28
POST-NORMAL SCIENCE

Roger Strand

Introduction

Post-normal science (PNS) is a critical concept originally developed to describe situations in which there are important or controversial public decision problems informed by an incomplete, uncertain or contested knowledge base. In this chapter, such problems will be called "post-normal problems". Writings on PNS are descriptive in the sense that they describe the characteristics of post-normal problems and how the relationship between science and policy changes in the face of such post-normal problems. They are frequently also normative in the sense that they interpret the situations created by post-normal problems and propose how the relationship between science and policy ought to change, above all in the direction of what has been called the democratisation of expertise.

This chapter introduces and explains the theoretical context of, and contribution by, PNS as a critical concept. The concept will be briefly compared to similar developments in post-empiricist history, philosophy and sociology of science before I end the chapter with a look at the practice dimension of PNS [see also Chapter 29]. The objective of this chapter is to indicate what the concept may have to offer ecological economists now and in the future.

Post-normal science: origin and overview

The philosophers Silvio Funtowicz and Jerome Ravetz (1990) coined the term "post-normal science". It has since caught on as an analytical concept but also as a label for a style of research practice. A growing community of researchers state that they are doing PNS, that is, producing knowledge or practising problem-solving strategies better suited for post-normal problems. These strategies are then seen as being logically implied, or at least legitimised, by analyses based on PNS as a critical concept. Often, the strategies include elements of public participation—"extension of the peer community"—and/or innovative approaches to the characterisation, communication and management of scientific uncertainty.

Research that claims to do PNS in practice, attends to a wide range of concrete topics and decision problems. However, the main focus is on public governance of the natural environment, sustainability and climate change. Figure 28.1 provides a rough indication of topics covered. This is a word cloud representing words in the title or abstract of explicitly "post-normal" research papers registered in Web of Science.

Figure 28.1 Post-normal science word cloud.

This reveals a considerable overlap between the topics of ecological economics and writings on PNS. That is no coincidence. After an initial focus mainly on technological risk, the seminal writings of Funtowicz and Ravetz (to be explained below) rapidly broadened their scope to the risks, uncertainties and complexities of environmental governance. Notably, Funtowicz and Ravetz (1994) themselves proposed that ecological economics is, or could become, a post-normal science. Silva and Teixeira (2011) performed a bibliometric analysis of all articles in the journal *Ecological Economics*, 1989–2009, and concluded "at least through the lens of EE [the journal], ecological economics has evolved towards a post-normal science" (ibid.: 849). Their conclusion is a matter of some contestation, because they employ a definition that includes much which is actually rejected by Funtowicz and Ravetz; as Spash (2013: 356) states, their claims "that ecological economics is now a post-normal science appear based upon the antithesis of the post-normal philosophy (e.g., the spread of mathematical formalism, abstract expert modelling, and low quality uncritical monetary quantification)". This indicates some confusion on the part of Silva and Teixeira as to what constitutes PNS. Still, as also noted by Spash (2012: 42–43), the concept of PNS has definitely played a role in the modern history of ecological economics.

The theory of PNS by Funtowicz and Ravetz

Three types of risk assessment

In 1985, five years before the first mentioning of "post-normal science", Funtowicz and Ravetz published a paper called "Three Types of Risk Assessment: A Methodological Analysis" (Funtowicz and Ravetz, 1985). The three types were:

1 applied science;
2 technical consultancy (to be called "professional consultancy" in later works);
3 total-environmental assessment (to be called "post-normal science" in later works).

They explained these concepts in a diagram reproduced here as Figure 28.2 and discussed further below.

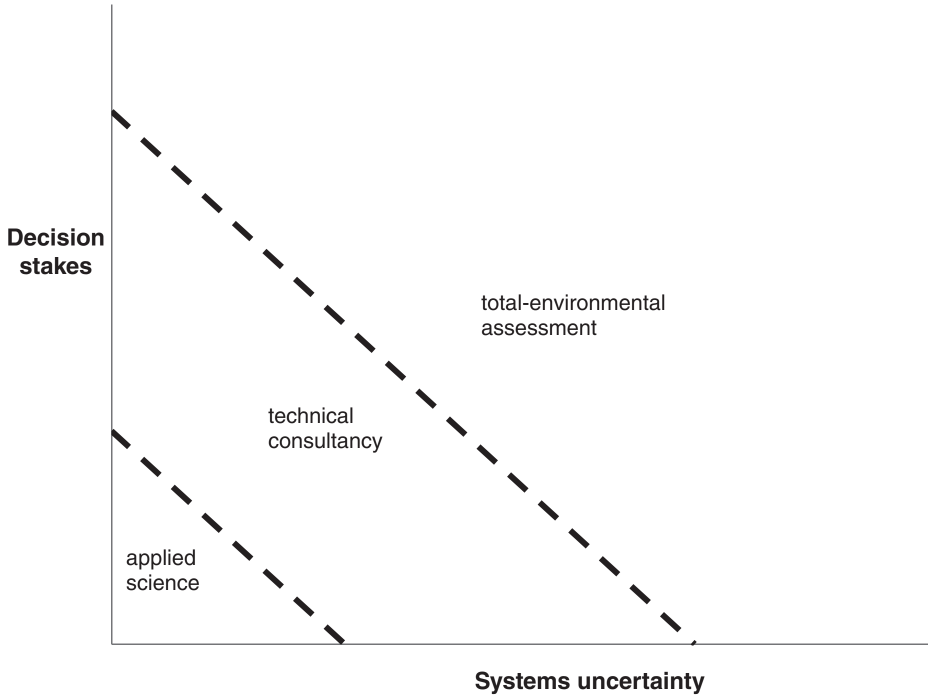

Figure 28.2 The original diagram of three types of risk assessment.
Source: redrawn after Funtowicz & Ravetz (1985).
Note: "Total-environmental assessment" would later be relabelled as "post-normal science".

The typology is better understood as a heuristic than a set of exact definitions. The underlying idea is that applied science works well only under conditions of relatively low uncertainty and decision stakes. For instance, one might be faced with technological risks (or rather hazards) for which there is no historical record that validly warrant probability estimates, or that result from interactions in systems that are too complex or too poorly characterised to be reliably captured in a model. This might be a trivial problem if the risk or hazard is quite small and there can be a cumulative learning cycle based on repeated failures, accidents and improvements. Ever more decision problems are, however, connected to issues of a larger magnitude and with higher decision stakes. The risk of major nuclear accidents is one example.

Decision stakes are rightly seen as a subjective property in the sense that they cannot be assessed without applying value judgements. Not by logical necessity but as a matter of fact, high decision stakes also tend to be connected to disagreements about the stakes and the underlying value judgements between different stakeholders and affected parties. Funtowicz and Ravetz (1985) note that there is a type of practical risk assessment (and management) taking place that may be heavily informed by science but still different from it: the type of work being done by engineers, medical doctors and other professional consultants who combine scientific training with practical skills and experience. This style of work may tackle higher uncertainties along the facts and values dimensions, in part because of a more flexible and adaptive work mode than allowed by the standards for scientific rigour, and in part because they reduce the uncertainty in the value dimension by loyalty. Those undertaking such work, by adopting the values and goals of their client—patient, company, governmental organisation—may reduce the scope of relevant concerns and accordingly exclude relevant system uncertainties.

The partly descriptive, partly normative claim of Funtowicz and Ravetz is that there is a growing class of decision problems in modern societies that belong to the top-right sector of Figure 28.2, characterised by high stakes and high system uncertainties. Energy supply is one of their examples. The analysis accepts that societal choices about energy supply can be informed by numerous rigorous facts and analyses. However, the choices cannot be reduced to questions that can be answered by applied science or technical consultancy, because the content of the choices themselves depends upon: the framing of the decision alternatives; the value judgements on the relevance and importance of the various risks and benefits; the time frame; the delimitation of affected parties; and similar factors. In other words, the definition of the issue itself is (or should be) a matter of continuous value-laden deliberations and negotiations. This is so for two reasons, regarded as inextricably linked concerns: democracy and the quality and effectiveness of knowledge. The recommended process of ongoing deliberation—framing and reframing the issue—is what Funtowicz and Ravetz originally call the "total-environmental assessment", because it cannot be reduced to a technical exercise, or a political decision. Potentially, any kind of descriptive or normative input to the process may be relevant.

The concern for democracy is readily justified if interpreted on the grounds of justice. We may imagine an industry that produces a certain product for profit and at the same time exposes its workforce and its surroundings to emissions with unclear consequences for human health and the environment. The industry itself may hire consultants to investigate the risks as defined from their perspective, for instance in terms of current occupational health standards, and with a view to liabilities and reputational risks. However, affected parties might be dissatisfied with this scope, and insist on the need for an independent inquiry where they have a say in the choice of risks to consider and the standards employed for comparison. In academic communities, such as that of ecological economics, as well as in the regulatory set-up of most modern states, there is little need to argue why these affected parties may rightly make such claims.

Clearly the concern for democratic principles is often neglected in reality—war, torture, oppression, exploitation. Risk assessment and management may often take place as an expert closed-down exercise within a narrow frame, excluding the concerns of affected parties, and without much transparency. Within the theory of PNS this is also seen as a quality problem. If the risk assessment fails to address relevant concerns of legitimate stakeholders, it is not fit for the purpose. The definition of the purpose is a matter of the framing of the problem, which is a democratic—deliberative, inclusive and participatory—concern, and consequently the judgement on quality is also, in this particular sense, a democratic concern. The scope of the risk assessment (or indeed any other production of knowledge relevant for a decision problem) is a function of the over-all decisions on problem framing, but typically also a myriad of smaller or larger methodological decisions, such as: the exact choice of variables, metrics and indicators, resolution of spatial and temporal grids, inclusion and exclusion of data variables, and so on. All of these decisions potentially involve value judgements of a partially political character. A key concept in PNS is therefore the extension of the peer community, defined as the inclusion for the production of knowledge of 'other' (non-expert) actors in methodological decisions and quality judgements.

A frequent criticism of the argument just presented is that it is too idealist in the sense of being too optimistic about the prospects of involving non-experts in processes of knowledge production. Believers in 'objective science' typically argue that stakeholders might be guided by their own interests and accordingly subjective and biased. Furthermore, following this line of argument, stakeholders may lack the knowledge required to make competent contributions to the deliberations upon methods and knowledge quality. One can find different stances towards this objection in the writings on PNS. One line of response is to see the objection as a question open to empirical inquiry. Wikipedia, for example, is an online encyclopaedia with a virtually open peer community and with a relatively light structure to organise and manage author inputs. In this sense it can be seen as an experiment of PNS that can help with casting light on the conditions under which an extension of the peer community appears to work well (or not so well). The response closest to the core of the analysis provided by Funtowicz and Ravetz, however, is to remind the critics of the direness of many post-normal situations. PNS is not about piece-meal improvements in the methods of risk assessment—enriching an already fine scientific risk assessment with valuable facts and values from the citizens and thereby also improving its democratic legitimacy. The theory of PNS focuses on the social reality of controversy, conflict and erosion of trust between experts and non-experts. Extension of the peer community is not an extra to improve chances of success but a necessity. It is required to address the implication of acknowledging that—under conditions of high uncertainty, high decision stakes and values in dispute—a number of dichotomies, central to the imaginary of modern societies, simply fail to match reality. A range of dichotomies and their divisions cannot be justified in an absolute sense: facts and values; science and politics; expertise and non-expertise. So in cases where a strict separation between "objective" expertise and "subjective" non-expertise is seen to be impossible or dysfunctional—in the sense that controversies are unresolved, trust is absent and important decisions are stalled—refusing to extend the peer community is unjustified (Funtowicz and Strand, 2007).

The theory of PNS is less about providing recipes for success and more about understanding why success recipes and technical fixes often prove inadequate and why the desire for technical fixes is part of the dysfunctionality encountered in post-normal situations. In this sense, the word science in PNS is misleading because it indicates a focus on science alone. The scope of this theoretical perspective is more correctly seen as the interface between science and (modern) society, and in particular the institutions and culture(s) that govern the interface between science and policy. Rather than a technical solution, PNS advocates institutional change and a change of individual and group mentality.

Normal and post-normal science

Why then, should we call it "post-normal science" rather than total-environmental assessment? In addition to being catchy, the term "post-normal science" offers a useful contrast to "normal science", which is a central concept in the philosophy of science associated with Thomas Kuhn in his *The Structure of Scientific Revolutions* (1962). There are many good alternatives to Kuhn's description of scientific practice—indeed, Ravetz (1971) himself offered a much more detailed description and explanation of the workings of science—so the main advantage of Kuhn's work is its wide dissemination. *Structure* is widely cited and has sold more than one million copies, and (even amongst those who have never read or heard of the book) the concepts of normal science, paradigm, paradigm shift, scientific revolution and incommensurability have made their way into popular as well as academic culture. Didactically, the comparison of Kuhn's normal science with PNS allows a set of easily explained contrasts, as shown in Table 28.1.

Didactics are important in this respect. To my knowledge, there are no educational programmes entirely devoted to PNS. It has been included as one of several theoretical frameworks in some study programmes within interdisciplinary environmental science (including ecological economics). A very important role for PNS, however, is its inclusion as a critical concept in normal science education. In the typology of research related to PNS provided by Turnpenny et al. (2011), one of the main genres of PNS is that of being a challenge for scientists. One may argue with Kuhn that science education is designed to socialise students

Table 28.1 Normal science and post-normal science compared

Normal science	*Post-normal problems*	*Post-normal science maxims*
.. takes time	.. call for urgent decisions	Focus on knowledge quality in terms of fitness for purpose!
.. replaces ignorance and uncertainty with certainty	.. may have irreducible uncertainties	Communicate and manage the uncertainties (rather than trying in vain to eliminate them)!
.. focuses on simple and idealized systems	.. emerge in complex systems (non-linear; nature-culture)	Include a multitude of perspectives!
.. is puzzle-solving	.. have high stakes	Extend the peer communities!
.. is practiced within a paradigm of agreed methods and shared values	.. may involve both facts and values in dispute	Acknowledge that methods and facts are value-laden!

into the normal science paradigms, teaching them not only skills, theories and other forms of positive knowledge, but also discipline. That is, disciplining them to focus their cognitive efforts inside the borders of the paradigm and discouraging reflection and questions critical to or outside of the paradigm.

In Kuhn's analysis, this state of affairs is simply a matter of fact, an inevitable effect of the organisation of normal science. For his interlocutors Karl Popper (1963) and Paul Feyerabend (1975), the repressive and dogmatic character of normal science was something that could and should be fought by open criticism. They thought this could be achieved either by cherishing the classic scientific ideals of fallibilism and humility (Popper) or opening up established science to what currently falls outside of it (Feyerabend). The role of PNS, in normal science education, can then be thought of as akin to Feyerabend's philosophy, trying to open up scientific minds by showing them how important sources of uncertainty and complexity remain invisible from within their paradigms.

PNS and the science policy interface

The theory of PNS places itself among several other theoretical developments that criticise empiricist and objectivist portrayals of scientific practice and scientific knowledge. There are several sources and origins of these strands of thought. In non-Anglophone philosophy of science, post-empiricist/post-positivist thinking has been around for a long time (since Husserl and Heidegger and the Frankfurt School in Germany; with Bachelard, Canguilhem and Foucault in France; Ortega y Gasset in Spain; and in various shapes in Russian philosophy). In the 1960s, neo-Marxist influence was present in the critical science/radical science movements with which PNS shares its theoretical and practical interests. Similarly there is cross-over of interest with early science studies and "Science, Technology and Society" (STS), in the 1960s (Sardar and van Loon, 2011). STS later became more theoretically sophisticated and institutionalised into "Science and Technology Studies", a research field of its own, with study programmes and research departments around the world. Within (and around) STS, the typical topics of PNS—strong uncertainty, complexity, the importance of the framing of issues and the value of knowledge outside of normal science—are by no means underrepresented in theoretical and empirical work. When I teach courses on PNS I always include work by STS researchers such as Sheila Jasanoff and Brian Wynne, and frequently also the actor-network theory with reference to Bruno Latour or Michel Callon.

Furthermore, one can encounter critical concepts in STS and the history, philosophy and sociology of science that share the heuristic and didactic value and some of the content of PNS. One prominent example is the "Mode 1/Mode 2" distinction made by Gibbons et al. (1994). Mode 1 is a mode of knowledge production quite close to normal science: disciplinary, taking place within academic institutions, motivated by curiosity at the research forefront as defined by the discipline itself, and subject to quality standards defined and sanctioned by the disciplinary peer community. Mode 2, on the other hand, is conceived as transdisciplinary problem-solving by scientific researchers outside of academia, for example in industry and the service sector, where quality is a matter of solving the problem to the satisfaction of the employer or client rather than meeting disciplinary standards. In their original work, Gibbons et al. (1994) hardly touched the aspect of civil society and the democratisation of expertise. However, their follow-up (Nowotny et al., 2001), does cover the issue of accountability and interaction with broader society, and develops their concept of "socially robust knowledge", thereby providing a suite of concepts that are similar to professional consultancy and PNS.

In sum, there is no strict demarcation between the theory of PNS and these areas of STS and related work. PNS is also a moving target, depending upon which authors are read, and which of their texts. The texts by Funtowicz and Ravetz also, unsurprisingly, developed throughout the 1980s, 1990s and 2000s, departing from a focus on exceptionally complex cases of technological risk and towards a general diagnosis of late modern societies. As a rule of thumb, PNS has tended to place itself somewhere between scientific realism and the full-blown social constructivism of, say, early Sociology of Scientific Knowledge (SSK). This is not so extraordinary anymore either, because several of the radical SSK researchers later modified their positions to more moderate constructivist positions in the course of the 1990s. One might perhaps conclude that a weak, overlapping consensus on moderate constructivist positions is gradually emerging from the diversity of research on science and technology (Sardar and van Loon, 2011). Perhaps the main difference between PNS and STS is that the former has been much more engaged in the development of concrete problem-solving practices; taking part in practice and not just studying it.

Future directions

As noted above, the Web of Science (as of 1 August 2015) counts 259 research papers with the words 'post normal' or 'post-normal' in the title or abstract. Ten per cent of these papers were published in the 1990s. Most of them can be classified as philosophical or conceptual in their outlook, and most of them were published either by Funtowicz, Ravetz or their close collaborators. In the 2000s, this changed. More and more of the post-normal research papers (in the sense that they call themselves post-normal) are predominantly empirical and practical in their orientation, claiming to apply or do PNS in some sense, in particular in order to improve some aspect of environmental governance (e.g., concerning climate change, ecosystem or natural resource management, nature preservation).

This change is quite in line with the recommendations of the early writings of Funtowicz and Ravetz. Indeed, in their much cited book *Uncertainty and Quality in Science for Policy,* Funtowicz and Ravetz (1990) developed the NUSAP system, which is a flexible notational system for the characterisation, communication and management of scientific uncertainty. NUSAP is an acronym for numeral, unit, spread, assessment and pedigree [see also Chapter 29]. Normal statistical notation can provide options for NUSA (no P)—for example, standard deviation would be a measure for spread while p-values or confidence levels provide information on assessment. Pedigree, however, is the post-normal or mildly constructivist innovation in this system. The pedigree of a number or a fact can be elicited and communicated through tailored sets of descriptors, for instance on the level of expert agreement, quality of data, the use of models, and the level of maturity of underlying scientific theory. Some of those conducting PNS apply the NUSAP system or similar methodological developments along that line (for instance Petersen et al., 2011). The other major type of work on PNS consists of attempts at extending the peer community by including local knowledge, stakeholder involvement and/or citizen participation in some way. The types may of course be (and often are) combined in the same study or exercise.

Extending the peer community means moving towards transdisciplinarity in the sense of including non-academic knowledge sources and methods. The use of Web of Science in this chapter as a source of data on PNS can therefore be criticised as being far too narrow. Indeed, while Web of Science provides only 259 "hits", a similar search on Google Scholar gives 11,000. Furthermore, if one wanted to write a handbook on PNS, relevant methods and experiences go beyond those who use the label or reference PNS in their writings. From an (explicit) PNS perspective, practice is becoming increasingly post-normal, regardless of whether the actors are aware of that label or the underlying theory. For example, one may interpret the phenomenon of wikis (such as Wikipedia.org) as typical of PNS in the sense of being transdisciplinary knowledge production that essentially relies on a wide extension of the peer community. For some proponents of the theory of PNS, it is this broader trend that is important and not the label of "post-normal science" itself. Others are more active and eager to promote PNS as an identity marker in order to develop and consolidate a community of researchers and practitioners at the science–policy interface. If I were to make guesses about the future, however, I would not expect a similar development as that of STS, that is, going from a quite heterogeneous community who shared a broad concern to a disciplinary project of building academic institutions. After all, the core of PNS is that it is not normal science. It is unlikely that a majority of its proponents will forget this core.

Concluding remarks

Funtowicz and Ravetz (1994) proposed in their paper "The Worth of a Songbird: Ecological Economics as a Post-normal Science" that the main challenge for ecological economics is not to provide yet another attempt at a technical solution to reduce and eliminate the descriptive and normative uncertainties and complexities involved in human interaction with Nature, but rather to identify, understand and manage them. In that sense, they envisioned ecological economics as an opportunity not to become just another normal science. One may discuss the extent to which ecological economics in fact is post-normal. On one hand, we may observe that with the exception of *Futures*, no journal has published more papers on PNS than *Ecological Economics*. However, as noted earlier, while Silva and Teixeira (2011) concluded that ecological economics is indeed becoming post-normal, their analysis appears to be based upon a different and perhaps somewhat superficial concept of PNS. Furthermore, the extent to which publications in *Ecological Economics* actually pay attention to the messages of PNS is highly debatable. For example, Spash (2007) provides an example applying the framing of PNS to critically appraise the Stern review on climate change, but the majority of climate research published in *Ecological Economics* appears to adopt a mainstream neoclassical economic framing (Anderson and M'Gonigle, 2012), and thus remains within normal science.

We have argued above that PNS is a term that is used in a variety of ways; and while one can meaningfully say that one is "doing post-normal science" for instance by applying methods of strong uncertainty management or extending peer communities, describing PNS as a type of science or research is inappropriate. PNS is a term, a convenient label for a description and a diagnosis of the science–society interface, and for providing some ideas of how to improve that interface. In order to say something more specific about what PNS "is", the best option is to employ the initial term "total-environmental assessment", meaning the ongoing negotiation, framing and reframing of decision issues. Along this line of argument it does not make much sense to ask if ecological economics is "a post-normal science". PNS is not science; if ecological economics pretends to be a scientific discipline, it is not PNS.

What makes sense, however, is to analyse how actors and institutions of ecological economics envision the interface between their own production of knowledge and the society (and nature) for and in which it is produced, notably institutions of policy and governance. To the extent that ecological economics offers, or wishes to offer, venues at the science–policy interface for keeping fundamental philosophical debates open—what are problems? what may count as solutions? who should judge and assess? what are the quality criteria?— we can call it a "post-normal science". Rather than a set of techniques and practices, at least some of us who identify with the label "post-normal", identify with a certain engagement, mentality and attitude. A main dimension of that attitude is reflexivity and humility (Strand and Cañellas-Boltà, 2006). This, I believe, poses a challenge. Clearly this challenges orthodox economists who may have been socialised into the belief that they are producing 'Truth in order to speak to Power'. However, the challenge may be found equally demanding for ecological economists of all heterodox colours who have had to build such hard strategies against the hegemony of orthodox thought that they can only rarely allow themselves to doubt their own epistemic and moral virtues. This, however, is at the core of PNS: the willingness to open up one's cherished facts and values to inclusive processes of total-environmental assessment.

Acknowledgements

I thank the editor, Clive Spash, for comments that significantly improved this contribution. I also wish to express my gratitude to Ina Hannestad Nygaard for editing text and figures.

Key further reading cited

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

Funtowicz, S., Ravetz, J. (1994). The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics* 10, 197–207.

Funtowicz, S., Strand, R. (2007). Models of Science and Policy, in Traavik, T. and Lim, L.C. (eds.), *Biosafety First: Holistic Approaches to Risk and Uncertainty in Genetic Engineering and Genetically Modified Organisms* (pp. 263–278). Tapir Academic Press, Trondheim.

Sardar, Z. and van Loon, B. (2011). *Introducing Philosophy of Science: A Graphic Guide*. London: Icon Books.

Turnpenny, J., Jones, M., Lorenzoni, I. (2011). Where now for post-normal science? A critical review of its development, definitions, and uses. *Science, Technology, & Human Values* 36, 287–306.

Other references cited

Anderson, B., M'Gonigle, M. (2012). Does ecological economics have a future? Contradiction and reinvention in the age of climate change. *Ecological Economics* 84 (December), 37–48.

Funtowicz, S., Ravetz, J. (1985) Three types of risk assessment: a methodological analysis, in: C. Whipple, V. T. Covello (eds): *Risk Analysis in the Private Sector* (pp. 217–231), New York and London: Plenum Press.

Feyerabend, P. (1975) *Against Method*. London: Verso.

Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, M. Trow (1994). *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

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

Nowotny, H., Scott, P., Gibbons, M. (2001). *Re-Thinking Science. Knowledge and the Public in an Age of Uncertainty*. Cambridge: Polity Press.

Petersen, A.C., A. Cath, M. Hage, E. Kunseler, J.P. van der Sluijs (2011). Post-normal science in practice at the Netherlands Environmental Assessment Agency. *Science, Technology, & Human Values* 36, 362–388.

Popper, K. (1963) *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Routledge.

Ravetz, J. (1971), *Scientific Knowledge and Its Social Problems*. Oxford: Clarendon Press.

Silva, C.M., Teixeira, A. (2011). A bibliometric account of the evolution of EE in the last two decades: Is ecological economics (becoming) a post-normal science? *Ecological Economics* 70, 849–862.

Spash, C.L. (2007). The economics of climate change impacts à la Stern: Novel and nuanced or rhetorically restricted? *Ecological Economics* 63(4), 706–713.

Spash, C.L. (2012). New foundations for ecological economics. *Ecological Economics* 77(May), 36–47.

Spash, C. L. (2013). The shallow or the deep ecological economics movement? *Ecological Economics* 93(September), 351–362.

Strand, R., Cañellas-Boltà, S. (2006). Reflexivity and Modesty in the Application of Complexity Theory. In A. Guimarães Pereira, S. Vaz, S. Tognetti (eds.), *Interfaces between Science and Society* (pp. 100–117). Sheffield, UK: Greenleaf Publishing.

29

THE NUSAP APPROACH TO UNCERTAINTY APPRAISAL AND COMMUNICATION

Jeroen P. van der Sluijs

Introduction

The knowledge base available for decision-making on contemporary environmental and sustainability issues is often characterised by an imperfect understanding of the complex systems involved. Decisions will need to be made before conclusive scientific evidence is available, while at the same time the potential error costs of wrong decisions can be huge. This societal context of knowledge production and use for decision-making and risk management implies an urgent need for explicit appraisal and consideration of all dimensions of scientific uncertainty (Funtwicz and Ravetz, 1990, 1993; Van der Sluijs, 2002; Van der Sluijs et al., 2008, Saltelli et al., 2013); [see also Chapters 26–28].

The transdisciplinary nature of science for sustainability poses additional requirements with regard to the systematic analysis, documentation and communication of uncertainty. When quantitative information is produced in one disciplinary context and used in another, we often see that important caveats tend to be ignored, uncertainties compressed and numbers used at face value (Wynne, 1992). Knowledge utilisation for sustainability issues 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).

The past record of science for policy has shown that omitting uncertainty assessment and communication can undermine public trust in the science (e.g. Keepin and Wynne, 1984). An example concerns the Netherlands National Institute for Public Health and the Environment (RIVM). In early 1999, De Kwaadsteniet, a senior statistician, accused the RIVM of "lies and deceit" in their State of the Environment Reports and Environmental Outlooks. In a quality newspaper (*Trouw*) he criticised RIVM for basing their studies on the 'virtual reality' of poorly validated computer models, while RIVM presented these results as point values with unwarranted significant digits and without elaborating the uncertainties. A vehement public debate was triggered on the credibility and reliability of environmental numbers and models. The case got front page and prime time coverage in the mass media and led to a debate in the Netherlands parliament (Van der Sluijs, 2002). The RIVM went through a learning process that led to the development of guidance for uncertainty management in the institute (Van der Sluijs et al., 2008; Petersen et al., 2011).

In this chapter I will explain how such incidents have led to a revision of the approach to uncertainty. The next section outlines how the understanding of uncertainty has changed and the implications this has for policy. This background explains why new transdisciplinary methods for evaluating uncertainty have been developed. The chapter then focuses on the Numeral, Unit, Spread, Assessment and Pedigree (NUSAP) framework and explains its key features. Pedigree is given particular attention as the innovative aspect of this post-normal science system [see Chapter 28].

Understanding uncertainty

In the early phase of its development, the field of uncertainty analysis mainly evolved around mathematical statistical methods such as sensitivity analysis and Monte Carlo techniques for the assessment of error propagation in model calculations. These tools address quantitative dimensions of uncertainty using sophisticated algorithms (Saltelli et al., 2000, 2008). Although these quantitative techniques are essential in any uncertainty analysis, they provide only a partial insight into what usually is a very complex mass of uncertainties involving technical, methodological, epistemological and societal dimensions. 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. In their combination, the quantitative and qualitative methods provide a richer diagnosis of uncertainty than each of these methods alone.

Over the past decades, an increasing body of conceptual and theoretical work in the field of uncertainty management has been compiled. Key insights from the field include:

- Uncertainty is partly socially constructed and its assessment always involves subjective judgement;
- More research does not necessarily reduce uncertainty, it often reveals unforeseen complexities;
- Some uncertainty is irreducible (intrinsic or practically);
- High quality scientific knowledge for policy making does not require low uncertainty;
- Uncertainty is a multi-dimensional concept involving quantitative (technical inexactness) and qualitative dimensions (i.e., methodological unreliability, epistemological ignorance and societally limited robustness), and it can manifest itself at different locations (e.g., context, indicator choice, model structure, parameters and data).
- In problems that are characterised by high systems uncertainties, ignorance, and high decision stakes the qualitative dimensions of uncertainty may well dominate the quantitative dimensions.

Most of present day methodologies and practices for addressing uncertainty focus exclusively on quantitative uncertainty in model parameters and input data. Methods to address qualitative and societal dimensions of uncertainty are absent or in their early stage of development. Uncertainties relating to model structure, model assumptions and model context are largely ignored.

Scientists, policymakers and stakeholders now widely hold that uncertainty management in environmental assessment is essential. However, in the practice of uncertainty management there is little appreciation for the fact that uncertainty is more than a number. There are many different dimensions of uncertainty and there is a lack of understanding about their different characteristics, relevance and relative importance. Even within the different fields of decision support (such as integrated assessment, environmental risk assessment, environmental impact assessment, policy analysis, engineering risk analysis and cost-benefit analysis), there is neither a commonly shared terminology nor agreement on a generic typology of uncertainties (Walker et al., 2003).

A better understanding of the various dimensions of uncertainty is needed in order to provide an improved theoretical foundation for uncertainty assessment. Improved conceptualisation of uncertainty is desirable for a number of reasons. First, it will aid better communication amongst the many disciplines involved. In the current situation, different analysts use different terms for the same kinds of uncertainty, and some use the same term to refer to different kinds. This makes it difficult for those who have not participated in the actual work to understand what has been done. Improved conceptualisation of uncertainty will further provide for better communication among scientists, policymakers and stakeholders. A common belief is that policymakers expect scientists to provide certainties and hence dislike uncertainty in the scientific knowledge base. However, uncertainty is a fact of life and a better understanding of its key dimensions and their implications for policy choices would be likely to lead to more trust in the scientists providing decision support, and ultimately to better policies. Finally, a better understanding of the different dimensions of uncertainty and their potential impact on the relevant policy issues at hand would help in identifying and prioritising effective and efficient research and development activities for improving the knowledge base.

Van der Sluijs (1997) concludes that in the practice of uncertainty management in integrated modelling of climate change, major gaps exist in the systematic analysis of unreliability of the knowledge about input data, model parameters and model assumptions, and also in the analysis of uncertainty about model structure. A major obstacle is that tools for assessing these types of uncertainty and how these might affect the outcomes of assessments, are either not available or in their early stage of development. Only recently have new tools been developed that focus on the qualitative (methodological and epistemological) dimensions of uncertainty using methods of expert elicitation, quality assistance checklists (Risbey et al., 2005), Pedigree analysis (using multiple criteria [Chapter 30]) assessing the strength of various underpinning components of the knowledge base by self-review, peer review or extended peer review (Funtowicz and Ravetz, 1990; Van der Sluijs et al., 2005a, 2005b, 2005c), and methods for the systematic identification and characterisation of critical assumptions in models (Kloprogge et al., 2011).

Addressing the multiple dimensions of uncertainty

Whereas quantitative methods are well developed, standardised and supported by handbooks (Morgan and Henrion, 1990, Saltelli et al., 2000, 2008) and software (@Risk, Crystal ball, Simlab, Analytica), qualitative and multi-dimensional methods have been demonstrated and tested but have not yet been widely disseminated and adopted. Multi-dimensional methods are those that address qualitative and quantitative aspects in a coherent way. They do so by assessing the technical (inexactness), methodological (unreliability), societal (social robustness), and epistemological (border with ignorance) dimensions of uncertainty, as shown in Table 29.1.

Numeral, Unit, Spread, Assessment, Pedigree (NUSAP)

NUSAP is a notational system, proposed in the context of post-normal science by Funtowicz and Ravetz (1990), which aims to provide an analysis and diagnosis of uncertainty in science for policy. The NUSAP system structures the systematic appraisal and communication of the various dimensions of uncertainty. It provides an heuristic for good practice addressing uncertainty in quantitative information. NUSAP has extended the statistical approach to uncertainty with methodological and epistemological dimensions by adding expert judgement of reliability (Assessment) and systematic multi-criteria evaluation of the underpinning of numbers (Pedigree). Examples of Pedigree criteria are empirical basis, methodological rigour, theoretical understanding, degree of validation and peer acceptance.

Table 29.1 Dimensions of uncertainty

Dimension	*Type*	*Can stem from or can be produced by*
Technical	Inexactness	*Intrinsic uncertainty:* Variability; stochasticity; heterogeneity *Technical limitations:* Error bars, ranges, variance; Resolution error (spatial, temporal); Aggregation error; Linguistic imprecision, unclear definitions
Methodological	Unreliability	*Limited internal strength of the knowledge base in*: Use of proxies; Empirical basis; Theoretical understanding; Methodological rigour (including management of anomalies); Validation
Epistemological	Ignorance	*Limited theoretical understanding* *System indeterminacy*: Open-endedness of system under study; Chaotic behaviour *Intrinsic unknowability with active ignorance*: Model fixes for reasons understood; Limited domain of validity of assumptions; Limited domains of applicability of functional relations; Numerical error; Surprises type A (some awareness of possibility exists) *Intrinsic unknowability with passive ignorance*: Bugs (software error, hardware error, typos); Model fixes for reasons not understood; Surprises type B (no awareness of possibility)
Societal	Limited social robustness	*Limited external strength of the knowledgebase in*: Completeness of set of relevant aspects; Exploration of rival problem framings; Management of dissent; Extended peer acceptance/stakeholder involvement; Transparency; Accessibility *Bias/Value ladenness:* Value laden assumptions; Motivational bias (interests, incentives); Disciplinary bias; Cultural bias; Choice of (modelling) approach (e.g. bottom up, top down); Subjective judgement

The NUSAP framework provides a means for synthesis and integration of findings on each of these dimensions, combining formal Monte Carlo and mathematical sensitivity analysis techniques with systematic qualitative uncertainty assessment. NUSAP enables providers and users of knowledge to be clear and transparent about its various uncertainties. This promotes critical reflection on the strengths and weaknesses of the underlying knowledge base by users of all sorts (e.g., experts, lay public) and thereby supports an extended peer review process. It aims to provide those who produce, use and are affected by policy-relevant knowledge with a set of diagnostic tools for a critical self-awareness of their engagement with that knowledge.

NUSAP extends the statistical approach to uncertainty (inexactness) by incorporating the methodological (unreliability) and epistemological (ignorance) dimensions using expert judgement of reliability (Assessment) and systematic multi-criteria evaluation of the process by which numbers are produced (Pedigree). Numbers are provided with a separate qualification for each dimension of uncertainty, allowing nuances of meaning about quantities to be conveyed concisely and clearly, to a degree that is quite impossible with reliance on statistical methods alone. NUSAP captures both quantitative and qualitative dimensions of uncertainty and enables one to display these in a standardised and self-explanatory way. The basic idea is to qualify quantities using the five aspects of the NUSAP system: (i) Numeral, (ii) Unit, (iii) Spread, (iv) Assessment and (v) Pedigree. Each of these dimension, or numeric qualifiers, is discussed in turn.

First is the Numeral, which is normally an ordinary number, but, when appropriate, 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 is the Unit, which may be of the conventional sort, but which may also contain extra information, such as the date on which the Unit is evaluated (e.g. a common qualification for monetary values subject to inflation). The third category is Spread, which is a generalisation from the 'random error' of experiments or the variance of statistics. Although Spread is usually conveyed by a number (either +, % or 'factor of') it is not an ordinary quantity, because its own inexactness is of a different sort from that of measurements. Methods to address Spread can be statistical data analysis, sensitivity analysis or Monte Carlo analysis, possibly in combination with expert consultation.

The remaining two qualifiers constitute the more qualitative side of the NUSAP framework. 'Assessment' expresses qualitative judgements 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 qualities of optimism or pessimism. In some experimental fields, information is supplied qualified by two + 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 the use of Assessment in NUSAP. A frequently observed pitfall is to wrongly think that systematic error must always be less than any experimental error, or else a stated error bar would be meaningless or misleading. However, in many real life cases systematic error can be well estimated only in retrospect, and then it can produce surprising results that are far outside the error bar of the previously published number(s).

The fifth and final aspect of NUSAP is the Pedigree. This 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 as a set of criteria and assessed using qualitative expert judgement. Risbey et al. (2001) document a method to draft Pedigree scores by means of expert elicitation, on which Knol et al. (2010) provide guidance as to good practice. Arbitrariness and subjectivity in measuring strength are minimised by using a Pedigree matrix to code qualitative expert judgements for each criterion into a discrete Numeral scale from 0 (weak) to 4 (strong) accompanied by linguistic descriptors or modes.

Pedigree and its assessment

Each special sort of information has its own aspects that are key to its Pedigree, so different Pedigree matrices using different criteria can be used to qualify different sorts of information. An overview of the literature on Pedigree matrices and examples of questionnaires for eliciting Pedigree scores is available online at http://www.nusap.net. Ellis et al. (2000) have developed a Pedigree calculator to assess propagation of Pedigree in a calculation in order to establish Pedigree scores for quantities calculated from other quantities. Table 29.2 gives an example of a Pedigree matrix for emission monitoring data. Next I will briefly elaborate the four criteria employed.

Proxy

Sometimes measuring the thing we are interested in directly, or representing it by a parameter, is impossible, so some form of proxy measure is used. Proxy refers to how good or close a measure of the quantity that we measure or model is to the actual quantity we seek or represent. Examples are first order approximations, over simplifications, idealisations, gaps in aggregation levels, differences in definitions, non-representativeness and incompleteness issues.

Table 29.2 Example pedigree matrix for emission monitoring data

Score	*Proxy*	*Empirical basis*	*Methodological rigour*	*Validation*
4	An exact measure of the desired quantity	Controlled experiments and large sample direct measurements	Best available practice in well-established discipline	Compared with independent measurements of the same variable over long domain
3	Good fit or measure	Historical/field data uncontrolled experiments small sample direct measurements	Reliable method common within est. discipline Best available practice in immature discipline	Compared with independent measurements of closely related variable over shorter period
2	Well correlated but not measuring the same thing	Modelled/derived data Indirect measurements	Acceptable method but limited consensus on reliability	Measurements not independent proxy variable limited domain
1	Weak correlation but commonalities in measure	Educated guesses indirect approx. rule of thumb est.	Preliminary methods unknown reliability	Weak and very indirect validation
0	Not correlated and not clearly related	Crude speculation	No discernible rigour	No validation performed

Source: Risbey et al. (2001) adapted from Ellis et al. (2000a, 2000b).

Empirical basis

This typically refers to the degree to which direct observations, measurements and statistics are used to estimate a parameter. Sometimes directly observed data are unavailable, and the parameter or variable is estimated based on partial measurements or calculated from other quantities. Parameters or variables determined by such indirect methods have a weaker empirical basis and will generally score lower than those based on direct observations.

Methodological rigour

Parameter or variable estimates employ a method to collect, check, and revise the data. Methodological quality refers to the norms for methodological rigour in this process, as applied by peers in relevant disciplines. Well-established and respected methods for measuring and processing data would score high on this metric, while untested or unreliable methods would tend to score low.

Validation

This metric refers to the degree to which the analyst has been able to cross-check the data and assumptions used to produce the Numeral of the parameter against independent sources. In many cases, independent data for the same parameter over the same time period are unavailable and other data sets must be used for validation. This may require a compromise in the length or overlap of the data sets, or may require use of a related, but different, proxy variable for indirect validation, or perhaps use of data that has been aggregated on different scales. The more indirect or incomplete the validation, the lower it will score on this metric.

Visualising Pedigree analysis

In general, Pedigree scores will be established using expert judgements from more than one expert. Two ways of visualising results of a Pedigree analysis are discussed here: radar diagrams and kite diagrams (Risbey et al., 2001; Van der Sluijs et al., 2002), as exemplified in Figure 29.1. Both representations use polygons with one axis for each criterion, having 0 in the centre and 4 on each corner point. In the radar diagrams a coloured line connecting the scores represents the scoring of each expert, whereas a black line represents the average scores.

The kite diagrams follow a traffic light analogy. The minimum scores by a group of experts for each Pedigree criterion span the green kite; the maximum scores span the amber kite. The remaining area is red. The width of the amber band represents expert disagreement on the Pedigree scores. In some cases the size of the green area can be strongly influenced by a single deviating low score given by one of the experts. In those cases the light green kite shows what the green kite would look like if that outlier had been omitted. Note that the algorithm for calculating the light green kite is such that outliers are evaluated per Pedigree criterion, so that outliers defining the light green area need not be from the same expert. A web-tool to produce kite diagrams is available from http://www.nusap.net.

The kite diagrams can be interpreted as follows: the green coloured area reflects the (apparent minimal consensus) strength of the underpinning of each parameter. The more green, the stronger is the underpinning. The orange coloured zone shows the range of expert disagreement on that underpinning. The remaining area is red. The more red, the weaker is the underpinning (all according to the assessment by a group of experts). A kite diagram captures the information from all experts in the group without the need to average expert opinion. Averaging expert opinion is a controversial issue in elicitation methodologies (Knol et al., 2010). Another advantage is that it provides a fast and intuitive overview of parameter strength, preserving the underlying information. However, kite diagrams can be misleading because the amount of red and green surface area can be sensitive to the order of the criteria in the diagram. As an alternative, bar charts can be used with error-bars to reflect the range of expert opinion; see Kloprogge et al. (2007) and Wardekker et al., (2008) for further guidance.

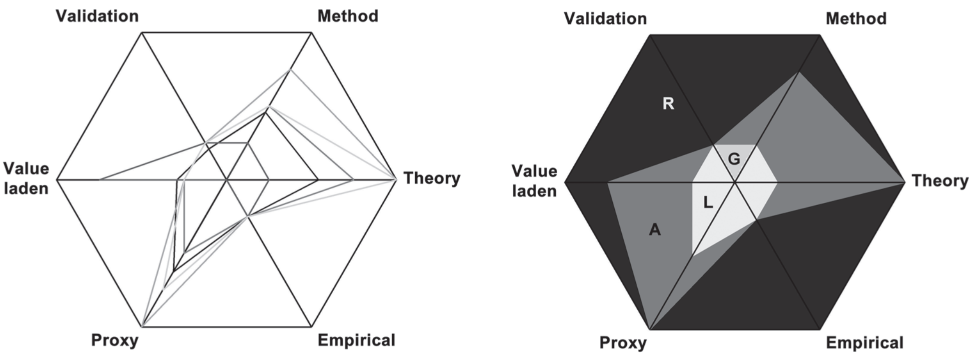

Figure 29.1 Graphic representation of pedigree scoring.
Notes: Example of representations of same pedigree results as scored by six different experts by radar diagram (left) and kite diagram (right) G=green, L=light green, A=amber, R=red.
Source: Van der Sluijs et al (2002).

Diagnostic Diagrams

There are two independent metrics that can be used for diagnostic purposes. First is the method chosen to address the Spread qualifier (typically sensitivity analysis or Monte Carlo analysis) which provides, for each input quantity, a quantitative metric of uncertainty contribution, or sensitivity (e.g., the relative contribution to the variance in a given model output). Second are the Pedigree scores that can be aggregated (by dividing the sum of the scores of the Pedigree criteria by the sum of the maximum attainable scores) to produce a metric for parameter strength. These two independent metrics can be combined in a NUSAP Diagnostic Diagram (see Figure 29.2).

The Diagnostic Diagram is based on the notion that neither Spread nor strength can alone provide a sufficient measure of quality. Robustness of model output to parameter strength could be good even if parameter strength is low, provided that the model outcome is not critically influenced by the Spread in that parameter. In this situation our ignorance of the 'true value' of a parameter has no immediate consequences because it has a negligible effect on calculated model outputs. Alternatively, model outputs can be robust against parameter Spread—even if its relative contribution to the total Spread in a model is high—provided that parameter strength is high. In the latter case, the uncertainty in the model outcome adequately reflects the inherent irreducible uncertainty in the system represented by the model. In other words, the uncertainty then is a property of the modelled system and does not stem from imperfect knowledge about that system. Mapping model parameters in the Assessment diagram thus reveals the weakest critical links in the knowledge base of the model with respect to the model outcome assessed, and helps in setting the priorities for model improvement.

Most of the Pedigree assessments in the literature have addressed uncertainties located in inputs and parameters, thereby focussing on the internal strength of the knowledge base. Kloprogge et al. (2011) extended Pedigree analysis to assess assumptions in models. Examples of putting the approach into practice include Laes et al. (2011) for evaluating the external costs of nuclear energy, De Jong et al. (2012) for quantified health risks of overhead power lines, and Boone et al. (2009) and Bouwknegt et al. (2014) for quantitative microbial risk assessment. Van der Sluijs et al. (2015) have further extended the application of NUSAP to assess modelling assumptions in a chain of integrated models in the context of decisions concerning local adaptation to climate change impacts.

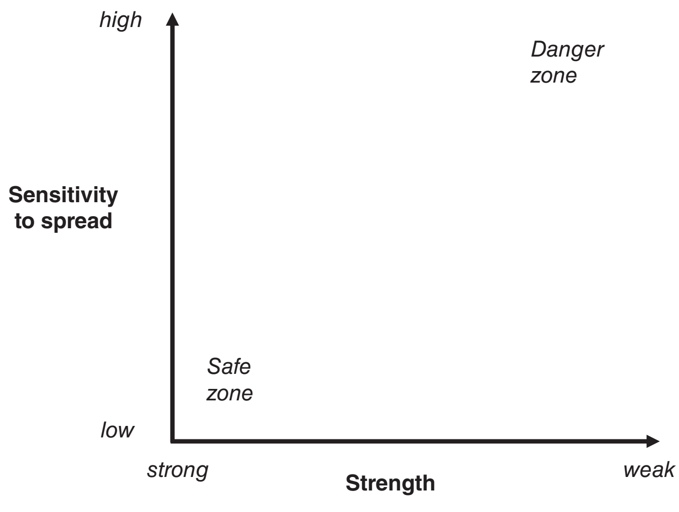

Figure 29.2 NUSAP diagnostic diagram.

Corral (2000), in his Ph.D. thesis, extended the Pedigree scheme to address uncertainties located in the socio-political context, focussing on the external strength of the knowledge base (i.e., its relationship to the world outside of science). The criteria that Corral used to assess the Pedigree of the processes of knowledge utilisation and institutional context of the analysts were inter alia: accessibility, terminology, completeness, source of information, verification, colleague consensus, extended peer acceptance, legitimation, experience, and flexibility.

Future directions

The NUSAP approach has a great potential to systematise the appraisal and consideration of uncertainty at the science–governance interface. Further tailoring and standardisation of Pedigree matrices and procedures for the elicitation of Pedigree scores is desirable but the main challenge is in dissemination. Successful pilots with inclusion of NUSAP in M.Sc. and Ph.D. teaching curricula at the universities of Utrecht, Bergen and Versailles Saint-Quentin-en-Yvelines can be scaled up. An open access course in knowledge quality assessment is also now available online (https://proxy.eplanete.net/galleries/broceliande7/KQA).

Concluding remarks

Overall, NUSAP has a strong foundation in the theory of knowledge and the philosophy of science connecting to post-normal science [see Chapter 28]. It provides a framework to systematically and coherently address and communicate three of the dimensions of uncertainty, namely: technical (inexactness), methodological (unreliability) and epistemological (bordering with ignorance). It provides a framework for synthesising qualitative and quantitative assessments of uncertainty and can act as a bridge between the quantitative mathematical disciplines/traditions and the qualitative discursive and participatory disciplines/traditions in the field of uncertainty management.

Key further readings cited

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

van der Sluijs, J.P., Craye, M., Funtowicz, S.O., Kloprogge, P., Ravetz, J.R. and Risbey, J. (2005a). Combining quantitative and qualitative measures of uncertainty in model based environmental assessment: the NUSAP system. *Risk Analysis* 25(2), 481–492.

van der Sluijs, J.P., Risbey, J. and Ravetz, J.R. (2005b). Uncertainty assessment of VOC emissions from paint in the Netherlands. *Environmental Monitoring and Assessment* 105, 229–259.

van der Sluijs, J.P., Petersen, A.C., Janssen, P.H.M., Risbey, J.S. and Ravetz, J.R. (2008). Exploring the quality of evidence for complex and contested policy decisions. *Environmental Research Letters*, 3, 024008 (9 pp.).

van der Sluijs J.P. and Wardekker, J.A. (2015). Critical appraisal of assumptions in chains of model calculations used to project local climate impacts for adaptation decision support: The case of Baakse Beek. *Environmental Research Letters* 10, 045005.

Other literature cited

Boone, I., Van der Stede, Y., Bollaerts, K., Vose, D., Maes, D., Dewulf, J., Messens, W., Daube, G. Aerts, M. and Mintiens, K. (2009). NUSAP method for evaluating the data quality in a quantitative microbial risk assessment model for Salmonella in the pork production chain. *Risk analysis* 29(4), 502–517.

Bouwknegt, M., Knol, A.B., van der Sluijs, J.P., Evers, E.G. (2014). Uncertainty of population risk estimates for pathogens based on QMRA or epidemiology: a case study of Campylobacter in the Netherlands. *Risk analysis*, 34(5), 847–864.

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

De Jong, A., Wardekker, J.A., van der Sluijs, J.P. (2012). Assumptions in quantitative analyses of health risks of overhead power lines. *Environmental science & policy* 16, 114–121.

Ellis, E.C., Li, R. G., Yang, L. Z. and Cheng, X. (2000). Long-term change in village-scale ecosystems in China using landscape and statistical methods. *Ecological Applications* 10, 1057–1073.

Funtowicz, S.O. and Ravetz, J.R. (1993). Science for the post-normal age. *Futures* 25(7), 735–755.

Keepin B. and Wynne, B. (1984). Technical analysis of IIASA energy scenarios. *Nature* 312, 691–695.

Kloprogge, P., van der Sluijs J.P., and Wardekker, A. (2007). Uncertainty communication: issues and good practice, report NWS-E-2007-199, Department of Science Technology and Society, Copernicus Institute, Utrecht University. 60 pp.

Kloprogge, P., van der Sluijs J.P., and Petersen A.C. (2011). A method for the analysis of assumptions in model-based environmental assessments. *Environmental Modelling & Software* 26(3), 289–301.

Knol, A.B., Slottje, P., van der Sluijs, J.P., and Lebret, E. (2010). The use of expert elicitation in environmental health impact assessment: a seven step procedure. *Environmental Health* 9, 1–16.

Laes, E., Meskens G., and van der Sluijs, J.P. (2011). On the contribution of external cost calculations to energy system governance: The case of a potential large-scale nuclear accident. *Energy Policy* 39(9), 5664–5673.

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

Nowotny, H. (1999). The place of people in our knowledge. *European Review* 7(2), 247–262.

Petersen, A.C., Cath, A., Hage, M., Kunseler, E., and van der Sluijs, J.P. (2011). Post-normal science in practice at the Netherlands Environmental Assessment Agency. *Science Technology & Human Values* 36(3), 362–388.

Risby, J. S., van der Sluijs, J.P., Ravetz, J. (2001). *Protocol for Assessment of Uncertainty and Strength of Emission Data*, Department of Science Technology and Society, Utrecht University, report nr. E-2001-10, 22 pp.

Risbey, J., van der Sluijs, J. P., Kloprogge, P., Ravetz, J., Funtowicz, S. and Corral Quintana S. (2005). Application of a checklist for quality assistance in environmental modelling to an energy model. *Environmental Modelling & Assessment* 10(1), 63–79.

Saltelli, A., Chan, K., Scott, E.M. (eds.). (2000). *Sensitivity Analysis*. Chichester: John Wiley & Sons.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. Chichester: Wiley.

Saltelli, A., Guimarães Pereira, Â., van der Sluijs, J.P. and Funtowicz, S. (2013). What do I make of your Latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy* 9(2/3/4), 213–234.

Van der Sluijs, J.P. (1997). *Anchoring Amid Uncertainty: On the Management of Uncertainties in Risk Assessment of Anthropogenic Climate Change*, Ph.D. dissertation, Utrecht: Utrecht University.

Van der Sluijs, J.P. (2002). A way out of the credibility crisis around model-use in Integrated Environmental Assessment. *Futures* 34, 133–146.

van der Sluijs, J.P., Potting, J., Risbey, J., van Vuuren, D., de Vries, B., Beusen, A., Heuberger, P., Corral Quintana, S., Funtowicz, S., Kloprogge, P., Nuijten, D., Petersen, A., Ravetz, J. (2002). *Uncertainty assessment of the IMAGE/TIMER B1 CO2 emissions scenario, using the NUSAP method*, Dutch National Research Program on Climate Change, Report no: 410 200 104, Bilthoven (available from www.nusap.net).

van der Sluijs, J. P., Craye, M., Funtowicz, S., Kloprogge, P., Ravetz, J. and Risbey, J. (2005c). Experiences with the NUSAP system for multidimensional uncertainty assessment in Model based Foresight Studies. *Water Science and Technology* 52(6), 133–144.

Walker, W.E., Harremoës, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P, Krayer von Krauss, M.P. (2003) Defining uncertainty, a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment* 4(1), 5–17.

Wardekker, J.A., van der Sluijs, J.P., Janssen, P.H.M., Kloprogge, P., Petersen, A.C. (2008). Uncertainty communication in environmental assessments: Views from the Dutch science-policy interface. *Environmental Science and policy* 11, 627–641.

Wynne, B. (1992). Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm. *Global Environmental Change* 6, 111–127.