a partial insight in what usually is a very complex mass of uncertainties involving technical, methodological, epistemological and societal dimensions. For the class complex problems that we are concerned with in this paper, it is often the case that unquantifiable uncertainties may well dominate the quantifiable ones, which implies that these techniques are of limited value for this particular class of problems. Quantitative methods can however be complemented with new qualitative approaches addressing aspects of uncertainty that are hard to quantify and were therefore largely under-addressed in the past. These new approaches will be discussed in the following.

## **4. Reflective approaches to uncertainty**

In the school of post-normal science, several new multidimensional and reflective approaches to Knowledge Quality Assessment have been developed. I will discuss two key examples here, the NUSAP system and the checklist approach to knowledge quality assessment that was recently adopted in the Netherlands Environmental Assessment Agency (MNP) in their Guidance for Uncertainty Assessment and Communication.

### ***4.1 The NUSAP system***

NUSAP is a notational system proposed by Funtowicz and Ravetz (1990), which aims to provide an analysis and diagnosis of uncertainty in the knowledge base of complex (environmental) policy problems. It captures both quantitative and qualitative dimensions

of uncertainty and enables one to communicate these in a standardized and self-explanatory way. The basic idea is to qualify quantities using the five qualifiers of the NUSAP acronym: Numeral, Unit, Spread, Assessment, and Pedigree.

The first qualifier is Numeral; this will usually be an ordinary number; but when appropriate it can be a more general quantity, such as the expression "a million" (which is not the same as the number lying between 999,999 and 1,000,001). Second comes Unit, which may be of the conventional sort, but which may also contain extra information, as the date at which the unit is evaluated (most commonly with money). The middle category is Spread, which generalizes from the "random error" of experiments or the "variance" of statistics. Although Spread is usually conveyed by a number (either  $\pm$ , % or "factor of") it is not an ordinary quantity, for its own inexactness is not of the same sort as that of measurements. Methods to address Spread can be statistical data analysis, sensitivity analysis or Monte Carlo analysis possibly in combination with expert elicitation.

The remaining two qualifiers constitute the more qualitative side of the NUSAP expression. Assessment expresses qualitative judgments about the information. In the case of statistical tests, this might be the significance level; in the case of numerical estimates for policy purposes, it might be the qualifier "optimistic" or "pessimistic". In some experimental fields, information is given with two  $\pm$  terms, of which the first is the spread, or random error, and the second is the "systematic error" which must be estimated on the basis of the history of the measurement, and which corresponds to our assessment. It might be thought that the "systematic error" must always be less than the "experimental error", or else the stated "error bar" would be meaningless or misleading. But the "systematic error" can be well estimated only in retrospect, and then it can give surprises.

Finally there is P for Pedigree, which conveys an evaluative account of the production process of information, and indicates different aspects of the underpinning of the numbers and scientific status of the knowledge used. Pedigree is expressed by means of a set of pedigree criteria to assess these different aspects. Assessment of pedigree involves qualitative expert judgment. To minimize arbitrariness and subjectivity in measuring strength, a pedigree matrix is used to code qualitative expert judgments for each criterion into a discrete numeral scale from 0 (weak) to 4 (strong) with linguistic descriptions (modes) of each level on the scale. Each special sort of information has its own aspects that are key to its pedigree, so different pedigree matrices using different pedigree criteria can be used to qualify different sorts of information. Pedigree assessment can be further extended to also address societal dimensions of uncertainty, using criteria addressing different types of value loading, quality of problem frames etc. (Craye et al, 2005, Kloprogge et al., 2005, Corral, 2000).

NUSAP provides insight on two independent properties related to uncertainty in numbers, namely spread and strength. Spread expresses inexactness whereas strength expresses the methodological and epistemological limitations of the underlying knowledge base. The two metrics can be combined in a Diagnostic Diagram mapping strength of for instance model parameters and sensitivity of model outcome to spread in these model parameters. Neither spread alone nor strength alone is a sufficient measure for quality. Robustness of model output to parameter strength could be good even if

parameter strength is low, if the spread in that parameter has a negligible effect on model outputs. In this situation our ignorance of the true value of the parameter has no immediate consequences. Alternatively, model outputs can be robust against parameter spread even if its relative contribution to the total spread in the model is high provided that parameter strength is also high. In the latter case, the uncertainty in the model outcome adequately reflects the inherent irreducible uncertainty in the system represented by the model. Uncertainty then is a property of the modelled system and does not stem from imperfect knowledge on that system. Mapping components of the knowledge base in a diagnostic diagram thus reveals the weakest spots and helps in the setting of priorities for improvement.

Experiences so far, as reviewed in Van der Sluijs et al. (2005), have shown that the NUSAP method is applicable not only to relatively simple calculation schemes but also to complex models in a meaningful way and that it is useful to assess not only parameter uncertainty but also (model) assumptions. Especially when extended to include societal dimensions of uncertainty such as problem framing and value loadings, it promotes reflexivity and collective learning. The task of quality control in the knowledge base of complex and controversial (environmental) policy problems is a complicated one and the NUSAP method disciplines and supports this process by facilitating and structuring a creative reflexive process and in depth review of the limitations of a given knowledge base. NUSAP promotes to make explicit, and to systematically reflect upon, the various dimensions of uncertainty. It provides a diagnostic tool for assessing the robustness of a given knowledge base for policymaking and promotes criticism by clients and users of all sorts, expert and lay and will thereby support extended peer review processes.

#### ***4.2 The RIVM MNP Uncertainty Guidance***

After the fore-mentioned "De Kwaadsteniet affair" in 1999 a national and international review of the Netherlands Environmental Assessment Agency (RIVM/MNP) was undertaken. The auditors recommended that MNP should start a project to systematically address terminology, methodology, interpretation and communication of uncertainty. Following these recommendations, MNP commissioned Utrecht University to develop a practical guidance for uncertainty assessment and communication in environmental assessment studies. This was done in consultation with an international team of uncertainty experts. It was judged that the scope of the guidance system should extend beyond the mere quantitative assessment of uncertainties in model results per se, and should focus instead on the entire process of environmental assessment.

The RIVM/MNP Guidance for Uncertainty Assessment and Communication (Van der Sluijs et al., 2003, Petersen et al., 2003, Janssen et al, 2003, Van der Sluijs et al., 2004, Janssen et al., 2005) aims to facilitate systematic reflection on uncertainties throughout the whole scientific assessment process. It is structured around six foci: problem framing, stakeholder participation, indicator selection, appraisal of the knowledge base, mapping and assessment of relevant uncertainties, and reporting of the uncertainty information (see table 1).

<b>Foci</b>	<b>Key Issues</b>
<b>Problem Framing</b>	Other problem views; interwovenness with other problems; system boundaries; role of results in policy process; relation to previous assessments
<b>Involvement of Stakeholders</b>	Identifying stakeholders; their views and roles; controversies; mode of involvement
<b>Selection of Indicators</b>	Adequate backing for selection; alternative indicators; support for selection in science, society, and politics
<b>Appraisal of Knowledge Base</b>	Quality required; bottlenecks in available knowledge and methods; impact of bottlenecks on quality of results
<b>Mapping and Assessing Relevant Uncertainties</b>	Identification and prioritization of key uncertainties; choice of methods to assess these; assessing robustness of conclusions
<b>Reporting Uncertainty Information</b>	Context of reporting; robustness and clarity of main messages; policy implications of uncertainty; balanced and consistent representation in progressive disclosure of uncertainty information; traceability and adequate backing

Table 1 Foci and key issues in knowledge quality assessment. Transparent and effective uncertainty management in science-for-policy asks for systematic reflection and argued choice.

The Guidance assists scientific advisers in being as transparent as possible in their treatment of uncertainties. It explicitly addresses institutional aspects of knowledge development, openly deals with indeterminacy, ignorance, assumptions and value loadings. It thereby facilitates a profound societal debate and a negotiated management of risks. The Guidance is not set up as a protocol. Instead, it provides a heuristic that encourages self-evaluative systematization and reflexivity on pitfalls. It also provides some diagnostic help as to where uncertainty may occur and why. This can contribute to more conscious, explicit, argued, and well-documented choices.

Following a checklist approach (inspired by Risbey et al, 2005), the Guidance consists of a layered set of instruments (Mini-Checklist, Quicksan, and Detailed Guidance) with increasing level of detail and sophistication. It can be used by practitioners as a (self-)elicitation instrument or by project managers as a guiding instrument in problem framing and project design. Using the Mini-Checklist and Quicksan Questionnaire, the analyst can flag key issues that need further consideration. Depending on what is flagged as salient, the analyst is referred to specific sections in a separate Hints & Actions document and in the Detailed Guidance. Since the number of cross references between the documents comprising the Guidance is quite large, an interactive web application has been implemented on the MNP Intranet and a public version on the Internet ([www.mnp.nl/guidance](http://www.mnp.nl/guidance)). Additional advantages of this web application are that it offers a prioritized todo list of uncertainty assessment actions, it generates reports of sessions (traceability and documentation), and it enables internal and external review.

In order to facilitate communication about the different types of uncertainty that arise in scientific assessments, a uncertainty typology makes part of the Guidance. The typology is based on a conceptual framework that resulted from a process involving an

international group of uncertainty experts most of whom participated in developing or reviewing the Guidance (Walker et al, 2003). Uncertainty can be classified along the following dimensions: its “location” (where it occurs), its “level” (whether it can best be characterized as statistical uncertainty, scenario uncertainty or recognized ignorance) and its “nature” (whether uncertainty primarily stems from knowledge imperfection or is a direct consequence of inherent variability). In addition, the typology distinguishes the dimensions “qualification of knowledge base” (what are weak and strong parts in the assessment) and “value-ladenness of choices” (what biases may shape the assessment). The typology is presented as a matrix. This uncertainty matrix is used as an instrument for generating an overview of where one expects the most important (policy-relevant) uncertainties to be located (the first dimension), and how these can be further characterized in terms of four other uncertainty dimensions. The matrix can be used as a scanning tool to identify areas where a more elaborate uncertainty assessment is required. The different cells in the matrix are linked to available uncertainty assessment tools suitable for tackling that particular uncertainty type. These tools are described in a Tool Catalogue that aims to assist the analyst in choosing appropriate methods.

The Tool Catalogue provides practical (“how to”) information on state-of-the-art quantitative and qualitative uncertainty assessment techniques, including global sensitivity analysis, NUSAP, expert elicitation, scenario analysis, and model quality assessment. A brief description of each tool is given along with its goals, strengths and limitations, required resources, as well as guidelines for its use and warnings for typical pitfalls. It is supplemented by references to handbooks, software, example case studies, web resources, and experts. The tool catalogue is a “living document,” available on the web, to which new tools can be added.

## **5. Conclusion**

Complex environmental problems have characteristics that require a post-normal science approach in which uncertainty, assumptions and value loadings are subject to explicit and systematic analysis and communication. For this class of problems, Knowledge Quality Assessment should be at the heart of the science-society interface, in order to promote a better awareness of the limits of science in relation to the task of knowledge producers to provide a scientific basis for policy debate. In combination with a widening in focus from "reducing uncertainties" to "coping with untameable uncertainties and complexities", this can help to avoid misunderstandings and undue expectations of the role and competence of science in complex environmental problems.

Tools and approaches for Knowledge Quality Assessment such as NUSAP and the checklist based Guidance for Uncertainty Assessment and Communication of the Netherlands Environmental Assessment Agency, have now been developed tested and made available. NUSAP helps to systematically assess the technical, methodological and epistemic uncertainties in knowledge claims and helps to focus research efforts on the potentially most problematic parameters and assumptions in models, identifying at the same time specific weaknesses and biases in the knowledge base.

MNP's Uncertainty Guidance structures the tasks of uncertainty management, promotes reflection and forces deliberate choice on how uncertainties are handled. It helps to avoid pitfalls in the assessment and communication of uncertainty.

Similar to a patient information leaflet alerting the patient to risks and unsuitable uses of a medicine, Knowledge Quality Assessment enables the delivery of policy-relevant quantitative information together with the essential warnings on its limitations and pitfalls. It thereby promotes the responsible and effective use of the information in policy processes.

## **6. References**

Beck B (2002). Model evaluation and performance, in: El-Shaarawi AH and Piegorsch WW (eds.) *Encyclopedia of Environmetrics*, Vol. 3, John Wiley & Sons, Ltd, Chichester, pp 1275–1279

Beven K. (2002). Towards a coherent philosophy for modelling the environment. *Proc Roy Soc London, A*; 458 (2026) 2465–84.

D. Collingridge and C. Reeve, *Science Speaks to Power*, Printer, London, 1986.

Corral Quintana, SA (2000). *Una Metodología integrada de exploración y comprensión de los procesos de elaboración de políticas públicas*. Ph.D. thesis, University of La Laguna

Craye M., Van der Sluijs JP and Funtowicz S (2005). A reflexive approach to dealing with uncertainties in environmental health risk science and policy. *International Journal for Risk Assessment and Management* **5** (2), p. 216-236.

Douglas, M. (1966). *Purity and danger: An analysis of the concepts of pollution and taboo*. Routledge, London & New York.

Funtowicz SO and Ravetz JR (1990). *Uncertainty and quality in science for policy*, Dordrecht, Kluwer Academic Publishers.

Funtowicz SO and Ravetz JR (1993). Science for the Post-Normal Age. *Futures* 25: 735-755. Hornberger and Spear (1981)

Hellström T (1996) The Science-Policy dialogue in transformation: model-uncertainty and environmental policy. *Science and Public Policy*, **23** (2), 91-97.

Houghton, J., Jenkins, G. and Ephraums, J. (eds) (1990). *Climate Change, The IPCC Scientific Assessment*. Cambridge University Press, Cambridge.

International Geosphere Biosphere Programme (IGBP) (1992). *Reducing Uncertainties*. Royal Swedish Academy of Sciences, Stockholm, Sweden.

Janssen PHM, Petersen AC, Sluijs JP van der, Risbey JS and Ravetz JR (2003). *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Quickscan Hints & Actions List*. Bilthoven, The Netherlands: RIVM/MNP. ISBN 90-6960-105-2.

Janssen, P.H.M., Petersen, A.C., Van der Sluijs, J.P., Risbey, J., Ravetz, J.R. (2005). A guidance for assessing and communicating uncertainties. *Water science and technology*, **52** (6) 125-131.

Jasanoff S (1990). *The Fifth Branch, Science Advisers as Policy Makers*. Harvard: Harvard University Press.

- Keepin B and Wynne B. (1984). Technical analysis of IIASA energy scenarios. *Nature* 312: 691-695.
- Kloprogge, P., Van der Sluijs, J. and Petersen, A. (2005). *A method for the analysis of assumptions in assessments applied to two indicators in the fifth Dutch Environmental Outlook*, Research Report. Department of Science Technology and Society, Utrecht University, Utrecht, The Netherlands.
- MacKenzie D (1990), *Inventing Accuracy*, MIT Press, Cambridge, MA, 1990.
- Michaels D (2005), Doubt is their product. *Scientific American*, June 2005, 96-101
- Nakicenovic, N. and R. Swart (Eds.) (2000), *Emissions Scenarios. 2000*, Special Report of the Intergovernmental Panel on Climate Change, Cambridge: Cambridge University Press.
- Nowotny, H. (1999). The place of people in our knowledge. *European Review*, 7 (2), 247–262.
- Nowotny, H., Scott, P. and Gibbons, M. (2001). *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*, Polity Press, Cambridge.
- Oreskes N, Shrader-Frechette K and Belitz K (1994). Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science* 263: 641-646.
- Petersen AC, Janssen PHM, Sluijs JP van der, Risbey JS and Ravetz JR (2003). *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Mini-Checklist & Quickscan Questionnaire*. Bilthoven, The Netherlands: RIVM/MNP. ISBN 90-6960-105-1.
- Petersen AC (2005) Simulation Uncertainty and the Challenge of Post-Normal Science, in: Küppers G, Lenhard J, and Shinn T (eds). *Yearbook Sociology of the Sciences*, Springer.
- Refsgaard JC, Van der Sluijs JP, Brown J, and Van der Keur P (in press). A Framework For Dealing With Uncertainty Due To Model Structure Error, *Advances in Water Resources* (Prepublished online 5 January 2006).
- Saltelli, A. (2002). Sensitivity analysis for importance assessment, *Risk Analysis* 22 (3), pp. 579-590
- Saltelli A, Chan K and Scott EM (eds.) (2000). *Sensitivity Analysis*. John Wiley & Sons Publishers. Probability and Statistics Series.
- Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. (2004). *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, John Wiley & Sons publishers.
- Schneider SH (1997). Integrated Assessment Modeling of Global Climate Change: Transparent Rational Tool for Policy Making or Opaque Screen Hiding Value-Laden Assumptions. *Environmental Modeling and Assessment* 2, 229-248.
- S. Shackley and T. Skodvin, IPCC Gazing and the Interpretative Social Sciences, *Global Environmental Change*, 5 (3), 1995, p. 175-180.

- Smits, M. (2002). *Monsterbezwering, De culturele domesticatie van nieuwe technologie*, Boom, Amsterdam.
- Smits, M. (2004). *Taming Monsters. The cultural domestication of new technology*, University of Eindhoven, Eindhoven, The Netherlands.
- Stirling A (1999). *On science and precaution in the management of technological risk: Volume I - a synthesis report of case studies*. Seville: European Commission Institute for Prospective Technological Studies. Report number EUR 19056 EN.
- Stirling A (2001). *Inclusive deliberation and scientific expertise: precaution, diversity and transparency in the governance of risk*. PLA Notes 40: 67-71.
- Van Asselt MBA (2000). *Perspectives on uncertainty and risk: the PRIMA approach to decision support*. Dordrecht: Kluwer.
- J.C.M. van Eijndhoven and P. Groenewegen, The Construction of Expert Advice on Health Risks, *Social Studies of Science*, **21**, 1991, p. 257-278.
- Van der Sluijs, JP (1997). *Anchoring Amid Uncertainty: On the Management of Uncertainties in Risk Assessment of Anthropogenic Climate Change*, PhD thesis, Utrecht University, Utrecht, The Netherlands.
- Van der Sluijs, JP (2002). A way out of the credibility crisis of models used in integrated environmental assessment. *Futures*, **34**, 133–146.
- Van der Sluijs, JP (2005), Uncertainty as a monster in the science policy interface: four coping strategies. *Water science and technology*, **52** (6), 87–92.
- Sluijs JP van der, Eijndhoven JCM van, Shackley S and Wynne B (1998). Anchoring devices in science for policy, the case of consensus around climate sensitivity. *Social Studies of Science* **28** (2): 291-324.
- Van der Sluijs, J., Risbey, J., Kloprogge, P., Ravetz, J., Funtowicz, S., Corral Quintana, S., Guimaraes Pereira, A., De Marchi, B., Petersen, A., Janssen, P., Hoppe, R. and Huijs, S. (2003). *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Detailed Guidance*, Copernicus Institute for Sustainable Development, Utrecht University, and RIVM-MNP, Utrecht, The Netherlands, Available at [www.nusap.net](http://www.nusap.net).
- Van der Sluijs JP, Janssen PHM, Petersen AC, Kloprogge P, Risbey JS, Tuinstra W, Ravetz JR (2004), *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Tool Catalogue for Uncertainty Assessment*, Utrecht University & RIVM.
- Van der Sluijs JP, Craye M, Funtowicz S, Kloprogge P, Ravetz J, and Risbey J (2005) Combining Quantitative and Qualitative Measures of Uncertainty in Model based Environmental Assessment: the NUSAP System, *Risk Analysis*, **25** (2). p. 481-492
- Walker WE, Harremoës P, Rotmans J, Sluijs J P van der, Asselt MBA van, Janssen P and Kreyer von Krauss MP. Defining Uncertainty; A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment* **4**(1): 5-17.
- Weinberg A (1972). Science and Transcience. *Minerva*, **10**, 209-222.
- Williamson, P. (1994). Integrating Earth System Science. *Ambio*, **23** (1), 3.

Wynne B (1992), Uncertainty and Environmental Learning, in: *Global Environmental Change*, **2**, 111-127.

Yearley S (1996). *Sociology, Environmentalism, Globalization*, London: SAGE Publications.