SDRN Rapid Research and Evidence Review on
Emerging Methods for Sustainability Valuation and
Appraisal

A report to the Sustainable Development Research Network

Final Report

Sigrid Stagl
January 2007
Acknowledgments

I am profoundly grateful to all those who have contributed inputs of various kinds to this work. Particular thanks should go to Malcolm Eames, previously at PSI, now at Brunel University, for his invaluable guidance and support throughout; to interviewees from various government departments for their time and for sharing their views and insights; and to Kate McGeevor and Ben Shaw for excellent organisational support and editing.

Numerous individuals have provided advice, comment, criticism and suggestions for improving the work and improving its clarity and scope. I owe debts of gratitude to all eighteen interviewees for their time and giving me insight into their specific context within which decisions for more sustainable development are prepared in the UK. Special thanks also to the participants of the kick-off and expert workshops and reviewers from academia and the policy arena for their time and effort; they were Rebecca Anderson, Victor Anderson, Stephen Bass, Paul Caskie, Richard Clarkson, Peter Costigan, Ruth Coward, Rob Curry, Geoff Dawe, Michael Doble, Eleanor Field, Karin Gavelin, Keith Goulding, Virginia Hall, Claire MacAulay, Sasha Maguire, Emily McKenzie, Sarah Moon, Shaun Mowat, Paula Orr, Alan Pearman, Tony Pike, Diana Pound, Domah Preetum and Steve Preston.
Executive summary

1. As sustainable development rises in importance on the political agenda and the appraisal of policies, programmes and projects against sustainability criteria becomes more widespread, theorists and practitioners are asking fundamental questions about the nature of appraisal and its role in the political process.

2. The type of evaluation and the institutional structures in which an appraisal is embedded influences its outcome. Therefore the choice of valuation and appraisal methods is not a wholly technical question, but one of ‘institutionalising social choice’.

3. The aims of the review are: (1) to provide an accessible introduction to the theoretical background and underpinnings of recently developed multi-criteria, participatory and deliberative methods; (2) to provide practical case-studies of these new decision support tools; and (3) to explore the utility, relevance and limitations of such approaches for sustainability valuation and appraisal in the UK context.

4. A range of new sustainability valuation and appraisal methods have been developed and tested over the last 10 years. They combine interpersonal deliberation with quantitative methods.

5. Six new sustainability valuation and appraisal methods are covered in this report. While there are some differences between these methods and the problems they are suited to, a greater difference exists between this group of methods and monetary environmental valuation on the one hand and deliberative methods on the other hand. Compared to monetary valuation (cost-benefit analysis) they focus more on ‘opening up’ the issues that is exploration of the complexities of the issue and also allow for incommensurability of values. Compared to established deliberative methods (e.g. citizens’ juries or panels, consensus conferences, future search, deliberative polling) the methods reviewed here contain a stronger element of analysis and a more structured approach for ‘closing down’, that is analysis and deliberation over ‘best option(s)’ and information for policy decisions. As with more established deliberative methods, the quality of facilitation is of vital importance to the success of these new methods.

6. The six methods reviewed here build on a number of common principles. They all: account for different types of knowledge (monetary and non-monetary; quantitative and qualitative data); consider seriously the issue of inter-generational equity; provide opportunities for learning during the appraisal process; ensure transparency of each step of the appraisal process; and have a strong element of public and stakeholder engagement.

7. The six methods for sustainability valuation and appraisal reviewed in this report are -

i. **Deliberative monetary valuation**

   The use of formal deliberation concerning an environmental impact to express value in monetary terms for policy purposes, and more specifically as an input to cost-benefit-analysis. Cost-benefit analysis was developed for the appraisal of infrastructure projects. Deliberative monetary valuation differs from cost-benefit-analysis primarily in that preferences are constructed during the deliberation process. Like cost-benefit analysis, deliberative monetary valuation is most suitable for the appraisal of projects whose impacts are relatively well understood, where the impacts do not reach far into the future, and which do not affect complex ecosystem services such as biodiversity.

ii. **Social multicriteria evaluation**

   The combined use of participatory techniques and multicriteria analysis to aid decision making about a number of policy, programme or project options while taking conflicting interests and multiple criteria into account. It highlights transparency and social learning during the appraisal process. This method is
most suitable for the appraisal of policies, programmes or projects whose impacts are not well understood yet and therefore benefit from a multidisciplinary modelling of impacts. Social multicriteria evaluation was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

iii. Three-stage multicriteria analysis

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy, programme or project options. Here, the sequencing and choice of participants is based on the ‘co-operative discourse’ model. In this model, stakeholders select the evaluation criteria, experts present information and measure impacts, and citizens explore values. This method is most suitable for the appraisal of policies, programmes or projects whose impacts are reasonably well understood by experts and which contain a significant technical element. Three-stage multicriteria analysis was developed for the social appraisal of technologies with particular emphasis on the role of risk and uncertainty.

iv. Multicriteria mapping

An interview-based multicriteria analysis that focuses on eliciting and documenting detailed technical and evaluative judgements concerning the performance of alternative options. Individual specialists and stakeholders appraise the performance of core and discretionary options against their own sets of criteria. It highlights the systematic exploration of uncertainties and the sources of variability between diverse viewpoints. This method is most suitable for the appraisal of policies, programmes or projects to which stakeholders had some exposure and where views not only about data, but also about options and criteria, are controversial. Multicriteria mapping was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

v. Deliberative mapping

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy options. In addition to measuring the specific performance of each option against criteria, this method highlights the need to explore the arguments participants use to justify their judgements. Specialists and small groups of citizens follow the same assessment process. This method is most suitable for the appraisal of policies, programmes or projects where views are controversial and where value judgements are particularly important. Deliberative mapping was originally developed to address complex issues and to deal with uncertainty and contested values in the appraisal of controversial technologies.

vi. Stakeholder decision / dialogue analysis

The combined use of group deliberation techniques and (a qualitative form of) multicriteria analysis to aid decision making about policy options. It highlights the framing of problems, scoping options, eliciting criteria and making judgements through facilitated deliberation. This method is most suitable for the appraisal of policies, programmes or projects where it is important to work first on a common problem understanding and for which a rough impact assessment is sufficient as input in the decision process. Stakeholder decision analysis was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

8. While these new methods will benefit from further testing and development, they have been shown to be effective tools of sustainability appraisal and valuation under a range circumstances.
9. The Treasury’s Green Book offers detailed guidance for monetary valuation. The sections on alternative methods are extremely short and guidance on when to use which method is missing. Most departmental guidance documents build explicitly on the Green Book.

10. The manner in which monetary, and other quantitative and qualitative impacts are currently summarised in Regulatory Impact Assessments (RIAs), resembles an impact matrix (options are appraised against a range of criteria). In a multicriteria appraisal the next steps would be to weight the criteria and to apply an algorithm that calculates a ranking of options. Under the current RIAs, decision-makers are required to do these final two steps by themselves without technical support, which the decision making literature would consider an excessive demand on human judgement. As a result, the danger arises that not all criteria are fully taken into account.

11. While the major guidance documents highlight the benefits of monetary valuation (ease of comparison, expression of opportunity costs etc.), they acknowledge that monetary valuation may sometimes be difficult and leave it up to the appraiser to judge how benefits and costs can be best represented. While there are few formal limitations to applying the new appraisal methods at the policy, programme or project level, in several departments there is a culture of “monetary valuation is the preferred appraisal method”. At the same time many difficulties in applying monetary valuation across all policy areas are known and expressed in these departments. The guidance in the Green Book could be improved by presenting information on a range of appraisal methods, comparing and contrasting them, and giving advice on which methods are most suitable under which circumstances. The guidance from the RIA could be improved by retaining the existing openness to different types of impacts and the ways they are measured, but developing a framework which gives more support to decision-makers.

12. Since there is no one method suited to appraising all types of policies, programmes and projects, a more differentiated approach should be favoured. In order to develop cultures that support the search for the most suitable methods, departments need to be familiar with the different methods at hand and provide guidance on them so appraisers can choose the most appropriate method.
Table 1. Summary of the six methods included in this review

<table>
<thead>
<tr>
<th>Method</th>
<th>Origin of criteria</th>
<th>Transparency</th>
<th>Public and stakeholder engagement</th>
<th>Focus on opening up</th>
<th>Focus on closing down</th>
<th>Robustness</th>
<th>Approach to Uncertainty</th>
<th>Outputs that the approach is good at producing</th>
<th>Indicative costs (thousands £)</th>
<th>Indicative time (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deliberative monetary valuation</td>
<td>citizens panel/s and researchers</td>
<td>**</td>
<td>*</td>
<td>****</td>
<td>**</td>
<td></td>
<td>Sensitivity or scenario analysis</td>
<td>Monetary value</td>
<td>40-50</td>
<td>4 – 6</td>
</tr>
<tr>
<td>Social multicriteria evaluation</td>
<td>stakeholders and / or research team</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>Fuzzy numbers; sensitivity or scenario analysis</td>
<td>Complete or partial ranking</td>
<td>30-50</td>
<td>4 – 8</td>
</tr>
<tr>
<td>Three-stage multicriteria analysis</td>
<td>stakeholders</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>Sensitivity or scenario analysis</td>
<td>Complete ranking</td>
<td>30-40</td>
<td>4 – 12</td>
</tr>
<tr>
<td>Multicriteria mapping</td>
<td>stakeholders and research team</td>
<td>****</td>
<td>**</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>Optimistic and pessimistic scores; sensitivity or scenario analysis</td>
<td>Map of perspectives and ranking, plus discourse analysis</td>
<td>20-30</td>
<td>2 – 6</td>
</tr>
<tr>
<td>Deliberative mapping</td>
<td>stakeholders and research team</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>Optimistic and pessimistic scores; Sensitivity or scenario analysis</td>
<td>Map of perspectives and ranking, plus discourse analysis</td>
<td>30-40</td>
<td>4 – 12</td>
</tr>
<tr>
<td>Stakeholder decision analysis</td>
<td>stakeholders</td>
<td>****</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td>Qualitative analysis</td>
<td>Complete or partial ranking</td>
<td>20-30</td>
<td>4 – 12</td>
</tr>
</tbody>
</table>

* = worst; ***** = best
Notes:

1. An initial list of criteria is often suggested by the research team, which can then be altered by participants. Some methods (MCM, DM) allow each individual participant to add criteria to the common list, but most methods require a common list of criteria to be used by all participants.

2. Transparency is rated higher, if participants have the opportunity to go through the whole process and if the tools applied are simple and can be easily explained in a workshop.

3. All reviewed methods include a substantial amount of citizen, expert and/or stakeholder engagement. However, with some methods a wider range of elements (criteria, options, weights, policy suggestions) is drawn from the deliberations than in others.

4. Like with public and stakeholder engagement, all methods reviewed include processes of opening up the issues and then supporting closing down and arriving at recommendations. However, the focus lies in some methods more on the one or on the other aspect.

5. Robustness: being rigorous about data as well as framing assumptions, includes “social robustness” by involving a variety of participants.

6. The way the methods address uncertainty is briefly characterised; for most methods this includes some form of sensitivity / scenario analysis, but several go further than this.

7. The types of outputs that the approaches are good at producing are briefly characterised; a ranking of options is partial if incomparability between options arises; a ranking of options is complete, if no such incomparabilities (are assumed to) arise; the stronger the assumptions on which the algorithm is based the more likely that a complete ranking can be calculated.

8. The costs represent very rough estimates for a medium size standard application and may vary significantly for specific applications.

9. Time required is influenced mostly by whether the method requires several stages of public and stakeholder engagement or not.
1. Introduction

Background to the review

The appraisal of projects, plans, programmes, and technologies has been an increasingly important element in the quest for moving toward sustainable development. From the late 1960s, many countries passed environmental policies which enacted legislation for environmental impact assessment of projects. More recently, and partly in response to the limitations of a project-based approach, practices such as strategic environmental assessment (SEA) and sustainability appraisal have been promoted. This shift has been accompanied by high-profile policy developments, such as the adoption of the European ‘SEA Directive’ (EC 2001). Additionally, in the UK the Green Book and the Regulatory Impact Assessment give general guidance on how to assess projects, programmes and policies. At the same time, fundamental questions are being asked by theorists and practitioners, about the nature of appraisal and its role in the political process (Owens et al., 2004). These questions will be explored further in section 2.

This study differs from previous ones on related topics in various ways. The recently completed study on ‘Valuing Our Natural Environment’ by eftec (2006) covers a much broader suite of valuation methods and therefore did not review hybrid methodologies in detail. ‘Multi-criteria Analysis: A Manual’ commissioned by DLTR (Dodgson et al., 2000) covers the different types of multicriteria algorithms in much greater depth, but does not review systematically the hybrid methodologies. The only direct overlap is multicriteria mapping, which is covered in both reports. A number of studies have been commissioned over the last few years reviewing environmental valuation and appraisal methods for applications in specific areas, such as urban water management (Starkl et al., 2004; 2005). These studies tend to cover only the specific section of sustainability valuation and appraisal methods that is relevant for the issue at hand.

Scope and content

This review aims to -

1) provide an accessible introduction to the theoretical background and underpinnings of recently developed multi-criteria, participatory and deliberative methods;
2) provide practical case-studies of the use of these new decision support tools; and
3) explore the utility, relevance and limitations of such approaches for sustainability valuation and appraisal in the UK context.

The methods covered in this report are listed below. A glossary of terms is included in Appendix 1 to this report and each of the methods is described in more detail in chapter 3.

Deliberative monetary valuation

The use of formal deliberation concerning an environmental impact to express value in monetary terms for policy purposes, and more specifically as an input to cost-benefit-analysis (Spash 2001). It differs from cost-benefit-analysis primarily in that preferences are constructed during the deliberation process.

Social multicriteria evaluation

The combined use of participatory techniques and multicriteria analysis to aid decision making about a number of policy options while taking conflicting interests and multiple criteria into account. This method highlights transparency and social learning during the appraisal process (Munda 2004).
Three-stage multicriteria analysis

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy options. Here, the sequencing and choice of participants is based on Renn’s and Webler’s (1998; 1993) ‘co-operative discourse’ model. In this model, stakeholders select the evaluation criteria, experts present information and measure impacts, and citizens explore values.

Multicriteria mapping

An interview-based multicriteria analysis that focuses on eliciting and documenting detailed technical and evaluative judgements concerning the performance of alternative policy options. Individual specialists and stakeholders appraise the performance of core and discretionary options against their own sets of criteria. It highlights the systematic exploration of uncertainties and the sources of variability between diverse viewpoints (Stirling 1997).

Deliberative mapping

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy options. Besides measuring the specific performance of each option against criteria, this method highlights the need to explore the arguments participants use to justify their judgements. Specialists and small groups of citizens follow the same assessment process (Davies et al., 2003).

Stakeholder decision / dialogue analysis

The combined use of group deliberation techniques and (a qualitative form of) multicriteria analysis to aid decision making about policy options. This method highlights the framing of problems, scoping options, eliciting criteria and making judgements through facilitated deliberation (Burgess 2000).

Definitions

In this project, the key terms are defined as follows:

Based on Dunn’s (1981) and Owens’ (2004) definitions of (policy) appraisal, sustainability appraisal means the application of a variety of methods of inquiry and argument to produce policy-relevant information that is then utilised to evaluate the consequences of human actions against the normative goal of sustainable development.

Sustainability valuation means methods and procedures for capturing changes in a social-ecological system or subsystems; the valuation results are fed into the appraisal process.

Sustainability valuation and appraisal are more comprehensive than mere data collection and consultation methods. They differ also from ‘decision-making methods’ (e.g. in Operations Research) by acknowledging that resolving public problems and decision-making about public policies is a process in which sustainability valuation and appraisal methods should have an input, but they do not replace this process.

This report focuses on ‘hybrid methodologies’ of valuation and appraisal in the context of sustainable development. The Oxford English Dictionary defines ‘hybrid’ as “[d]erived from heterogeneous or incongruous sources; having a mixed character; composed of two diverse elements”. Hybrid methodologies combine interpersonal deliberation and quantitative methodologies to produce both depth and breadth in valuation and appraisal processes. They have the potential to “resituate specialist knowledge claims through attention to their framing conditions and boundaries of uncertainty, while co-producing new forms of citizen and stakeholder expertise, thus opening up the appraisal of projects, plans, programmes, and technologies to other forms of framing and reasoning” (Davies, 2006: 235).
Methodology

The project was conceived primarily as a desk study, making use of existing material, developing new thinking, communicating with academics and stakeholders and drafting papers. Its three major components are –

i) Synthesis of research and thinking on the role of sustainability valuation and appraisal. This involves bringing together insights from the author’s own recent and current work with those of others, based on an extensive review of relevant literatures and official discourse, and dialogue with academics, practitioners and policymakers in the field.

ii) Rigorous analysis of the above material, to address the themes and questions outlined above; to distil key findings; and to identify outstanding questions and emergent issues.

iii) Development of a set of key issues and suggestions for using sustainability valuation and appraisal methods in the UK context.

For this purpose a review template was developed and applied to six novel sustainability valuation and appraisal methods.
Table 2. Review Template: Appraisal Methods

<table>
<thead>
<tr>
<th>Summary</th>
<th>Description of the method and associated tools.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Description of the method and associated tools.</td>
</tr>
<tr>
<td><strong>Framing</strong></td>
<td>Perception of the decision situation. Conceptual background and assumptions made. Was the method explicitly designed for sustainability appraisal and what dimensions of sustainable development are included in the appraisal?</td>
</tr>
<tr>
<td><strong>Nature of participation</strong></td>
<td>Type of involvement (expert/stakeholder/public), timings in the process, issues of representation, outcome or impact in the project or subsequent action.</td>
</tr>
<tr>
<td><strong>Decision context</strong></td>
<td>Is the method intended for “opening up” and/or “closing down” a choice of options?</td>
</tr>
<tr>
<td><strong>Treatment of criteria and values</strong></td>
<td>Measurement of impacts and way of capturing different societal interests. Includes discussion of type of data used (quantitative / qualitative; derived from models / elicited from experts or stakeholders).</td>
</tr>
<tr>
<td><strong>Applicability to different types of evidence bases</strong></td>
<td>Applicability to different states of knowledge: (1) Risk: the different possible outcomes are known exactly and a probability can be assigned to each possibility; (2) Ambiguity: the probabilities are known but the outcomes to which they attach are not known exactly; (3) Uncertainty: the different possible outcomes are known but probabilities cannot be assigned to them; (4) Ignorance: the definition of a complete set of possible outcomes is problematic and probabilities cannot be assigned.</td>
</tr>
<tr>
<td><strong>Treatment of equity and distributional issues</strong></td>
<td>How does the method address the question of equity between and within human generations?</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td>How are the results presented? Cardinal or ordinal ranking? With or without description of the reasonings?</td>
</tr>
<tr>
<td><strong>Human and financial resources</strong></td>
<td>Resource requirements of a typical application.</td>
</tr>
<tr>
<td><strong>Key studies</strong></td>
<td>Brief description and references for exemplary studies.</td>
</tr>
</tbody>
</table>

The report is structured as follows: The following section gives a concise overview of how the field of sustainability valuation and appraisal has developed in recent decades. Section three describes and reviews six sustainability valuation and appraisal methods. Section four explores the institutional and policy context of sustainability appraisal in the UK. Section five concludes by highlighting key messages.
2. Development of sustainability valuation and appraisal methods

As sustainable development rises in importance on the political agenda and the appraisal of policies, programmes and projects against sustainability criteria becomes more widespread, theorists and practitioners ask fundamental questions about the nature of appraisal and its role in the political process.

Sustainable development is a multidimensional concept: How can potentially conflicting impacts that are measured in different units (monetary / non-monetary) and in different ways (quantitative / qualitative) be brought together to aid decision-making (Martinez-Alier et al., 1998)?

The interface between science and policy is complex: How can appraisal tools aid decision-making when the field of policy analysis increasingly rejects the concept of neutral, objective advice (Owens et al., 2004)?

Natural and human systems are adaptive: How can valuation and appraisal tools account for the dynamic and uncertain nature of natural and human systems (Folke et al., 2005)?

The aim of sustainable development puts special demands on valuation and appraisal methods. The fact that sustainable development is (i) a multidimensional concept, (ii) crucial for human survival, and (iii) a long-term issue, makes it necessary to test the applicability of existing valuation and appraisal tools in this context.

It is often argued that for the environment to be taken seriously in government and business decisions, it must be assigned a monetary value. Environmental economists have developed and used a range of methods including travel cost, hedonic pricing, production function analysis, contingent valuation and choice modelling (Hanley and Spash, 1993) to attribute monetary values to the ‘environment’ in decision-making processes. This approach has been successful in that many national and international agencies are performing monetary valuation exercises as part of their overall assessment of projects. However, there has also been criticism of the monetary valuation of environmental goods and services. Critiques can be broadly grouped into those concerned with the theoretical foundations of economic valuation, and those looking at the validity of the specific numbers produced and the tools employed.

Cost-benefit analysis claims that benefits and costs can be expressed in monetary terms and hence made comparable or commensurable. A significant literature in ethics, political philosophy and economics has arisen which doubts this assumption. For example, market boundaries should exist because trading children, certain drugs or weapons of mass destruction is morally wrong. The mix of moral and economic in these choices is just what critics of cost-benefit analysis claim characterizes environmental policy. Hence, the surrogate markets of cost-benefit analysis are regarded as similarly unacceptable in these contexts (Anderson, 1993; O’Neill, 1993; Sunstein, 1997; Vatn and Bromley, 1994). If CBA is inappropriate for political decisions concerning say, abortion policy, then it is argued to be inappropriate for much environmental policy too. Another strand of academic literature argues that incommensurability arises when a rational agent is unable to attach a monetary value to certain “goods” (e.g. environmental assets) for legitimate reasons (see Aldred 2006 for a more comprehensive review of these issues). It is widely recognized that the incommensurability problems facing monetary valuation are particularly noticeable, and particularly acute, in environmental contexts. Thus David Pearce acknowledged, “the issue of ‘incommensurables’ grew to be the single most controversial issue in cost-benefit analysis, and it remains so today” (Pearce, 2000: 51). These concerns from the academic arena have so far had limited impact on the practice of sustainability appraisal.

In another area of critical analysis, cost-benefit-analysis, along with mainstream microeconomics, has been criticised for building upon axioms of choice which are inconsistent with theories of modern psychology and empirical evidence. Kahneman and Tversky’s (1979) prospect theory revealed that people value gains and losses asymmetrically,
which can explain the observed gap between willingness-to-pay and willingness-to-accept measures. Knetsch (1995) showed the frequent occurrence of behaviours that are inconsistent with accepted economic norms but commonly dismissed by economic models. The refusal to make trade-offs has been shown to arise in cost-benefit-analysis studies, both amongst those who protest against the use of monetary valuation of the environment and also those prepared to acknowledge that it can make a contribution (Spash 2000).

Bateman and Mawby (2004) illustrated that practical matters such as interviewer appearance (interviewer wearing either formal or more casual clothing) have a significant impact upon stated willingness-to-pay for an environmental good. Alvarez-Farizo and Hanley (2006) found that preferences change significantly when people were given more information, and time to think and discuss, compared to individuals taking part in a conventional valuation survey.

→ Valuation and appraisal tools that do not require monetary valuation of all aspects are more suitable for issues of sustainable development.

Another challenge to cost-benefit analysis arises from the long-term perspective that is necessary in sustainability valuation. It has been argued that with all but very small discount rates, long-term consequences are left out of decisions. When applying the long-term discount factors that are recommended in the Green Book over a period of 100 years (3.5% for years 0–30, 3% for years 31–75 and 2.5% for years 76–100), costs or benefits that are worth £100 today, have a value of merely £5 in 100 years. The result raises issues of inter-generational equity. This concern was addressed in the recent Stern Review (Stern, 2006, esp. section 2A.2) by using a discount rate of 0.1%.

→ Valuation and appraisal tools that use very small or no discount rates are more suitable for issues of sustainable development.

Researchers and policymakers in the UK are increasingly recognising that the interface between science and policy is complex. The old technical-rational model of appraisal in which ‘objective assessment’ was assumed to lead straightforwardly to better decisions, has proved theoretically, politically, and practically inadequate. Instead, attention has been drawn to the complexities of appraisal practices, and to the different, sometimes subtle, ways in which they might secure legitimacy, influence outcomes, and lead to the adjustment of policies. Owens et al., (2004) argue that an important objective for appraisal should be to foster learning of more than one type and, potentially, to modify the belief systems and behaviour of individuals and organisations over time.

→ Valuation and appraisal tools that support social learning processes have more potential to aid decision-making for sustainable development.

In the field of policy analysis the concept of neutral, objective advice is increasingly rejected. Instead it is increasingly acknowledged that the interface between ‘science’ and policy is complex. In this context the framing of problems is of fundamental importance. For example, in the quest to identify which energy (or agricultural, or transport, or waste management) system is the ‘most sustainable’, answers are invariably contingent (Stirling, 1999). When dealing with complex systems there is no uniquely rational way to aggregate different dimensions of value along a single metric (Munda, 2004; Stirling and Mayer, 2001).

Under these circumstances public and stakeholder engagement as well as transparency of the decision-process is essential for increasing legitimacy of decisions.

Natural and human systems are adaptive and impacts of changes in the systems are characterised by uncertainty (Allen 2001; Anderson et al., 1988; Folke et al., 2005; Pimm, 1984). All these factors redefine the role of experts, the meaning of knowledge and how decision processes need to be designed to make more effective policy. Post-normal science argues that under these circumstances we should shift away from the sole focus on outcome towards the quality of decision processes (Funtowicz and Ravetz, 1990). They have also called for more transparent, deliberative, and inclusive processes for informing policy and decisions.
Valuation and appraisal tools that include public and stakeholder engagement and that are transparent tend to perform better in decision-making for sustainable development.

The characteristics of adaptiveness and uncertainty also lead to difficulties in the measurement all impacts in quantitative terms. In situations where important impacts can only be measured in qualitative terms, methods are needed that can make use of both types of knowledge and bring them together in a systematic way.

Valuation and appraisal tools that can draw on both quantitative and qualitative data and bring them together in a systematic way are more suitable for issues of sustainable development.

In the last 10 years researchers and practitioners working in the fields of ecological economics, institutional economics, sustainability science, decision analysis and policy appraisal have sought to address the various challenges laid out above and developed an alternative toolbox to aid more sustainable decision-making and the articulation of public values. Most prominently, the ‘hybrid methodologies’ combine interpersonal deliberation and quantitative methods.

However, sustainability appraisal is not merely a methodological issue. The type of evaluation and the institutional structures in which the appraisal is embedded influences the outcome. Therefore the choice of valuation and appraisal methods is a process one of ‘institutionalising social choice’.

The method of inclusion of environmental resources and ecosystem services in decision processes determines how far the environment is taken into account with results affecting the quality of our lives and those of future generations; the same applies to social aspects of sustainability. Valuation and appraisal methods determine -

- who participates in the decision-making process;
- how they participate and in what capacity (consumer, stakeholder, citizen);
- what counts as data;
- which data processing and aggregation procedures are used.

Therefore valuation methods can be seen as ‘value-articulating institutions’ (Jacobs, 1997; Vatn, 2004). The type of valuation and the institutional structures in which the appraisal is embedded influence the outcome. For example, if an appraisal framework requires impacts to be measured in quantitative terms only, aspects such as landscape aesthetics or community development are likely to be excluded from the analysis. Another example is the extent to which respondents of a valuation study are given time and resources to learn about the issue at hand. More generally, which (methodological or substantive) issues are considered and possibly taken into account depends (i) on the level of familiarity of appraisers with the whole toolbox from which they can draw when conducting an appraisal and (ii) on how results from different appraisal tools would – in their view – be received by their peers and management.

The process of ‘institutionalising social choice’ would be helped by (i) in-depth familiarity with different appraisal tools by those responsible for devising rules for appraisal, (ii) an organisational culture of curiosity to identify the appraisal method that is most suitable for the issues at hand, and (iii) more academic research which compares and contrasts different appraisal tools for specific problem situations.
Alternatives to the ‘technical-rational’ model of appraisal attempt to deal with the problem of ‘institutionalising social choice’ and to establish a ‘socially robust’ (Gibbons 1999) framework for appraisal. Hence, all the methods reviewed in this report build on the following principles –

- accounting for different types of knowledge (monetary and non-monetary; quantitative and qualitative data);
- considering seriously the issue of inter-generational equity;
- providing opportunities for learning during the appraisal process;
- ensuring transparency of each step of the appraisal process; and
- having a strong element of public and stakeholder engagement.
3. Methods and examples

3.1 Deliberative monetary valuation

Definition

The use of formal deliberation concerning an environmental impact to express value in monetary terms for policy purposes, and more specifically as an input to cost-benefit-analysis (Spash, 2001). It differs from cost-benefit-analysis primarily in that preferences are constructed during the deliberation process.

Framing

For example, consider a proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. A group of citizens are selected and meet to discuss information about environmental damages associated with the proposed development. Known costs and benefits (discounted) are presented, while those pertaining to environmental damages are deliberated. The citizens form a jury aiming to provide a monetary value for environmental damages which might be in terms of an individual willingness-to-accept to allow the project to proceed. The result would then be incorporated into a net present value calculation to determine the viability of the project.

Alternatively, deliberative processes can be used to complement monetary valuation techniques. For example, (1) group deliberation can help test the monetary valuation survey design (e.g. wording and comprehension of questions), validate the information content and help identify design biases; (2) deliberative methods can be used for determining the policy options or the institutional context, upon which the valuation survey will be based; or (3) outcomes of an environmental valuation can be validated by a participatory process (Kallis et al., 2004).

The combination of monetary valuation with deliberative processes was advocated by several authors (Brown et al., 1995; Jacobs, 1997; Kenyon and Nevin, 2001; MacMillan et al., 2002; McDaniel et al., 2003; Ward, 1999) to account within monetary valuation more explicitly for the fact that preferences are socially constructed. The notion of value construction suggests that respondents do not have well-defined preferences for many complex environmental goods prior to the elicitation process, but that these preferences are constructed during this elicitation process itself (Gregory et al., 1997; O'Connor, 2000; Payne and Bettman, 1999). Otherwise the assumptions are similar to those of cost-benefit analysis:

1) Consequences are key: A good (moral) act is one from which good consequences arise.
2) Trade-offs are normal: The moral act has good consequences which outweigh bad ones.
3) Values arise from individual preferences: Value is how much an individual prefers a consequence.
4) Values are commensurable and can be aggregated: Good and bad consequences can be added and subtracted for any individual and across all individuals.

Deliberation forums can be used to set the scope of cost-benefit-analysis (boundaries, alternatives considered etc.), define the objectives of cost-benefit-analysis, control the results of cost-benefit-analysis, etc. The combination of deliberative methods and monetary valuation is employed most successfully in parallel, and when recommendations from the cost-benefit-analysis are used in conjunction with the recommendations from the deliberative process.

A criticism of deliberative monetary valuation, which applies in particular to the combined use (rather than the complementary use) of deliberative and monetary techniques, refers to the fundamental differences between deliberative forums (e.g. citizens’ juries) and monetary
valuation (e.g. contingent valuation) (Niemeyer and Spash, 2001). These include the different approaches taken to theoretical foundations (individual and social ontology, preference basis, rationality theory), practical issues (justification, framing, value representation, institutional setting), and political issues (manipulation, representation, social impact). In short, there are significant difficulties in incorporating the views of the public from the deliberative process into the monetary valuation part of the exercise. Whether the combined version of deliberative monetary valuation could therefore be described as an improvement on the contingent valuation approach is currently contested (O’Brien, 2003).

Nature of participation

The deliberative method of deliberative monetary valuation is limited to determine a monetary value, and is therefore primarily instrumental in approach. Thus, it is restrictive in allowing participants to look beyond economic considerations. Difficulties could occur in this type of approach if the citizens’ jury decided to diverge from discussions of economic valuation.

Policymakers need to know how a population will respond to various policy instruments if they are to be effective. When drawing such conclusions, it is important to represent the population adequately. Monetary valuation surveys aim to represent a cross-section of a population, that allow answering questions about the relationships between willingness-to-pay and other socio-economic variables, such as class, gender, income and maybe ethical beliefs, to be answered. The sampling procedure may be either random or stratified sampling. This focus on statistical power has been criticised for failing to offer an adequate account of either interpretation or explanation in the social sciences. Interpretative activity is already implicitly presupposed in arriving at statistical generalisations between socio-economic variables. At the heart of a number of criticisms of the contingent valuation of environmental goods is the claim that it fails to take this interpretative dimension seriously. Deliberative monetary valuation tries to respond to this criticism by deriving data from small-scale deliberative processes, offering the potential for both the interpretative and explanatory depth that large-scale statistical studies lack.

- Deliberative monetary valuation offers interpretative depth - insights into the meaning of the responses offered.
- Deliberative monetary valuation also offers possible explanatory depth - insights into why individuals and groups respond as they do.

The degree to which these benefits are harnessed depends on whether the design intends this data to merely inform the process or to also form part of the final report.

Decision context

Although deliberative monetary valuation includes a certain element of opening up (deliberation about costs and benefits), the method focuses on closing down (calculating a net present value to determine the viability of the project).

Treatment of criteria and values

The impacts are measured by responses from citizens to willingness-to-accept or willingness-to-pay questions. All criteria are represented on a single comparative term (monetary value). Different societal interests are captured by choosing respondents from different socio-economic groups.

Applicability to different types of evidence bases

Risk is calculated using expected values. This involves multiplying the outcome of a given state of the world by its probability, then summing across all probability-outcome products. This treatment of risk valuation has been criticised for some environmental risks, especially those related to very low probability but very high cost events (e.g. toxic waste disposal). Deliberative monetary valuation addresses uncertainty principally through sensitivity and / or scenario analysis. Hanley and Spash (1993) suggest that where uncertainty exists about environmental impacts, there may be a case for weighting environmental costs by a factor greater than that applied to monetary benefits, as a precautionary measure.
Treatment of equity and distributional issues

- **Intragenerational equity and distribution** - The use of market and surrogate market prices as a basis of environmental choice leaves the poor underrepresented, because willingness-to-pay and willingness-to-accept are income dependent. The use of raw willingness-to-pay measures will give greater weight to the preferences of the rich – ‘the poor sell cheap’ (Guha and Martínez-Alíer, 1997). This problem is addressed by weighting responses by income (Brent, 1984; Layard and Walters, 1994; Tol, 2001). Another criticism of monetary valuation is that the distribution of resources and property rights is presupposed; instead they could themselves be the subject of public deliberation (O’Neill, 2001).

- **Intergenerational equity and distribution** – In willingness-to-pay / willingness-to-accept studies, the interests of nonhumans and future generations are only represented indirectly through the preferences of current consumers (O’Neill, 1993). The distribution of costs and benefits between current and future generations is influenced significantly by the choice of the discount rate (Hanley and Spash, 1993, chapter 8; Stiglitz, 1994).

Results

The results can be presented with a focus on the calculated net present value or with a focus on explaining the underlying ethics, belief structures and conflicts that were potentially explored in the deliberative process.

Human and financial resources

The costs for a deliberative monetary valuation study consist of the sum for the deliberative process (organisation and support, payment and expenses of panel and witnesses, hire of venue, costs of facilitation, recording and transcribing proceedings and costs for any related analysis) and of the sum for the monetary valuation exercise (collection of scientific information, conduct of survey, data entry, analysis). Evidently, the team conducting a deliberative monetary valuation study needs to be proficient in both deliberative methods and environmental monetary valuation. A deliberative monetary valuation will therefore have significantly higher costs than a deliberative or monetary valuation study alone.
Key studies

Forest management in Scotland

Kenyon and Nevin (2001) assess the development of a forest project in Scotland. They undertook a contingent valuation study and used a citizens’ jury (CJ) to provide information for policy-makers and describe the advantages in using these two techniques to obtain a broader view. They also suggest that the CJ could very well provide adequate information for policy-makers without the need to resort to monetary valuation.

National park management in Australia

The study by James and Blamey (2004) is concerned with the management of national parks in New South Wales, Australia. This CJ formed part of an examination of the potential use of the technique in environmental decision-making in Australia. The CJ included the following elements: a conventional option choice charge which did not involve deliberative valuation; two approaches to deliberative valuation, involving a further charge in which the jurors were asked to determine societal willingness-to-pay for a specified programme involving environmental improvement; and a stated preference (choice modelling) survey, completed by the individual jurors before and after the CJ proceedings. A process of group discussion and decision-making was used to arrive at a valuation (willingness-to-pay) for a posited programme involving environmental improvement. The study explored different options for the national park management as well as financing these works. The citizens came to a consensus on one of the management options and decided by majority rule on a 0.1 per cent levy on income tax, resulting in an additional A$109.7m (1997/8 income data) for the park management. The study also discusses methodological issues and protocols.

European Water Framework Directive

Alvarez-Farizo and Hanley (2006) combine a citizens' jury with choice modelling (a stated-preference technique) in 'valuation workshops'. The study analyses different approaches for water quality improvements under the Water Framework Directive. The authors compare their results with those of a conventional choice experiment and find that the choice experiment format can be successfully implemented in a ‘jury’ setting. In particular the study found that (a) preferences change significantly when people were given more information and time to deliberate, than when individuals take part in a conventional survey; (b) values, as measured by implicit prices, changed in moving from a conventional survey setting to a valuation workshop; and (c) moving from individual views to collective/citizen views produced changes in both values and preferences.

In summary, cost-benefit analysis was developed for the appraisal of infrastructure projects. Deliberative monetary valuation differs from cost-benefit-analysis primarily in that preferences are constructed during the deliberation process. This method is most suitable for the appraisal of projects whose impacts are rather well understood, where the impacts are relatively short-term, and which do not affect complex ecosystem services such as biodiversity.
3.2 Social multicriteria evaluation

Definition

The combined use of participatory techniques and multicriteria analysis to aid decision making about a number of policy options while taking conflicting interests and multiple criteria into account. It highlights transparency and social learning during the appraisal process (Munda, 2004).

Framing

For example, consider again a proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. The team of researchers starts by conducting an institutional analysis and possibly a historical trend analysis in order (a) to understand the ecosystem of the wilderness area and the socio-economic context of the road construction and (b) to develop policy options and appraisal criteria. For this purpose, they use information from secondary data, focus groups and interviews with stakeholders and citizens. Then the impacts of the policy options are modelled by researchers from relevant disciplines and the results are collected in an impact matrix. The impacts can be measured in quantitative or qualitative terms. This stage is supported by expert discussion groups and interviews. Finally, a ranking of policy options is calculated and presented to stakeholders. Stakeholders are encouraged to interrogate the data, weightings of criteria and resulting rankings. The iterative process can lead to social learning among stakeholders.

The assumptions of social multicriteria evaluation are –

1. Systems are complex: The need for public or stakeholder participation arises from the nature of the problems at hand, from their urgency, their interdisciplinarity, their uncertainty and their irreversibility.

2. Trade-offs are normal: The moral act has good consequences that outweigh bad ones.

3. Values arise from individual as well as group preferences: Value is determined by an interplay of individual and group preferences.

4. Values are incommensurable: Irreducible value conflict is unavoidable but compatible with rational choice employing practical judgement.

The structure of a social multicriteria evaluation consists of six main steps:

1. characterising a wide range of relevant alternative ways to achieve a particular policy aim (‘options’); here social multicriteria evaluation emphasizes the need for institutional analyses to understand well the decision-making context and the need for stakeholder and public participation in order cover the main option as well as ensuring ‘buy-in’ from relevant groups of society,

2. developing a set of ‘criteria’ to represent different viewpoints on the issues that are relevant to the appraising of those options,

3. evaluating options against each criterion based on models or expert judgement from various disciplines (impact matrix) and specifying the preference function for each criterion,

4. assigning a quantitative ‘weighting’ to each criterion, in order to reflect its relative importance under the viewpoint in question.

5. calculating an overall performance rank for each option under all the criteria; this can be presented either as an overall ranking based on group weights or separate rankings for particular viewpoints or individuals.

6. analyzing the potential for conflicts and coalitions between participating stakeholders (‘equity matrix’).
Deliberation among citizens or stakeholders over alternative development options has the potential to generate new ideas. The iterative process of social multicriteria evaluation is flexible and allows for new options to be added as the social learning process proceeds.

Incommensurability does not imply incomparability. The method allows that different options are weakly comparable, that is comparable without recourse to a single type of value. Multicriteria analysis does, however, require a clear distinction between options and criteria and the criteria should be independent of each other.

There are a number of multicriteria algorithms and corresponding software packages available. For a discussion of the different types of algorithms and their respective advantages and disadvantages see for example Dodgson et al., (2000) or De Montis et al. (2005). Most algorithms require the definition of indifference or gradual degrees of preference and these have to be associated to the deviations observed between the evaluations. This is a challenging task for which it is difficult to obtain evidence. Another challenge which is common to most multicriteria algorithms is the establishment of criteria weights; a range of interpretations of weights and corresponding weighting procedures exist (Choo et al., 1999).

The main weakness of sophisticated methods such as NAIADE¹ is their lack of transparency which may lead to difficulties of acceptance among the citizens and stakeholders participating in the MCA process.

Nature of participation

Public and stakeholder engagement starts in social multicriteria evaluation already with the definition of policy options and includes deliberations about criteria, input in the impact matrix and exploration of results with different parameters. Hence, public and stakeholder engagement in social multicriteria evaluation is comprehensive and spans the whole appraisal process. However, all involved need to agree on a common framing (the same criteria against which the same options are assessed).

Decision context

By engaging the public and stakeholders at an early stage and by being transparent throughout the appraisal process, social multicriteria evaluation supports the opening up of relevant issues. The need to agree on a common structure (criteria and options) and the ranking then supports the process of closing down.

Treatment of criteria and values

There are different legitimate values and points of view in society. This creates social pressure to take into account various perspectives of the issue at hand, e.g. economic, political, social, cultural, and religious. The research team translates these dimensions into objectives and criteria. Hence, the criteria in social multicriteria evaluation capture different viewpoints of different groups in society.

Related to this are the questions of who should attach criterion weights and how. To answer this question we have to accept a basic assumption: to weigh different criteria implies to give weights to different groups in society. It is widespread practice to derive weights from deliberative processes, e.g. by visual ranking by importance, by individuals, or by groups. However, Munda (2004) argues that such processes are technically very difficult (e.g., which elicitation method should be used? Which statistical index is a good synthesis of the results obtained? Do average values of weights have meaning at all?) and practically difficult as they

¹ NAIADE was developed by Munda Munda, G. 1995 Multicriteria Evaluation in a Fuzzy Environment: Theory and Applications in Ecological Economics, Heidelberg: Physica-Verlag. at Joint Research Centre (JRC). NAIADE stands for “Novel Approach to Imprecise Assessment and Decision Environments” and is a discrete multi-criteria method, based on the partial comparability axiom and uses pairwise linguistic evaluation of alternatives. In an attempt to avoid overprecision NAIADE uses fuzzy relations, based on “semantic” distance between linguistic qualifiers (e.g. “very good”, “good”, “moderate” etc.). NAIADE applies equal weights to all criteria, because Munda saw major difficulties in deriving weights from stakeholders or otherwise. The software is available on CD from the JRC of the EU Commission and is free for non-commercial use. Version 4.0 is currently being tested at JRC and a new manual is expected for the end of 2006.
could cause conflicts among the various social actors. Instead he suggests using a plurality of ethical principles (e.g., economic development attaching more weight to the economic dimension, precautionary principle giving a bigger weight to the environmental dimension, or sustainability which might imply an equal weighting of all dimensions). There is general agreement in the literature that eliciting weights is the most difficult element of a multicriteria appraisal. But there is still disagreement about the best way to proceed.

Applicability to different types of evidence bases

Data sources can be scientific studies, expert or stakeholder estimates, or traditional knowledge. Many multicriteria algorithms (e.g. Naiade, Promethee, Regime) allow for quantitative as well as qualitative data formats. In most cases qualitative data are translated in fuzzy numbers, i.e. ranges of numbers with no crisp endpoints. Fuzzy numbers are also used to address uncertainty in data.

Treatment of equity and distributional issues

- **Intragenerational equity and distribution** – The impact matrix, which captures the various impacts of each policy option, is complemented by an 'equity matrix' (Munda, 1995). The latter uses information from the preferences of different respondents; it helps to analyse the position of different stakeholders and the potential formation of coalitions among the actors to defend or veto a policy option (representation in a dendrogram of coalitions). As in probably all methods discussed in this review the willingness and capacity to speak and to be heard are unevenly distributed across class, gender, income, and ethnicity (O'Neill, 2001). This poses a particular challenge for the design and facilitation of deliberative processes.

- **Intergenerational equity and distribution** – Discount rates are, if at all, only applied to the financial criteria. In relation to environmental impacts, future generations are considered equal to the current generation.

Sometimes intra- or intergenerational equity is included as a separate criterion making the impacts of policy options among one generation or on other generations more explicit and offering the opportunity to give more importance to this dimension. For example, Munda found in his Palermo case study (Munda, 1995) taking distributional issues explicitly into account increased the transparency of the study and enhanced the interaction with various social actors. However, the number of studies applying this procedure has been small so far.

Results

The results include a ranking of policy options as well as an analysis of the different perspectives about the options held among respondents. Depending on the application and requirements of decision-makers, the ranking may be complete or partial; the latter includes the natural avowal of incomparable alternatives (e.g. one policy option being much better in the social criteria than another which is much better in the environmental criteria). The results

---

2 PROMETHEE is a widely used multi-criteria assessment algorithm Brans, J.-P., Mareschal, B. and Vincke, P. 1986 'How to select and how to rank projects. The PROMETHEE method', European Journal of Operational Research 24(2): 228-38.. PROMETHEE I and II (Preference Ranking Organisation Method for Enrichment Evaluations) aggregate the information by an outranking procedure and finally rank the options (here: scenarios). PROMETHEE I identifies incomparability between options, if the positive and the negative flow of one option is higher than those of another one. This means, that these two options are not comparable. PROMETHEE II uses the net flow (the difference of positive minus negative flows), which delivers a complete ranking of all options. Additional features of PROMETHEE’s the software, Decision Lab, are images of stability intervals, which are intervals of weights throughout which the ranks remain the same. The larger these intervals are the more stable the results. Decision Lab is commercial software and costs currently €990. The REGIME method was developed by Hinloopen et al. Hinloopen, E., Nijkamp, P. and Rietveld, P. 1983 'Qualitative discrete multiple criteria choice models in regional planning', Regional Science and Urban Economics 13: 73–102., assessed and refined by Hinloopen Hinloopen, E. 1985 ‘De Regime Methode’ Interfaculty Actuarial and Econometrics, Amsterdam: Free University Amsterdam. and Hinloopen and Nijkamp Hinloopen, E. and Nijkamp, P. 1990 ‘Qualitative multiple criteria choice analysis’, Quality and Quantity 24: 37-56.. REGIME is a qualitative multiple criteria method. Qualitative multiple criteria methods aim to provide a tool for analysing complex situations, which are impossible to model by means of quantitative information.
should also include a sensitivity analysis and a clear view of the conflicting character of the criteria and the influence of a particular set of weights.

**Human and financial resources**

The costs of a social multicriteria evaluation study consist of the sum for the deliberative process (organisation and support, payment and expenses of panel and witnesses, hire of venue, costs of facilitation, recording and transcribing proceedings and costs for any related analysis) and the sum for the analytical tasks (institutional analysis, modelling, multicriteria analysis, sensitivity analysis). The team conducting a social multicriteria evaluation study needs expertise in deliberative methods, the relevant disciplines for the impact modelling and multicriteria analysis. Social multicriteria evaluation requires considerable amounts of data to evaluate each policy option against each criterion, making it rather resource intense. This requirement can be eased somewhat by accepting rough estimates of impacts and by using flexible tools that can make use of both qualitative and quantitative data.

**Key studies**

**Water management in Sicily**

http://alba.jrc.it/valse/

De Marchi et al., (2000) analysed water shortage in Troina (North-eastern Sicily, Italy) and possible ways to resolve it. They found a complex set of water-related interests present in the case study area. Institutions supporting an effective dialogue were missing. The research team started by identifying the main actors, institutions and other structural issues in relation to the water problem. The following graph lays out the main steps of the appraisal process of the Troina study.

**Figure 1. Scheme of the appraisal process in the Troina case study**
The team explicitly designed the appraisal exercise as a social learning process for stakeholders. Eight policy options were assessed against nine criteria by use of an outranking\(^3\) algorithm called NAIADE. Due to time and resource constraints the impact matrix in this study was limited to qualitative expert assessments. The NAIADE software is also unique in offering a conflict analysis feature that visualises groups whose interests cluster in a dendrogram of coalition formation (see Error! Reference source not found.). Another way to account for the different interests, also applied in this study, is the comparison of rankings of policy options by actor groups in addition to calculating an overall ranking.

Figure 2. Dendrogram of the coalition formation process

---

Tourism and national park management in Spain
http://einstein.uab.es/c_ceambientals/Diafanis/default.html

Martí et al., (2000) developed and appraised options for tourism development in the buffer zone of the Aigüestortes i Estany de Sant Maurici National Park (Spain). The eight development options were appraised against 13 economic, social and environmental criteria within a transparent and participatory decision-making framework. For this purpose, the DIAFANIS project combined historical trends analysis, institutional analysis, GIS, dynamic systems modelling, financial analysis and multi-criteria analysis. Like the previous study, DIAFANIS used the NAIADE multicriteria algorithm, but it differed insofar as the impacts were modelled by the use of STELLA dynamic systems modelling and financial modelling. For a short description of the study in English see the project webpage or Munda (2004).

Water management in Spain
http://ecoman.dcea.fct.unl.pt/projects/advisor

As part of the ADVISOR project, Moral et al. (2006) study and evaluate alternative policies for improving the water supply for the Costa del Sol (Spain). They applied social multicriteria evaluation as an implementation of integrated evaluation and participation objectives of the Water Framework Directive. Eight policy options were evaluated against eleven socio-economic and environmental criteria. To understand the problem and the context, the research team combined in-depth interviews, questionnaires and observation at open meetings, with analysis of secondary material. The interviews and questionnaires were also used to identify policy options and criteria. On the basis of technical information and an opinion poll conducted to evaluate the ‘degree of social acceptance’ criterion, the impact matrix was completed. As in the previous social multicriteria evaluation studies, NAIADE was used to calculate the ranking and well as a dendrogram of stakeholder coalitions. The results showed that the demand-management options were the highest scoring ones. At the end a focus group meeting brought together the participating stakeholders to review the results; in this case, the inclusion of new policy options was suggested, which were then assessed against the same criteria and the ranking was reviewed again. The research team considered the Costa del Sol social multicriteria evaluation as a successful mutual learning experience for the participating stakeholders and the research team. In their view, the process also contributed to the clarification of conflicts and the creation of conditions necessary to construct a more permanent dialogue process among the stakeholders involved and to work towards a resolution.
Renewable energy in Austria
http://www.project-artemis.net

Kowalski et al. (2006) combined scenario building and social multicriteria evaluation in the context of renewable energy use in a multi-scale study. By projecting different possible pathways into the future, scenarios play an important role in decision-making about long-term consequences. Scenario analysis is a common approach to account for more complexity of real-world situations than analysing individual projects, technologies, or actions could do. Five renewable energy scenarios for Austria for 2020 were evaluated against 17 qualitative and quantitative socio-economic, environmental, and technical criteria. An equivalent process was undertaken on the local level, where four renewable energy scenarios were evaluated against 15 criteria. On both scales, the scenario development consisted of two stages: first an exploratory stage and second a modelling stage with forecasting-type scenarios. The scenarios consist of a narrative part (storyline) and a modelled quantitative part. The weights were derived in interviews and group workshops, respectively, from Austrian energy stakeholders on the relevant scales. The rankings for the various stakeholders were calculated by use of the outranking algorithm PROMETHEE I (see Error! Reference source not found.). In this study importance weights were applied on the basis of individual and group preferences. The stakeholder learning experiences were studied in depth (Garmendia and Stagl, 2006).

Figure 3. Typical ranking of scenarios (national level)

In summary, this method is most suitable for the appraisal of policies, programmes or projects whose impacts are not yet well understood and therefore benefit from a multidisciplinary modelling of impacts. Social multicriteria evaluation was developed to address complex issues and to deal with uncertainty in the context of sustainable development.
3.3 Three-stage multicriteria analysis

Definition

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy options while taking conflicting interests and multiple criteria into account. Here the sequencing and choice of participants is based on Renn’s and Webler’s (1998; 1993) ‘co-operative discourse’ model. In this model, stakeholders select the evaluation criteria, experts present information and measure impacts, and citizens explore values.

Framing

For example, consider again a proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. Stakeholders, representing different interests (e.g. economic, political, social, cultural, and religious), select concerns and related criteria that they consider relevant when evaluating options to deal with the road project. Experts translate these criteria into measurable indicators, identify different policy options, present information and assess the performance of each policy option against each of the criteria. In a final stage, citizen panels consider the resulting information and explore values and preferences. In this process the citizens are encouraged to compare the results from the numerically based analysis with their own or their group’s holistic judgement. Differences between these two assessments are discussed before a final recommendation is made.

The assumptions of three-stage multicriteria analysis are –

1. Systems are complex: The need for public and stakeholder participation arises from the nature of the problems at hand, from their urgency, their interdisciplinarity, their uncertainty and their irreversibility. In the participatory process specific roles are assigned to stakeholders, experts and citizens.

2. Trade-offs are normal: The moral act has good consequences that outweigh bad ones.

3. Values arise from individual as well as group preferences: Value is determined by interplay of individual and group preferences.

4. Values are commensurable: Values can be captured in a utility function. However, conflicts about values prevail and therefore deliberation complements the analysis based on utility theory.

The appraisal process that is based on the ‘cooperative discourse’ entails the following three steps (Renn, 2006):

1. Identifying and selecting concerns and evaluative criteria.

All relevant stakeholder groups are asked to reveal their values and criteria for judging different options (step 1 in Error! Reference source not found.). It is crucial that all relevant value groups be represented and that the value clusters are comprehensive and include economic, political, social, cultural, and religious values. A value-tree analysis is used to elicit the values and evaluative criteria used for judging different options (von Winterfeldt, 1987). The purpose of a value-tree is to elicit and represent the concerns of all relevant stakeholder groups. It structures the elicited values, criteria, and corresponding attributes in a hierarchy, with general values and concerns at the top, and specific criteria and attributes at the bottom. The aim is to represent the concerns of all affected parties in a list of hierarchically structured values. Depending on the political context and the nature of the decision to be made, the values of the various stakeholder groups may vary considerably. By giving each group the right to assign a weight of zero to each criterion that they regard irrelevant, it is possible to construct a joint or combined value-tree that accounts for all viewpoints and that can be verified by all participants.
To avoid strategic behaviour by stakeholders, Renn (2006) argues that the actual measurement of impacts should be left to a group of independent experts and weighting to an unbiased jury of uncommitted citizens. However, stakeholder groups may inform the experts about potential impacts they expect as a result of any one option and they can contribute their evaluation of these options to the citizen panel in their testimony.

2. Identifying and measuring the impacts and consequences related to different policy options.

The research team or an external expert group operationalises and transforms the evaluative criteria derived from the value-tree into indicators (step 2 in Error! Reference source not found.). The participating stakeholder groups review these operational definitions and indicators. Once approved by all parties, the indicators are used to evaluate the performance of each policy option on all value dimensions. Experts from relevant academic disciplines and with diverse perspectives on the topic of the discourse are asked to judge the performance of each policy option on each indicator. Webler et al. (1991) developed a modified Delphi method for this purpose. This method deviates from the original Delphi format by replacing written responses by group interactions. The objective is to reconcile conflicts about factual evidence and reach an expert consensus via direct confrontation among a heterogeneous sample of experts. The desired outcome is a specification of the range of scientifically
plausible and defensible expert judgements and a distribution of these opinions among the expert community with verbal justifications for opinions that deviate from the average (median) viewpoint. This information is used to complete the impact matrix.

3. **Conducting a discourse with randomly selected citizens as jurors and representation of interest groups as witnesses.**

The final step is the evaluation of potential solutions by one group or several groups of randomly selected citizens (Dienel, 1989; Dienel and Renn, 1995). These panels are asked to evaluate and design policy options based on knowledge of the likely consequences and their own values and preferences. The participating citizens are informed about the options, evaluative criteria, and performance profiles of options. Their involvement helps to elicit values and assign relative weights to the different value dimensions (step 3 in Error! Reference source not found.). The procedures used for this purpose are derived from Multiattribute Utility Theory (von Winterfeldt and Edwards, 1986). The participating citizens are first asked to use the criteria of the joint value-tree to rate each decision option on each criterion. Participants are free to add new values to the tree, but they may not delete any of the criteria elicited from the stakeholder groups. They can also modify the presented options or add options to the list. The rating of each option then proceeds on the basis of the profiles that the experts generated during the Group Delphi. Finally, each criterion is weighted against all other criteria resulting in a matrix of relative weights and utility measures for each option and each criterion.

Both tasks, the transformation of the expert data in utilities and the assignment of trade-offs, are performed individually and in small groups. Deviating from the established MAU (multiattribute utility model) procedure, the numerical results (i.e. for each option the sum over the utilities of each dimension multiplied by the weight of each dimension) of the decision process are not used as expression of the final judgement of the citizen participant, but as a structuring aid to improve the participants' holistic and intuitive judgement (Renn and Webler, 1998). Instead of breaking the options down into specific attributes, the options are assessed as a whole. By pointing out potential discrepancies between the numerical model and the holistic judgements, the participating citizens are forced to reflect upon their opinions and search for potential hidden motives or values that might explain the discrepancy. The final recommendations are always based on a holistic judgement by individuals or groups.

The process of this stage is similar to a jury trial with experts and stakeholders as witnesses and advisers on procedure as ‘professional’ judges (Crosby *et al.*, 1986). The representatives of interest groups and the experts take part in the process as witnesses; they provide their arguments and evidence to the panels who ultimately decide on the policy options. This deliberation process takes time; citizen panels are conducted as seminars over three to five consecutive days or over a longer period of up to six months. All participants are exposed to a standardised programme of information, including hearings, lectures, panel discussions, videotapes, and field tours.
The main strengths of utility-based multicriteria appraisal methods are their simplicity and strong theoretical grounding. However, being based on utility theory and rational expectations, these algorithms are subject to the same critique as the theories on which they are built (strong assumptions about preference functions and commensurability of values). Initially utility-based MCA algorithms were applied for production planning and financial portfolio choices and geared towards individual decision-makers. More recently they have also been applied in environmental decision-making and extended to group decision-making (Beinat and Nijkamp, 1998).

**Nature of participation**

Stakeholders, experts, and citizens have clearly defined roles in the “Three-Stage-Multicriteria Analysis”. Stakeholder groups are asked to help identify and select concerns and evaluation criteria. Experts present information and are engaged in a process of identification and measurement of impacts of different decision options. Citizens are asked to consider the resulting information and to explore values and preferences. While public and stakeholder engagement also spans the whole appraisal process, the contributions are here more compartmentalised. This may make the participatory process more efficient, but it may also reduce the opportunities for social learning. Like in the case of social multicriteria evaluation, the participants need to agree on a common framing in the final stage (same criteria against which the same options are assessed).
Decision context

By engaging stakeholders at the initial stage and citizens at a later stage and by being transparent throughout the appraisal process, the “Three-Stage-Multicriteria Analysis” supports the opening up of relevant issues, but to a lesser extent than social multicriteria evaluation, deliberative mapping or multicriteria mapping. The need to agree on a common structure (criteria and options) and the ranking then supports the process of closing down.

Treatment of criteria and values

There are different legitimate values and points of view in society. This creates social pressure to take various perspectives of the issue at hand into account, e.g. economic, political, social, cultural, and religious. The research team or external experts translate these criteria into measurable indicators. As with social multicriteria evaluation, the criteria capture different viewpoints of different groups in society.

Related to this are the questions of who should attach criterion weights and how. To answer this question we have to accept a basic assumption: to weigh different criteria implies to give weights to different groups in society. In the three-stage multicriteria analysis, each criterion is weighed against all other criteria resulting in a matrix of relative weights for each option performed individually and in small groups. Hence, weights are indifference trade-offs (Choo et al., 1999).

Applicability to different types of evidence bases

Data sources can be scientific studies or expert judgements. Both, the multi-attribute utility framework and the holistic appraisal allow the use of quantitative as well as qualitative data formats.

Citizen participants and decision-makers are encouraged to interact directly with experts to get an impression of the uncertainty about consequences, ambiguity of the knowledge base, and dissent among experts. The three-stage multicriteria analysis also includes the technique of the Group Delphi which aims explicitly to explore uncertainty about data, to reconcile conflicts about factual evidence, and agree on a range of scientifically plausible and defensible expert judgements.

Treatment of equity and distributional issues

This is similar to that of Social Multicriteria evaluation, but the three-stage multicriteria analysis does not have an ‘equity matrix’. Instead distributional issues are accounted for by asking all relevant stakeholders and experts to help identify concerns and criteria. Also, like with the previous method, intra- or intergenerational equity is sometimes included as a separate criterion. For example, in the CoRWM study about radioactive waste the ‘burden on future generations’ was one of the eleven headline-criteria. This criterion was detailed by four sub-criteria: costs (reduce the financial liability – whole life costs – imposed on future generations), effort (reduce managerial effort for all aspects of implementation imposed on future generations, including the pre-operational phase), worker dose (reduce exposure of the workforce imposed on future generations), and environmental impact (reduce the environmental burden imposed on future generations, taking into account pollution, physical disturbance, use of natural resources, visual impact, noise, transport and surface land use).

Results

The results include a ranking of policy options as well as an analysis of the different perspectives about the options held among citizen participants and possibly among stakeholders and experts. The calculated ranking is complete, but this result is compared with a more holistic appraisal. The final recommendations are always based on a holistic judgement by individuals or groups and include a ranking and the reasoning for this ranking. The results should also include a sensitivity analysis and a clear view of the conflicting character of the criteria and the influence of a particular set of weights.
Human and financial resources

The costs for a three-stage multicriteria study consist of the deliberative processes with stakeholders, experts and citizens (organisation and support, payment and expenses of stakeholders, experts and citizen panel, hire of venues, costs of facilitation, recording and transcribing proceedings and costs for any related analysis) and of the sum for the analytical tasks (multicriteria analysis and sensitivity analysis). The team conducting a three-stage multicriteria study needs expertise in deliberative methods, access to experts from relevant disciplines and expertise in multicriteria analysis. Three-stage multicriteria analysis requires consulting repeatedly a wide range of experts and engaging with stakeholders and citizens repeatedly, which makes it rather resource intense.

Key studies

Waste management in Germany

Schneider et al. (1998) conducted a study in the South Germany that aimed to support the decision about a waste-management programme. A round table with 16 major stakeholder groups was organised in 1994 to develop waste reduction policies and to assess the recycling potential of the area (using the method of value-trees). The same group also was asked to find the most suitable technical solution for waste processing before final disposal. After these decisions were made, 200 randomly selected citizens from potential host communities were asked to find the most appropriate site for the types of facilities that had been previously chosen by the representatives of the round table. Panelists were even found to be willing to approve a siting decision that would affect their own community. All ten citizen panels reached a unanimous decision which involved the recommendation to construct a small state-of-the-art incinerator in the centre of the most populated town within the region. The reason for this surprising recommendation was that citizens wanted to present a visual reminder to their fellow citizens not to forget the need to reduce waste, to burden those who contribute most to the problem and to put the incinerator near the political power centre.

Radioactive waste management in the UK

http://www.corwm.org.uk

Recently, a three-stage multicriteria analysis was commissioned by the Committee for Radioactive Waste Management (CoRWM) to appraise the performance of the shortlisted options for the UK nuclear waste inventory (Phillips et al., 2006). The short-list of options and assessment criteria were informed by public, stakeholder and specialist views. In the value-tree the views and concerns were arranged under an initial set of broad headings that, at lower levels, are broken down into more specific issues. There were 11 ‘headline criteria’ and 27 ‘sub-criteria’. The policy options were formally assessed on the sub-criteria level. The assessment of the options was carried out for each waste stream. To ensure that specialist knowledge was fed into the process, the assessment of option performance against the criteria (‘scoring’) was done by experts. CoRWM convened specialist workshops for each main group of criteria. Participants came from the industry, regulators, consultancies, academia and NGOs. Briefing material was commissioned for each workshop. The judgements of criteria weighting was done by CoRWM’s Integration Group, informed by public and stakeholder views. The weighting procedure differed in this case study from the trade-offs that are usual in multi-attribute utility theory. While criteria have an intrinsic value or importance, the weights they are given in a relative assessment of options needs to take into account how much of a discriminator they are between these options. Hence, a ‘swing weighting’ process was applied (Hiview4 software based on M-MACBETH5).

4 Hiview is a commercial software based on MACBETH, which currently costs £850.
5 MACBETH (Measuring Attractiveness by a Categorical Based Evaluation Technique) is a multicriteria analysis method that requires only qualitative judgements about differences of value to help decision-maker(s) quantify the relative attractiveness of options. The approach, based on the additive value model, aims to support interactive learning about the evaluation problem and the elaboration of recommendations to prioritise and select options in individual or group decision making processes. M-MACBETH is a tool that converts a verbal descriptive scale into a numerical one for multicriteria analysis.
This involved making judgements on the relative importance of sub-criteria, once the
difference between the top and bottom of the scoring scales for those criteria was taken into
account. The conclusions drawn from the multicriteria analysis were: (1) Overall, geological
disposal options ranked higher than storage options. (2) The difference in ranking between
geological disposal and storage was substantial for most waste streams and for most of the
limiting case sector scenarios. (3) Generally, the borehole option was the lowest ranked
geological disposal option. Error! Reference source not found. summarises the
multicriteria process. A sample output of results can be found in Figure 7.

Figure 6. Overview of the CoRWM multicriteria process

In parallel to the multicriteria process, the team devised a holistic assessment based on
intuitive judgement of the options as a whole. The provisional preferences of the majority of
members taking part in the holistic assessment for some form of geological disposal for most
waste categories was consistent with the multicriteria results. Details about the results from
the different assessment strands, how they were brought together to a recommendation and
comments on these recommendations can be found in chapter 12 of CoRWM’s final report.

Figure 7. Overall result for the high-level waste model
In summary, this method is most suitable for the appraisal of policies, programmes or projects whose impacts are reasonably well understood by experts and which contain a significant technical element. Three-stage multicriteria analysis was developed for the social appraisal of technologies with particular emphasis on the role of risk and uncertainty.

### 3.4 Multicriteria mapping

**Definition**

An interview-based multicriteria analysis whereby individual specialists and stakeholders are invited to appraise the performance of core and discretionary options against their own sets of criteria. It focuses on eliciting and documenting detailed technical and evaluative judgements concerning the performance of alternative policy options. It devotes particular attention to the systematic exploration of uncertainties and the sources of variability between diverse viewpoints (Stirling, 1997).

**Framing**

For example, consider again the proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. Participants are recruited on the basis of a stakeholder analysis. The recruitment is conducted in a way that seeks to reflect a broad ‘envelope’ of relevant perspectives and to represent in some detail the main relevant dimensions in the policy debate. The team of researchers develops a set of core evaluation criteria and core options, which all participants are asked to consider. Participants are then guided by a researcher through their own individual analysis in separate 2-3 hour sessions. As part of this session the participant is asked to consider whether they would like to add any criteria or options that might be relevant for them for appraising the proposed road through the wilderness area. During the session, the researcher works interactively with a piece of computer software (MC Mapper⁶) to explore the performance of options, against their criteria, under different assumptions. In addition to the quantitative and textual documentation recorded using the software package, the interviews are also audio-recorded for later transcription and analysis. From the outset, the aim is not to achieve a consensus on how to proceed on the road proposal, but to expose the variety of views and to try to understand where the differences are most marked and why.

The assumptions of multicriteria mapping are –

1. **Systems are complex**: The need for public participation arises from the nature of the problems at hand, from their urgency, their interdisciplinarity, their uncertainty and their irreversibility.
2. **Trade-offs may or may not be possible**: Ethical or other reasons may prevent the weighing up of good consequences against bad ones.
3. **Values arise from individual as well as group preferences**: Value is determined by the interplay of individual and group preferences. These can be captured by individual stakeholder interviews.
4. **Values are incommensurable**: Irreducible value conflict is unavoidable, but compatible with rational choice employing practical judgement.

The structure of multicriteria mapping consists of six main steps, which are also summarised in Error! Reference source not found.:—

1. characterising a wide range of relevant alternative ways to achieve a particular policy aim (‘options’);

---

⁶ MCM Mapper is a software package that was developed by A Stirling at SPRU. It records during the interview the options and criteria (core and additional), the scores for each option using each criteria (one under the most favourable assumptions, one under the most pessimistic conditions) and the criteria weighting; it also allows to review the final picture of option performance and to make any necessary adjustments. The software is free for non-commercial purposes.
2. developing a set of ‘criteria’ to represent different viewpoints on the issues that are relevant to the appraisal of those options;

3. evaluating each criterion in turn with numerical ‘scores’, to reflect the performance of each option under each criterion for a given viewpoint;

4. exploring uncertainties in the data (by asking respondents for scores under optimistic and pessimistic conditions) and ambiguities in the assumptions (by analysing qualitative data collected during the interviews);

5. assigning a quantitative ‘weighting’ to each criterion that reflects the relative importance of their criteria to the interviewee. In contrast to the relatively technical business of scoring, this weighting process reflects intrinsically subjective judgements over priorities and values;

6. calculating an overall performance rank for each option under all the criteria taken together for a particular viewpoint. Multicriteria mapping uses the ‘linear additive weighting’ procedure, in which the rank simply represents the weighted sum of normalised scores. Next, the ranks and other elements of the appraisal process are considered. After seeing the ranking of options, participants are free to alter their weightings or scores in the light of this, with the objective of arriving at a final overall pattern of ranks, which they feel comfortable accurately represents their personal perspective. Sometimes, this review prompts participants to define new options or criteria, or even to reconsider aspects of scoring. In such cases, the interviewer should encourage the participants to justify their reasons for any changes.

Figure 8. Overview of the multicriteria mapping process

Source: Stirling (1997)
One of the advantages of multicriteria mapping is the simplicity of the algorithm and the subsequent transparency of the analysis process. Multicriteria mapping avoids the distinction between impacts and preference functions, which simplifies the algorithm and might increase the ‘buy-in’ of participants.

Rather than seeking to produce a single aggregate ‘answer’, the multicriteria mapping tool is used to explore how differing assumptions, priorities and value judgements shape participants’ individual appraisals. Depending on the viewpoint, this additional feature of multicriteria mapping may be seen as an advantage or as a disadvantage.

Nature of participation

The participatory process in multicriteria mapping differs significantly from those devised in other methods in this review because stakeholders are interviewed individually. It is assumed that, because of their role in an organisation related to the issue in question, they would have deliberated with others about the most important aspects of the issue before the interview. This is a characteristic that cannot generally be assumed for citizens.

In multicriteria mapping, the participants do not need to agree on a common framing, i.e. individual respondents can add criteria and options during the appraisal process, and these will be reported together with the appraisal of core criteria and core options.

Decision context

Multicriteria mapping focuses as much on ‘opening up’ as on ‘closing down’ a decision or policy process. It uses the appraisal process as a way to gain a systematic picture of the precise way in which different perspectives vary on the issues and options in question. This generates a rich body of information concerning the reasons for differing views, as well as their practical implications for the overall performance of the selected options. In this way, multicriteria mapping tries to span the divide between narrow quantitative methods (which directly address decision priorities, but which may be insensitive to wider considerations) and broader qualitative approaches (which can accommodate more diverse perspectives, but can have difficulty focusing on the context of the decision). Particular features of multicriteria mapping that allow this unusual combination include:

1. a core set of diverse options are precisely defined in advance by the research team for purposes of comparison, but participants are also free to redefine those options or add additional ones;

2. participants are entirely free to choose and define their own criteria (rather than having these imposed upon them), but this does not affect the comparability of the final results (which are in terms of ‘performance’);

3. careful attention is given to the exploring and documenting of ‘uncertainties’ – the way in which performance may vary for any individual participant, depending on assumptions or context;

4. a clear picture of performance is given under each individual viewpoint and the method does allow these to be aggregated across groups of participants or all participants taken together. But the primary focus is on exploring the resulting ‘map’ of how option performance varies across perspectives, rather than on revealing a single uniquely definitive view (Millstone et al., 2006).

By combining a tight focus on decision options whilst at the same time ‘opening up’ the practical implications of different real-world perspectives, multicriteria mapping tries to avoid a serious – but often neglected – problem suffered in common by economic, decision and risk assessment techniques, as well as by many more qualitative deliberative and participatory approaches. This problem concerns the way in which such methods claim, aspire or are interpreted to provide a single uniquely robust, rational or legitimate picture of option performance, irrespective of the divergent uncertainties, interests, priorities, and values associated with different expert and socio-political perspectives. Where they are used to ‘close down’ policy debates, such methods are being employed in a fashion that undermines their own fundamental founding principles of rationality or inclusion. To the extent that it
avoids such untenable attempts at ‘closing down’, multicriteria mapping is free to adopt the most straightforward of theoretically valid mathematical procedures used in decision analysis, thus enhancing the important qualities of accessibility to participants and transparency to third parties (Millstone et al., 2006).

**Treatment of criteria and values**

There are different legitimate values and points of view in society. This creates social pressure to take various perspectives of an issue into account, e.g. economic, political, social, cultural, and religious. Stakeholders who represent different interests in society express these values and points of view in the multicriteria mapping interviews.

Related to this is the issue of weighting. In multicriteria mapping participants are asked to indicate the relative importance of each of their appraisal criteria by means of a simple numerical weighting. Taken together, these weightings reflect the relative importance of the criteria to the interviewee. In contrast to the relatively technical process of scoring, this weighting process reflects intrinsically subjective judgements over priorities and values.

**Applicability to different types of evidence bases**

Since participants provide the scores and weights, these are the main data sources in multicriteria mapping. Both, quantitative and qualitative information is recorded (MCM Mapper software) and used in the analysis (MCM Analyst software†). The ranks are enriched by the results from the analysis of qualitative information recorded during the interview.

Interviewees are asked to provide ‘best’ and ‘worst’ performance scores. This captures the degree of uncertainty and variability around the performance of particular options under a given criterion. Participants are asked to talk about their assumptions behind these different scores, and this qualitative data is recorded and transcribed. This captures uncertainty about how well the option will actually work, variability within the option, and sensitivity to wider context conditions. This leads to rankings expressed as ranges of values instead of single numbers. These ranges express uncertainty in assigning scores, differences of opinion and variability in performance from context to context, for instance, the differences between good and bad implementation, or between appropriate and inappropriate applications.

**Treatment of equity and distributional issues**

The method treats equity and distributional issues in a similar way to Social Multicriteria Evaluation and Three-Stage Multicriteria Analysis, but individual participants chose whether to include such issues in their analysis when selecting their criteria.

**Results**

The results include rankings of policy options either per stakeholder, per perspectives (i.e. groups of participants), or averages of all pessimistic (left-hand end of bar) and of all optimistic (right-hand end of bar) ranks (i.e. combined weighted scores for all criteria) for core options and for additional options. The interpretation of the rankings is helped by the textual analysis of statements recorded in the software during the interview and from the interview transcripts. The results should also include a sensitivity analysis and a clear view of the conflicting character of the criteria and the influence of a particular set of weights.

**Human and financial resources**

The costs for a multicriteria mapping study consist of the sum for interviewing stakeholders (2-3 hours per interview, recording and transcribing proceedings and costs for any related analysis) and of the sum for the analytical tasks (multicriteria analysis and sensitivity analysis). The team conducting a multicriteria mapping study needs basic expertise in

† MCM Analyst is a software package, which was recently developed by A Stirling at SPRU. It includes a central database containing all data relating to all participants, interlinked with text reports for representing in narrative form various permutations in the qualitative data and a spread sheet to process and present quantitative data in the form of charts. The software allows the data to be examined individually as well as by perspectives (i.e. groups of participants), issues (i.e. groups of criteria) and/or clusters (i.e. groups of options). The software is free for non-commercial purposes.
interview techniques and expertise in multicriteria analysis. Compared to the other methods, resource requirements for multicriteria mapping are low.

Key studies

Genetically modified crops and foods in Europe

Stirling and Mayer (2001) studied the controversial issue of introducing genetically modified (GM) crops and foods in Europe. Claims of unprecedented economic benefits are qualified by concerns over the potential for serious irreversible harm. However, there is considerable scientific uncertainty over the form and magnitude of possible effects and, as yet (in contrast with chemical or nuclear risks), little accumulated practical experience to draw upon. This has led to the evolution of a set of controls which are intended to be `precautionary' in nature – accepting that action to avoid harm may be taken in the absence of scientific proof. But the regulatory appraisal process has failed to gain confidence, either on the part of NGOs, private industry, or the general public. This lack of confidence has arisen because, among other things, the scope of the regulatory appraisal is disputed.

Drawing on a variety of perspectives in the current UK debate, a range of agricultural strategies for the production of oilseed rape, including both GM and non-GM options were explored in this multicriteria mapping study. Participants were asked to consider and appraise six core policy options, which were identified and defined in advance by the researchers. They were: organic agriculture, integrated pest management (IPM), conventional agriculture, GM oilseed rape with segregation and present systems of labelling, GM oilseed rape with post-release monitoring, GM oilseed rape with voluntary controls on areas of cultivation. Participants were able to add up to six further options which they were free to define. Twelve participants were interviewed individually. The total set of appraisal criteria (117), which were developed by participants, reflects a wide range of considerations viewed from a disparate array of perspectives.

The criteria were ordered by the researchers into six groupings: `environment', `agriculture', `health', `social', `economic', and `other' issues. By contrast with many multicriteria analyses, participants were left relatively free to undertake the weighting process in whatever way they felt most comfortable, with the interviewer providing guidance where requested. Starting from a default position where equal weighting was assigned to each criterion, participants usually began by ordering the criteria simply in sequence of their relative importance. Starting with the least and the most important criteria, the intensity of the differences in importance between pairs of criteria were then addressed by altering the weightings criterion by criterion. This continued in an iterative fashion until a final set of weightings was arrived at with which the participant felt comfortable. Final rankings assigned by participants were calculated and plotted in bar diagrams (see Figure 9).

Stirling and Mayer found that the viewpoints of the different participants result in very different ranking orders across the six basic options. They also found uncertainties to be very important; under many perspectives, the worst options at their best rank were higher than the best options at their worst. A number of key findings emerged. First, the GM options clearly perform best overall under the perspectives of only three participants, who are all associated with government or industry bodies. Second, under the perspectives of two participants associated with government or industry bodies, the position was more equivocal, with non-GM options (notably organic) performing better under certain conditions. Third, the voluntary controls regime in general performed indifferently or worst among the regulatory strategies for GM crops under the perspectives both of industry and of public interest group participants. Fourth, the charts display some interesting idiosyncrasies in the rankings derived by individual participants. For example, the pattern arrived at by one participant ascribes a maximum rank to all GM options alike, a minimum rank to all other options, and no uncertainty. This is a result of the scoring of these options under a human health criterion, coupled with a weighting of 100% on that criterion. Finally, the conventional intensive cultivation option tends to perform rather poorly under all perspectives, both with respect to the GM options and with respect to the organic and IPM options (depending on the perspective).

Figure 9 Divergent views of risks of different agricultural options
McDowall and Eames (2006) used multicriteria mapping to delineate alternative configurations of a ‘sustainable hydrogen energy economy’, within a wider global context and realistic timescale. For this purpose, they combined backcasting with multi-criteria appraisal.
Deeply contested values embodied in the public debate suggest that a consensus view of a sustainable hydrogen economy may be impossible at present. Hence, the project team needed an open and transparent process for building scenarios, and used participatory approaches as a route to ‘social learning’. The team developed six ‘visions’ - central pipeline, forecourt reforming, liquid hydrogen, synthetic liquid fuel, ubiquitous hydrogen and electricity storage. They used a multicriteria mapping approach that was adapted to long-term scenarios. The adaptations brought the inclusion of a reference or ‘status quo’ vision, ex-ante sustainability grouping of criteria (environmental, social, economic, security/diversity) and qualitative exploration of ‘sideswipes’ (rapid climate change scenario and sustained oil and gas crisis scenario).

The multicriteria mapping aims were: (1) capturing quantitative and qualitative information about expectations of key stakeholders, (2) mapping divergent values, exploring divergent problem framings and (3) systematically exploring technological uncertainty. 15 expert stakeholders were interviewed. Participants developed in total 98 criteria. 3 participants appraised additional hydrogen futures, which were largely variations on the core visions. Aggregated ranks were calculated, but presented with the warning that these should be interpreted with caution as the emphasis of multicriteria mapping is to move away from aggregation. While there are significant overlaps, the electricity store vision performs best (offset by concerns about feasibility) and forecourt reforming does worst (performs poorly on carbon emissions). The appraisals were analysed by participants, perspectives and criteria groupings. The main findings include: hydrogen is not automatically a sustainable option; there is significant uncertainty over future performance of the technology; the sustainability of alternative hydrogen futures is deeply contested; there is a wide range of rationales for ordering the different options; differences in framing assumptions help explain variability of scores (e.g. about fossil fuel supplies and costs, likelihoods of technological success, consumer behaviour, etc); and that nuclear power is a key area of disagreement.

Obesity in Europe
http://www.sussex.ac.uk/spru/porgrow

The PorGrow project conducted a systematic process to identify key public policy interventions that could be implemented in response to the rising trend in the incidence of obesity in Europe. Millstone et al. (2006) used multicriteria mapping to gather quantitative and qualitative data from representatives of a broad range of organisations representing relevant stakeholder interest groups in nine EU Member States and analyse their varying perceptions and judgements.
During structured interviews (190 in total), stakeholders were invited to consider a set of 20 policy options and to propose additional policy options if they wished, and then to score those options assuming either favourable conditions or unfavourable conditions for the context in which the options could be implemented. Stakeholders provided scores by reference to criteria of their own choosing, and they were asked to describe these criteria and to provide relative weights to their criteria. The interviews were recorded and transcribed to provide qualitative data indicating the reasons for the choice of criteria, scores and weightings. The research team then analysed the data gathered in the interviews, and set the results in the context of the rising incidence of obesity in the Member States, the changing patterns of food consumption and physical activity, and the current debates about policy responses to obesity in the Member States and in the European context. The findings indicate that a comprehensive portfolio of policy measures, integrated into a coherent programme, would be well-supported by broad coalitions of stakeholders.

In summary, this method is most suitable for the appraisal of policies, programmes or projects to which stakeholders have had some exposure and where views not only about data, but also about options and criteria, are controversial. Multicriteria mapping was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

3.5 Deliberative mapping

Definition

The combined use of participatory techniques and multicriteria analysis to aid decision making about policy options while taking conflicting interests and multiple criteria into account. Besides measuring the specific performance of each option against the criteria, it highlights the need to explore the arguments participants use to justify their judgements. Specialists and small groups of citizens follow the same assessment process (Davies et al., 2003).

Framing

For example, consider again the proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. First, the team of researchers conducts open-ended interviews with specialists and stakeholders in order to understand the ecosystem of the wilderness area and the socio-economic context of the road construction. The team of researchers develop a set of core evaluation criteria and core options, which all participants are asked to consider. Then, facilitated group discussions with small groups of citizens help to clarify, discuss and then agree meanings, definitions and implications of the options to be appraised. The groups discuss and agree a shared set of criteria to be used by the citizen panel to judge the pros and cons of the different options. Citizens score options under the chosen criteria; the panel reviews the performance patterns, and decide what issues to take to a joint workshop.

In parallel, multicriteria mapping interviews are conducted with specialists. Specialists are guided by a researcher through their own individual analysis in separate 2-3 hour sessions. As part of this session the specialists are asked to consider whether they would like to add any criteria or options that might be relevant to them when appraising the project of the road through the wilderness area. During the session the researcher works interactively with a piece of computer software to explore the performance of options, against their criteria, under different assumptions. In addition to the quantitative and textual documentation recorded using the software package, the interviews are also audio-recorded for later transcription and analysis. In a joint workshop, specialists exchange views with citizens and respond to questions. In a further group meeting, panellists discuss workshop outcomes, re-score options, weigh criteria to reflect priorities, and discuss individual and full panel results. The specialists go through a second multicriteria mapping interview after the joint workshop to elicit any changes in their appraisals.

In a specialist workshop, specialists reflect on the various perspectives and emerging findings and evaluate the process. From the outset, the aim is not to achieve a consensus on how to proceed on the road proposal, but to expose the variety of views among and between
specialists and citizens, and to try to understand where the differences are most marked and why.

The assumptions of deliberative mapping are –

1. Systems are complex: The need for public participation arises from the nature of the problems at hand, from their urgency, their interdisciplinarity, their uncertainty and their irreversibility.

2. Trade-offs may or may not be possible: Ethical or other reasons may prevent the weighing up good consequences against bad.

3. Values arise from individual as well as group preferences: Value is determined by the interplay of individual and group preferences. To capture them, individual interviews need to be complemented by group processes.

4. Values are incommensurable: Irreducible value conflict is unavoidable but is compatible with rational choice through the employment of practical judgement.

Deliberative mapping uses a variety of strategies for the appraisal of options (Davies et al., 2003):

- open-ended interviews with specialists and stakeholders to explore the problem from their different perspectives;

- facilitated group discussions with small groups of up to 10 citizens, a one day joint workshop with all the citizens panels and the specialists, and a specialists’ meeting;

- structured quantitative appraisal using a scoring and weighting framework (based on multicriteria mapping), which are adapted for the different audiences;

- a joint workshop with an agenda set by the citizens’ panels; which includes the free exchange of views and information between citizens and specialists.

Figure 11 shows the overall structure of the deliberative mapping process. Citizens and specialists follow the same basic framework for appraisal. Each of a set of common options is assigned a score to reflect performance under freely defined criteria. These are then weighted to reflect their importance. This allows determination of a set of ranks, indicating the overall performance of each option. The result is a ‘map’ of the way performance varies under different perspectives. The patterns shown in the ‘maps’ are further interpreted through qualitative analysis of the interviews with individual specialists, and the discussions in the citizens’ panels.

Deliberative mapping aims to incorporate scientific expertise during the deliberation process itself. Firstly, by providing credible and authoritative information when demanded by panellists during their sessions and, secondly, by creating a space for deliberation between experts and lay persons in small workshops.
Nature of participation

Public and stakeholder engagement in deliberative mapping is comprehensive, multi-faceted and spans most stages of the appraisal process. While core options are defined by the research team, deliberations about criteria, scoring, self-organised engagement between specialists and citizens and the opportunity to review the various appraisal stages, makes for a comprehensive programme of participation. By integrating expert and citizen assessments and by emphasising diversity and social learning deliberative mapping aims for robust, democratic and accountable decision-making.

Decision context

Deliberative mapping focuses as much on the ‘opening up’ as on the ‘closing down’ of a decision or policy process. It uses the appraisal process as a way to gain a systematic picture of the precise way in which different perspectives vary on the issues and options in question. This generates a rich body of information concerning the reasons for differing views, as well as their practical implications for the overall performance of the selected options. Building on multicriteria mapping it spans the divide between narrow quantitative methods (which directly address decision priorities, but which may be insensitive to wider considerations) and broader qualitative approaches (which can accommodate more diverse perspectives, but can have difficulty focusing on the context of the decision). Additionally, perspectives from citizens and specialists are developed in parallel, but the two groups engage with each other before forming their final views.

Treatment of criteria and values

There are different legitimate values and points of view in society. This creates social pressure to take various perspectives of the issue at hand into account, e.g. economic, political, social, cultural, and religious. Citizens and specialists express these values and points of view in a combination of open-ended interviews, facilitated group discussions, structured quantitative appraisal using a scoring and weighting framework based on multicriteria mapping, a joint workshop between citizens and specialists.

Related to this is the issue of weighting. Similar to multicriteria mapping, deliberative mapping participants are asked to indicate the relative importance of each of their appraisal criteria by means of a simple numerical weighting. Taken together, these weightings reflect the relative importance of the criteria to the interviewee or group. In contrast to the relatively technical process of scoring, this weighting process reflects intrinsically subjective judgements over priorities and values.
Applicability to different types of evidence bases

Data sources can be scientific studies or expert judgements. The deliberative mapping framework allows the use of quantitative as well as qualitative data formats.

Citizen participants are encouraged to interact directly with specialists to appreciate any uncertainty surrounding the consequences, the ambiguity of the knowledge base, and any dissent among experts.

In the multicriteria mapping-type processes, specialist interviewees or citizen groups are asked to provide ‘best' and ‘worst' performance scores. This captures the degree of uncertainty and variability around the performance of particular options under a given criterion. Participants are asked to talk about the assumptions behind their different scores, and this qualitative data is recorded and transcribed. This captures any uncertainty about how well the option will actually work, the variability within the option, and any sensitivity to wider context conditions. This leads to rankings expressed as ranges of values instead of single numbers. These ranges express uncertainty in assigning scores, differences of opinion and variability in performance from context to context, for instance, the differences between good and bad implementation, or between appropriate and inappropriate applications.

Treatment of equity and distributional issues

The method treats equity and distributional issues in a similar way to Social Multicriteria Evaluation, Three-Stage Multicriteria Analysis and Multicriteria Mapping. However, it should be noted that the DM process was explicitly designed to facilitate and support participation by citizens from a range of socio-economic backgrounds.

Results

The results include a ranking of policy options as well as an analysis of the different perspectives about the options held among citizen participants and among experts. The interpretation of the rankings is helped by the textual analysis of statements recorded in the software during the interview and from the interview and group transcripts. The results should also include a sensitivity analysis and a clear view of the conflicting character of the criteria and the influence of a particular set of weights.

Human and financial resources

The costs of a deliberative mapping study are made up of the cost of deliberative processes with citizens and experts (organisation and support, payment and expenses of stakeholders, experts and citizen panel, hire of venues, costs of facilitation, recording and transcribing proceedings and costs for any related analysis) and the costs of the analytical tasks (multicriteria analysis and sensitivity analysis). The team conducting a deliberative mapping study needs expertise in deliberative methods, access to experts from relevant disciplines and expertise in multicriteria analysis. Deliberative mapping requires repeated engagement with a range of experts and citizens, making it relatively resource intense.

Key studies

Kidney transplants in the UK

http://www.deliberative-mapping.org

In the first deliberative mapping project, Davies et al. (2003) addressed the problem of the ‘kidney gap'; the fact that more patients need a kidney transplant than there are donor organs available, and demand for organ transplants will increase in the future. Options for reducing the kidney gap range from biotechnological innovations (such as stem cell based therapies and xenotransplantation), to changes in organ donation practices and improved health education. However, all the options may be disputed on scientific and technical grounds and may raise economic, cultural or ethical difficulties. Given this challenging context, the purpose of this study was to evaluate the strengths and weaknesses of deliberative mapping in allowing a fair, competent and transparent assessment of the options available to policy-makers.
Thirty-eight citizens from Camden, North London were recruited through a stratified social survey. Panellists, who agreed to attend six evening meetings and the one-day joint workshop, were chosen to reflect a wide range of occupations, ethnic backgrounds, ages and family circumstances. To cover views held by experts and to get access to information, 17 specialists were recruited to span a wide range of perspectives, disciplines and institutional affiliations.

Following the deliberative mapping process performances of options were analysed. Despite many differences in perspective, there was a remarkable degree of consistency between the appraisals made by the four citizens’ panels, and the 17 specialists. Of the six core options, two performed markedly better overall: improved transplant services, and encouraging healthier living. Two further options were generally ranked highly, but slightly lower or with more qualifications than the two ‘lead options’: presumed consent and altruistic living donation. There were correspondingly negative implications for certain technology-based options like xenotransplantation and embryonic stem cells. Qualitative analysis of the panel discussions showed that citizens used a variety of ‘public reasons’ rather than purely personal preferences to reach their judgements about the options. Figure 12 and Figure 13 show the resulting rankings from the citizens’ and experts’ perspectives.

**Figure 12 Citizens’ Panel Rankings**

![Figure 12 Citizens’ Panel Rankings](Source: Davies et al. (2003: 15))

**Figure 13 Specialist’s Rankings**

![Figure 13 Specialist’s Rankings](Source: Davies et al. (2003: 15))
Radioactive waste management in the UK

Burgess et al. (2004) conducted a trial deliberative mapping for the Committee for Radioactive Waste Management (CoRWM) to explore options for managing the UK’s radioactive waste. The trial brought 9 specialists and 16 citizens into a participatory option appraisal process.

In summary, this method is most suitable for the appraisal of policies, programmes or projects where views are controversial and where value judgements are particularly important. Deliberative mapping was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

3.6 Stakeholder decision / dialogue analysis

Definition

The combined use of group deliberation techniques and (a qualitative form of) multicriteria analysis to aid decision making about policy options while taking conflicting interests and multiple criteria into account. It highlights the framing of problems, scoping options, eliciting criteria and making judgements through facilitated deliberation (Burgess, 2000).

Framing

For example, consider again a proposal to build a new road through a wilderness area, which would destroy the habitat of a number of rare or threatened species. In a series of four workshops stakeholders are provided with relevant information about the proposed project and asked to come through a carefully designed combination of individual and collective processes to a ranking of options. The main steps (e.g. final rankings) must be agreed by all group members. The process allows for reconsideration and revision of results in a transparent manner.

The assumptions of stakeholder decision / dialogue analysis are –

1. Systems are complex: the need for stakeholder participation arises from the nature of the problems at hand, from their urgency, their interdisciplinarity, their uncertainty and their irreversibility.

2. Trade-offs may or may not be possible: Ethical or other reasons may prevent the weighing up good consequences against bad.

3. Values arise from individual as well as group preferences: Value is determined by the interplay of individual and group preferences. These can be captured by analysing group processes.

4. Values are incommensurable: Irreducible value conflict is unavoidable but compatible with rational choice employing practical judgement.

Stakeholder decision / dialogue analysis is a qualitative, multi-criteria process based on deliberative interactions in small groups over a period of several weeks. Stakeholder decision / dialogue analysis builds on the dialogic potential of small groups by inviting a range of stakeholders to work together in defining the problem requiring a decision; identifying options for resolving the problem; eliciting a set of criteria for appraisal of the options; undertaking the appraisal through a process of negotiating an agreed option score against each criteria; weighting criteria; and then reviewing outputs to agree final recommendations. In stakeholder decision / dialogue analysis, the specific multicriteria component acts as much as a heuristic to structure the discussions as a formal tool for analysis. In this sense, it is an adjunct to deliberation which helps with prioritising proposed actions or issues through consensus-building with stakeholder groups (Burgess, 2000; Clark et al., 1998). A similar approach, called ‘Stakeholder Preference Mapping’ was adopted by Butler et al. (2004).
Stakeholder decision / dialogue analysis may be used to seek expert opinion about priority issues, prioritise proposed options by the involvement of stakeholders, and / or build consensus on the priority of options.

Stakeholder decision / dialogue analysis involves facilitated discussions with a small group of people who meet for five sequential sessions (lasting 90 minutes each), to deliberate each stage of the appraisal process. The group members work interactively with one another and use low-tech pen and paper techniques to record their judgements about the performance of options against criteria.

The structure of stakeholder decision / dialogue analysis consists of 9 main steps –

1. Recruitment of a stakeholder group of key stakeholders with a professional interest or responsibility, or a local knowledge of, the issues. The stakeholder group is briefed in some detail about the background of the problem at hand and the purpose of the tasks to which they are to be applied. Each member of the stakeholder group is given a draft document in which the key issues identified or draft recommendations are set out.

2. Each stakeholder group member individually assesses gains, losses and the probability of an adverse impact occurring for each issue/recommended action or option.

3. Organisers convene a workshop at which all stakeholder group members discuss the draft document and discuss any omitted issues/options. The group then collates their individual lists of positive and negative impacts, and agrees a fully-inclusive list of benefits, costs and risks, as identified by the collective stakeholder group, for every issue/option presented.

4. At the end of the workshop group members are briefed about the next stage. Before the next workshop, they are asked to consider individually what sort of criteria might be helpful for prioritising the courses of action (options).

5. In a second workshop the stakeholder group produces a fully inclusive list of final criteria, based on the ones they thought of individually and brought to the workshop. The group is briefed for the next task.

6. Each group member must individually give each of the final criteria a ranking from 1-100 for its utility in assessing the priority issues/option. These rankings are sent to the organisers, who calculate the top 10 criteria from the group and calculate the score (weighting) for each.

7. In a third workshop group members work in small sub-groups and, using the top 10 criteria, assess each presented issue/option against each criterion. They agree assessments on a four-point scale: ‘high’ (3); ‘medium’ (2); ‘low’ (1); and, ‘not applicable’ (0). The scores are totalled for each issue and multiplied by the criterion weighting to produce a total score for each issue/option. The issues/options can then be ranked (prioritised) according to their scores. Options are placed into groups, a group being defined by a split between any two adjacently-ranked actions that have total scores differing by more than 10. The group is briefed for the final workshop.

8. Organisers send a final copy of the prioritised issues/options to each group member, for them to consider prior to the final workshop.

9. In a fourth workshop group members discuss the final prioritisation and work in small groups to agree the ranking. The exact rank order is not significant, but moving an issue/option between groups is. Members may decide to move an issue/option up or down to a new priority group, but if they do so, must move another one back in the opposite direction. All small-group members must agree to any changes. The full stakeholder group reassembles to discuss their decisions and the entire group must agree to any changes.

Clark et al. (1998) recommend using this method with no more than 24 issues/options and to involve the stakeholders in the process as early as possible. If possible, they should be involved in the generation of the list of issues/options to be prioritised.
Stakeholder decision / dialogue analysis uses a range of qualitative tools for facilitating and interpreting in-depth discussion groups, enabling consultation to transcend conventional emphasis on scientific knowledge and rational, utilitarian argument and to complement this with moral, aesthetic, emotional and local ways of knowing and valuing.

**Nature of participation**

Public and stakeholder engagement begins in stakeholder decision / dialogue analysis with an existing definition of policy options and includes deliberations about criteria, input in the impact matrix and exploration of results with different parameters. Hence, public and stakeholder engagement in stakeholder decision / dialogue analysis is comprehensive and spans the whole appraisal process. The process is also open, flexible and accessible.

**Decision context**

By engaging stakeholders at an early stage and by being transparent throughout the appraisal process, stakeholder decision / dialogue analysis supports the opening up of relevant issues. The need to agree on a common structure (criteria and options) and the ranking supports then the process of closing down.

**Treatment of criteria and values**

There are different legitimate values and points of view in society. This creates social pressure to take into account various perspectives of the issue at hand, e.g. economic, political, social, cultural, and religious. Stakeholders express these values and points of view in a combination of facilitated group discussions and structured qualitative appraisal using a scoring and weighting framework based on multicriteria analysis.

Related to this is the issue of weighting. Similar to multicriteria mapping, stakeholder decision / dialogue analysis participants are asked to indicate the relative importance of each of their appraisal criteria by means of a simple numerical weighting. Taken together, these weightings reflect the relative importance of the criteria to the interviewee or group. In contrast to the relatively technical process of scoring, this weighting process reflects intrinsically subjective judgements over priorities and values.

**Applicability to different types of evidence bases**

Data sources can be scientific studies, expert or stakeholder estimates, or traditional knowledge. The qualitative nature of the appraisal avoids over-precision in situations characterised by uncertainty in the data and ambiguities of assumptions.

**Treatment of equity and distributional issues**

The method treats equity and distributional issues in a similar way to Social Multicriteria Evaluation, Three-Stage Multicriteria Analysis, Multicriteria Mapping and Deliberative Mapping.

**Results**

The results include a ranking of (groups of) policy options as well as an analysis of stakeholders’ different perspectives about the options. The interpretation of the rankings is helped by the analysis of the qualitative data collected during the workshops. The results should also include a sensitivity analysis and a clear view of the conflicting character of the criteria and the influence of a particular set of weights.

**Human and financial resources**

The costs of a stakeholder decision / dialogue analysis study comprises the sum for the deliberative process (organisation and support, payment and expenses of panel and witnesses, hire of venue, costs of facilitation, recording and transcribing proceedings and costs for any related analysis). The team conducting a stakeholder decision / dialogue analysis study needs expertise in deliberative methods and multicriteria analysis. While stakeholder decision / dialogue analysis requires repeated interactions, its other data requirements are rather low. Hence, this is a relatively cost effective method.
Key studies

Local planning in England

Stakeholder decision / dialogue analysis has been applied in several local and regional environmental decision situations, including the New Forest Local Environment Agency Plan 1997-8 (Burgess, 2000; Clark et al., 1998). Local Environment Agency Plans (LEAPS) were developed by the UK Environment Agency to promote an integrated and sustainable approach to managing the natural environment. Stakeholder decision / dialogue analysis was applied to appraise the LEAP for the New Forest, England. Through a series of workshops selected interest groups produced a set of criteria for aiding discussion, evaluation and prioritisation of issues relating to the LEAP; this was combined with multicriteria analysis. The criteria employed covered environmental, social, economic and legal considerations. Scientific knowledge was accorded high value. Specific landscape and cultural values were recognised in a criterion that assessed local distinctiveness and quality of life issues. Economic concerns included criteria for maintaining the local economy in balance with social and environmental needs. The negotiated criteria changed the original set of priorities produced in the draft LEAP. The Agency agreed to abide by the decisions of the interest group, and duly issued the revised plan for wider public consultation. All LEAPs now incorporate a stakeholder decision / dialogue analysis in their planning phase.

In summary, this method is most suitable for the appraisal of policies, programmes or projects where it is important to work first on a common problem understanding and for which a rough impact assessment is sufficient as input in the decision process. Stakeholder decision analysis was developed to address complex issues and to deal with uncertainty in the context of sustainable development.
4. Institutions and policy context

4.1 Green Book

The Green Book is a best practice guide for all central departments and executive agencies, and covers projects of all types and size. It aims to make the appraisal process throughout government more consistent and transparent. Most departmental guidance documents (e.g. DfT’s WebTag) build explicitly on the Green Book. While it does not prescribe monetary valuation, its guidance for this type of valuation is much more detailed. The sections on alternative methods are extremely short and guidance on when to use which method is missing.

The Green Book provides guidance on how proposals should be appraised before significant funds are committed, and on how past activities should be evaluated. It is intended to ‘encourage a more thorough, long-term and analytically robust approach to appraisal and evaluation’. It aims to describe ‘how the economic, financial, social and environmental assessments of a policy, programme or project should be combined’ (p.1). The key sections on appraising options advise on the use of monetary valuation techniques. The presentation of alternative methods is extremely short. Regulatory Impact Assessment (discussed in the next section) only applies to proposed regulations, whereas Green Book guidance is more widely applicable.

The HM Treasury is responsible for the Green Book and although it provides guidance only, it has high status and there are expectations that it will be followed. The HM Treasury provides some direct guidance to departments and is available for consultation on matters relating to the Green Book.

The latest 2003 edition of the Green Book places ‘a stronger emphasis on the identification, management and realisation of benefits’, suggests a lower discount rate (3.5% instead of 6%), and advocates ‘a greater emphasis on assessing the differential impacts of proposals on various groups in society’ than preceding versions.

However, the 2003 edition focussed almost exclusively on monetary valuation techniques. While the Green Book does not demand monetary valuation, it does not offer substantive guidance on any alternative methods.

4.2 Regulatory Impact Assessment

Regulatory Impact Assessments (RIAs) were introduced in 1997, primarily as a tool for assessing economic impacts on business. Since then RIAs have developed to provide a framework for analysing likely impacts of a policy change and a range of options for implementing it. In the process the scope of RIAs has expanded, so that it now includes the appraisal of a full range of economic, environmental and social costs and benefits for policy proposals and includes a public or stakeholder consultation. RIAs are required for any form of regulation, such as formal legislation, codes of practice or information campaigns, which has an impact on business, the public sector or the voluntary sector. The Better Regulation Executive of the Cabinet Office is responsible for the RIA process across Whitehall; it provides guidance to departments, monitors departmental compliance with RIA requirements, is available for consultation on matters of RIA practice and gets involved in any RIAs judged to be significant in terms of established criteria. Within each department, a Better Regulation Unit coordinates RIAs, provides guidance and assistance to teams involved in regulation, and is responsible for liaison with the Better Regulation Executive (BRE). Departmental compliance with the requirement to produce an RIA for all significant legislative changes is very high.

In 2005 the House of Common’s Environmental Audit Committee expressed concern that the structure of RIAs and associated guidance was “ill-suited to the overriding need for policy

---

8 Environmental Impact Assessment (project appraisal) and Strategic Environmental Assessment (programme appraisal) are not covered by this review because they focus on environmental issues and therefore do not attempt to cover the spectrum of sustainability dimensions.
makers to be able to balance environmental impacts against social and economic impacts and to assess the extent of any trade-offs which need to be made”. The Committee recommended that the National Audit Office (NAO) carry out an analysis of recent RIAs to assess how they have responded to the change in procedures requiring explicit consideration of environmental and social impacts.

The NAO stated that in theory the RIA process provide an ideal way to ensure that sustainable development considerations were taken into account in policy making, because it requires: (a) the taking into account of all costs and benefits; (b) evidence-based policy making; and (c) stakeholder consultation. However, the NAO identified the following problems in the implementation of RIAs: (1) Some RIAs focussed on a technical detail rather than the substance of a new policy and therefore precluded or inhibited discussion of sustainable development concerns; (2) RIAs are too often used to justify rather than appraise and inform policy, because RIAs may be started late in the policy-making process and therefore have only limited relevance to the outcome of the policy; (3) Sustainable development was not well handled in most of the RIAs that the National Audit Office examined. Few RIAs identified all social or environmental impacts. Environmental and social impacts were often not analysed in sufficient depth. The presentation of most RIAs made it difficult to see if and how sustainable development had been considered, and some principles of sustainable development were less evident or under-treated in most RIAs, such as evidence-based policy making. As result, the National Audit Office recommended that guidance for officials should be improved, officials involved in the production of RIAs should receive more training on sustainable development and the BRE should improve its scrutiny of RIAs.

At the time of writing, the BRE is reviewing the future of RIAs and considering options for improving their usefulness in policy-making. In July 2006 they set out to consult on the following suggested revisions: (1) introducing a new template to promote greater accountability and transparency of decision-making; (2) introducing much shorter guidance; (3) removing the requirement for an impact assessment to show three policy options (requiring instead a scalable impact assessment for each policy option considered); (4) requiring publicly available impact assessments to be signed off by both the responsible Minister and the department’s Chief Economist; (5) removing mandatory requirements to include certain elements relating to business, social and environmental impacts except where they relate to the costs and benefits of the policy; and (6) changing the name from RIA to Impact Assessment.

Two main trajectories seem possible: (1) a limitation of RIAs in scope of the sustainability dimensions considered in order to make full monetary valuation possible (environmental and social dimensions are often measured easiest – and cheapest – in physical or qualitative terms and sometimes cannot be measured in monetary terms at all) or (2) suggesting a range of sustainability appraisal tools from which departments can choose depending on which is best suited for the specific context. The Better Regulation Executive plans to report their recommendations at the end of January 2007.

4.3 Magenta Book

The Magenta Book provides guidance on a range of designs and methods of policy evaluation for policy evaluators and analysts and for people who use and commission policy evaluation. When asked about the documents which influenced the appraisal process none of the interviewees referred to it actively, but most were familiar with it when asked. However, its impact on sustainability appraisal of policies, programmes and projects seems very much limited.

4.4 Integrated Policy Appraisal

A dedicated sustainability appraisal tool developed by Defra, which RIAs subsumed in April 2004.
4.5 Departmental Policy Appraisal Guidance

Many departments have developed own departmental policy appraisal guidance that is tailored to the types of issues relevant in the respective department. Most prominently these are the Department for Transport’s ‘Transport Analysis Guidance’, Defra’s ‘Project Appraisal Guidance’ (which also applies to the Environment Agency), and the Department for Culture, Media and Sport (DCMS)’s ‘White Book’.

The Environment Agency is currently exploring the possibility to use multicriteria appraisal for flood risk management more widely. A range of case studies have been commissioned and a decision on which appraisal methods will be used for which types of problems is expected in the year 2007.

In some departments Sustainable Development Action Plans contain a requirement for audit and appraisal in relation to the department’s performance in developing skills about sustainable development. For example, Defra uses for this purpose an own tool called ‘Stretching the Web’.
5. Key messages

1. The aim of sustainable development puts special demands on valuation and appraisal methods. The fact that sustainable development is (i) a multidimensional concept, (ii) crucial for human survival, and (iii) a long-term issue, makes it necessary to test currently used valuation and appraisal tools for their fitness for this context.

2. A range of new sustainability valuation and appraisal methods have been developed and road-tested over the last 10 years. They combine interpersonal deliberation with quantitative methods.

3. These methods build on the following principles: accounting for different types of knowledge (monetary and non-monetary; quantitative and qualitative data); taking inter-generational equity seriously; providing opportunities for learning during the appraisal process; ensuring transparency of each step of the appraisal process; and having a strong element of public and stakeholder engagement.

4. The key differences between the six reviewed methods are –

   Deliberative monetary valuation uses of formal deliberation concerning an environmental impact to express value in monetary terms for policy purposes. Like cost benefit analysis, deliberative monetary valuation is most suitable for the appraisal of projects whose impacts are rather well understood, where the impacts do not reach far into the future, and which do not affect complex ecosystem services such as biodiversity.

   Social multicriteria evaluation combines participatory techniques and multicriteria analysis to aid decision-making about a number of policy, programme or project options while taking conflicting interests and multiple criteria into account. This method is most suitable for the appraisal of policies, programmes or projects whose impacts are not well understood yet and therefore benefit from a multidisciplinary modelling of impacts. Social multicriteria evaluation was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

   Three-stage multicriteria analysis uses participatory techniques and multicriteria analysis to aid decision making about policy options. This method is most suitable for the appraisal of policies, programmes or projects whose impacts are reasonably well understood by experts and which contain a significant technical element. Three-stage multicriteria analysis was developed for the social appraisal of technologies with particular emphasis on the role of risk and uncertainty.

   Multicriteria mapping is an interview-based multicriteria analysis that focuses on eliciting and documenting detailed technical and evaluative judgements concerning the performance of alternative options. This method is most suitable for the appraisal of policies, programmes or projects to which stakeholders had some exposure and where views not only about data, but also about options and criteria, are controversial. Multicriteria mapping was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

   Deliberative mapping combines participatory techniques and multicriteria analysis to aid decision making about policy options. This method is most suitable for the appraisal of policies, programmes or projects where views are controversial and where value judgements are particularly important. Deliberative mapping was originally developed to address complex issues and to deal with uncertainty and contested values in the appraisal of controversial technologies.

   Stakeholder decision / dialogue analysis combines the use of group deliberation techniques and (a qualitative form of) multicriteria analysis to aid decision making
about policy options. This method is most suitable for the appraisal of policies, programmes or projects where it is important to work first on a common problem understanding and for which a rough impact assessment is sufficient as input in the decision process. Stakeholder decision analysis was developed to address complex issues and to deal with uncertainty in the context of sustainable development.

5. While the new sustainability valuation and appraisal methods show some differences between them and their application is suitable for slightly different problems, the main difference lies between this group of methods and monetary environmental valuation on the one hand and deliberative methods on the other hand.

6. While these new methods will benefit from further testing and development, they have been shown to be effective tools of sustainability appraisal and valuation under a wide range circumstances. Very few studies so far have compared different appraisal methods systematically for specific areas of application.

7. The type of evaluation and the institutional structures in which the appraisal is embedded influences the outcome. Therefore the choice of valuation and appraisal methods is not wholly a technical question, but one of ‘institutionalising social choice’.

8. The Green Book offers detailed guidance for monetary valuation. The sections on alternative methods are extremely short and guidance on when to use which method is missing. Most departmental guidance documents build explicitly on the Green Book.

9. The way how monetary, other quantitative and qualitative impacts are currently summarised in RIAs, resembles an impact matrix (options are appraised against a range of criteria). In a multicriteria appraisal the next steps would be to weigh the criteria and to apply an algorithm that calculates a ranking of options. Under the current RIAs, decision-makers are required to do these final two steps by themselves without technical support, which the decision making literature would consider an excessive demand on any human judgement. From this practice, the danger arises that not all criteria are fully taken into account.

10. While the major guidance documents highlight the benefits of monetary valuation (ease of comparison, expression of opportunity costs etc.), they acknowledge that monetary valuation may sometimes be difficult and leave it up to the appraiser to judge how benefits and costs can be represented best. While there are few formal limitations to applying the new appraisal methods on a policy, programme or project level, in several departments there is the belief that “monetary valuation is best under all circumstances”. At the same time many difficulties for applying monetary valuation across all policy areas are known and expressed in these departments. The guidance from the Green Book could be improved by presenting a range of appraisal methods, comparing and contrasting them and giving advice on which methods are most suitable under which circumstances. The guidance from the RIA could be improved by keeping the openness to different types of impacts and the ways how they are measured, but developing a framework which gives more support to decision-makers.

11. Since there is no one method suited to appraising all types of policies, programmes and projects, a more differentiated approach should be favoured. In order to develop cultures that support the search for the most suitable methods, departments need to be familiar with the different methods at hand and provide guidance on them so appraisers can choose the most appropriate method.

12. Together with other reports that were mentioned in the introduction, this report aims to contribute to a map of different valuation and appraisal methods. The methods highlighted in this report are conceptually and practically particularly suited for appraising policies, programmes and projects for which sustainability is an important component.
# Appendix 1 - Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithms / Multi-criteria algorithms</td>
<td>Process, or set of rules, usually one expressed in algebraic notation, now used especially in computing, machine translation and linguistics.</td>
</tr>
<tr>
<td>Choice modelling</td>
<td>A ‘stated preference’ technique that is used to estimate non-market environmental costs and benefits by use of a survey instrument. Each question, called a ‘choice set’, presents to respondents the outcome of usually three or four alternative strategies. The alternatives are described in terms of a common set of attributes. The alternatives are differentiated one from the other by the attributes taking on different levels.</td>
</tr>
<tr>
<td>Citizens’ jury / Consensus conference</td>
<td>Tool for community opinion while addressing controversial issues. They are based on the rationale that given adequate information and opportunity, all people, even if not professionally trained, can be trusted to take decisions regarded as legitimate and fair on behalf of the community.</td>
</tr>
<tr>
<td>Complex system</td>
<td>System of many parts, which are coupled in a nonlinear fashion. The behaviour of a complex system cannot be inferred from the behaviour of its components considered separately.</td>
</tr>
<tr>
<td>Cost-benefit analysis</td>
<td>Appraisal method which seeks to value the expected impacts of an option in monetary terms. These valuations are based on utility theory and uses willingness-to-pay or willingness-to-accept. This theory can act as a guide to how valuation should be achieved, and as a referee in disputes about valuation.</td>
</tr>
<tr>
<td>Deliberation</td>
<td>A careful consideration or discussion of reasons for and against (OED). Or: reasoned debate between individual actors whereby understanding is advanced and mutual agreement is reached (or not) via the quality and persuasiveness of argument rather than by coercion, manipulation or deception.</td>
</tr>
<tr>
<td>Discounting</td>
<td>The process by which a sum due at a future date is given a present value.</td>
</tr>
<tr>
<td>Ecological economics</td>
<td>The study of the human economy as part of nature's economy.</td>
</tr>
<tr>
<td>Environmental economics</td>
<td>The specialisation within neoclassical economics that is concerned with the economy's insertions into the natural environment.</td>
</tr>
<tr>
<td>Extended peer communities</td>
<td>Participants in the quality assurance processes of knowledge production and assessment in Post-Normal Science, including all stakeholders engaged in the management of the problem at hand.</td>
</tr>
<tr>
<td>Fuzzy number</td>
<td>A number whose value is imprecise, rather than exact as is the case with ‘ordinary’ (single-valued) numbers. Hence, they can capture situations where a value is only known ‘roughly’, ‘nearly’, ‘about’ or ‘crudely’. A fuzzy number can be thought of as a function whose domain is a specified set (usually a set of real numbers).</td>
</tr>
<tr>
<td>Governance</td>
<td>Governance encompasses collective decisions made in the public sector, the private sector, and civil society.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Ignorance</td>
<td>State of knowledge where the definition of a complete set of possible outcomes is problematic and probabilities cannot be assigned.</td>
</tr>
<tr>
<td>Incommensurability of values</td>
<td>The absence of a common unit of measurement across plural values; it entails the rejection of monetary reductionism but also any physical reductionism (e.g. eco-energetic valuation).</td>
</tr>
<tr>
<td>Intergenerational equity</td>
<td>The question of fairness between human generations.</td>
</tr>
<tr>
<td>Monetary commensurability</td>
<td>Goods exhibit monetary commensurability, if they can be precisely measured along the cardinal scale of money.</td>
</tr>
<tr>
<td>Monetary valuation</td>
<td>Estimation in monetary terms of the value of environmental services and goods.</td>
</tr>
<tr>
<td>Multiattribute Utility Theory</td>
<td>A multicriteria approach that is based on von Neumann-Morgenstern axioms of preference and thus upon a utility function, which allows the comparison of risky outcomes through the computation of expected utility. It involves direct questioning of decision-makers about trade-offs and making choices on the basis of an aggregate utility measure for each alternative. Software packages: PREFCALC and 'Decide Right'.</td>
</tr>
<tr>
<td>Multi-criteria mapping</td>
<td>A multi-criteria appraisal method developed by A Stirling, which emphasises that there is a wide range of perspectives and expertise, and produces an overview that 'maps' the debate. It combines the transparency of numerical approaches with the unconstrained framing of discursive deliberations.</td>
</tr>
<tr>
<td>Multicriteria appraisal (also: multicriteria analysis or multicriteria decision analysis)</td>
<td>A group of appraisal techniques which make the options and their contribution to the different criteria explicit, and all require the exercise of judgement. Formal MCA techniques usually provide an explicit relative weighting system for the different criteria and lead to a ranking of options.</td>
</tr>
<tr>
<td>Neoclassical economics</td>
<td>The currently dominant school of economics.</td>
</tr>
<tr>
<td>Outranking methods</td>
<td>A group of multicriteria methods (ELECTRE I, II, III, IV, IS, TRI; PROMETHEE I, II; REGIME; NAIADE) which builds on less stringent assumptions about the existence and characteristics of utility functions. An option outranks another if it outperforms the other on enough criteria of sufficient importance (as reflected by the sum of the criteria weights) and is not outperformed by the other option in the sense of recording a significantly inferior performance on any one criterion. Outranking methods report incomparability of options, if their performance patterns have fundamentally different structures. Aim to capture some of the political realities of decision-making by downgrading options that perform badly on any one criterion.</td>
</tr>
<tr>
<td>Post-Normal Science</td>
<td>Methodology that is appropriate when &quot;facts are uncertain, values in dispute, stakes high and decisions urgent&quot;.</td>
</tr>
<tr>
<td>Preference function</td>
<td>In multi-criteria analysis a preference function indicates the intensity of one alternative over another for each criterion.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Public participation</td>
<td>In broad terms refers to any of the several mechanisms (processes and tools) intentionally designed for involving lay public and / or stakeholders (the two groups together are often referred to as interested parties) in decision making. Includes deliberation and inclusion (Bloomfield et al., 2001).</td>
</tr>
<tr>
<td>Risk</td>
<td>State of knowledge where the different possible outcomes are known exactly and a probability can be assigned to each possibility.</td>
</tr>
<tr>
<td>Scenario</td>
<td>An internally consistent story about one way in which the future could unfold.</td>
</tr>
<tr>
<td>Social incommensurability</td>
<td>In policy contexts social actors refer to a set of contrasting and legitimate values, perceptions and interests. This implies that any decision is always associated with winners and losers.</td>
</tr>
<tr>
<td>Stakeholders</td>
<td>Stakeholders are those actors who are directly or indirectly affected by an issue and who could affect the outcome of a decision making process regarding that issue or are affected by it.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>Maintaining the capacity of the joint economy-environment system to continue to satisfy the needs and desires of humans for a long time into the future.</td>
</tr>
<tr>
<td>Sustainability appraisal</td>
<td>A variety of methods of inquiry and argument to produce policy-relevant information that is then utilised to evaluate the consequences of human actions against the normative goal of sustainable development.</td>
</tr>
<tr>
<td>Sustainability valuation</td>
<td>Methods and procedures for valuing changes in environmental, social and economic subsystems in an integrated way. Sustainability valuation methods are part of the sustainability appraisal process.</td>
</tr>
<tr>
<td>Sustainable development</td>
<td>A widely-used definition of sustainable development is: “development which meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED 1987). Alternatively, development that conserves the natural capital, limits population and total resource demand in scale, maintains the integrity of ecosystems and diversity of species, remedies social inequities and environmental damage, while maintaining a sound economic base, fulfils basic health and educational needs, and is based on participatory democracy (Harris 2003).</td>
</tr>
<tr>
<td>System</td>
<td>A set of interacting components.</td>
</tr>
<tr>
<td>Technical incommensurability</td>
<td>It is impossible to reduce to a single model and data set heterogeneous information referring to representations belonging to non-equivalent descriptive domains.</td>
</tr>
<tr>
<td>Transition to sustainability</td>
<td>Movement across a threshold into another phase (sustainable development) as the result of the interaction of relevant factors during the previous phase.</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>State of knowledge where there is confidence in the completeness of the defined set of outcomes, but it is acknowledged that no valid theoretical or empirical basis exists to assign probabilities to these outcomes.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Utilitarianism</td>
<td>The school of ethics according to which the moral correctness of an action depends on the balance of pleasure and pain that it produces. In utility theory value is assigned based on the utility gained by individuals from a project, policy or option.</td>
</tr>
<tr>
<td>Valuation</td>
<td>The process of assessing the contribution of a particular object or action to meeting a particular goal, whether or not that contribution is fully perceived by the individual.</td>
</tr>
<tr>
<td>Weak comparability</td>
<td>Comparable without recourse to a single type of value; irreducible value conflict is unavoidable but compatible with rational choice employing practical judgement (O’Neill 1993); it can be operationalised by use of multicriteria analysis (Martínez-Alíer et al., 1998).</td>
</tr>
</tbody>
</table>
Appendix 2 - Theories and concepts

Utility theory

The underlying idea of utilitarianism is the principle of the greatest happiness for the greatest number (Bentham 1780). People pursue things which provide pleasure and avoid things that produce pain; all individuals seek to maximise their total pleasure. According to utilitarianism, the moral correctness of an action depends on the balance of pleasure and pain that it produces. Actions that increase the totality of pleasure or reduce the totality of pain are morally correct, actions that reduce the totality of pleasure or increase the totality of pain are morally incorrect. The term 'utility' refers to the situation of an individual in regard to the balance of pleasure and pain - pleasure is that which increases an individual's utility, pain is that which reduces an individual's utility. While pleasure and pain considered as subjective, for making policy recommendations interpersonal utility comparisons have to be made. Such comparisons necessitate the precise measurement of utility in ordinal and ultimately in cardinal terms. Bentham chose money as his units of measurement. Since market prices tend to underestimate the pleasure for some consumers, the 'consumer surplus' (utility in excess of price) is used to measure the pleasure. If market prices do not exist, surrogate markets are designed and consumers asked about their values. The term 'welfare' is used for the totality of utility across individuals, and according to utilitarianism morally correct actions are those that increase welfare. Utilitarianism is the ethical basis for neoclassical economics.

Since financial resources are involved, decisions about public policy are considered as an economic process and participants are addressed in their role as consumers.

Ecological concepts of values

Ecological economists (Martinez-Alier et al., 1998) have argued that when dealing with complex systems there is no uniquely rational way to aggregate different dimensions of value along a single metric. The absence of a common unit of measurement across plural values is called incommensurability. It entails the rejection of monetary reductionism but also any physical reductionism (e.g. eco-energetic valuation). Recently Aldred (2006) explored in depth the consequences of incommensurability for monetary valuation.

Decisions about public policy are considered as political acts and participants are addressed in their role as citizens.

Deliberative democracy

The assumptions of representative democracy have been progressively undermined by the scale and complexity of contemporary societies and their rates of change. Elected representatives can rarely capture the diverse values and social and economic interests of their constituents, while the uncertainties generated by novel threats argue for the inclusion of a wider range of knowledges in decisionmaking (Dryzek 1990). Deliberation is defined as 'careful consideration' or 'the discussion of reasons for and against'. It follows that deliberation is not necessarily inclusive. Deliberation is less likely to be effective in large groups because individuals would have less time to express themselves and learn from others. Hence, there is a tension between deliberation and inclusion, which needs to be carefully considered when designing appraisal processes.

Claims for a more deliberative style of decision-making have been strongly influenced by the writings of Habermas. He emphasizes that deliberation in the public sphere has benefits far beyond decision-making (Habermas 1984). By increasing our understanding of the interests of others, it allows us to appreciate our own interests better.

Decisions about public policy are considered as political acts and participants are addressed in their role as citizens.
## Appendix 3 – Interviews

18 people were interviewed from a range of governmental departments and organisations, together with one consultant.

<table>
<thead>
<tr>
<th>Organisation</th>
<th>Individual</th>
<th>Date of Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defra, ACS</td>
<td>Shaun Mowat</td>
<td>26 October 2006</td>
</tr>
<tr>
<td>Defra, EBC</td>
<td>Ruth Coward</td>
<td>30 October 2006</td>
</tr>
<tr>
<td>Defra, EP</td>
<td>Bob Davies</td>
<td>7 November 2006</td>
</tr>
<tr>
<td>Defra, LLSD</td>
<td>Peter Costigan</td>
<td>9 November 2006</td>
</tr>
<tr>
<td>Defra, SDU</td>
<td>Virginia Hall</td>
<td>25 September 2006</td>
</tr>
<tr>
<td>Defra, SDU</td>
<td>Arik Dondi</td>
<td>25 September 2006</td>
</tr>
<tr>
<td>Defra, SDU</td>
<td>Matt Jackson</td>
<td>25 September 2006</td>
</tr>
<tr>
<td>Defra, SQASD</td>
<td>Prashant Vaze</td>
<td>27 October 2006</td>
</tr>
<tr>
<td>DTI, ESIU</td>
<td>Dr Preetum Domah</td>
<td>27 October 2006</td>
</tr>
<tr>
<td>English Heritage</td>
<td>Geoff Dawe</td>
<td>30 October 2006</td>
</tr>
<tr>
<td>Environment Agency</td>
<td>Rob Curry</td>
<td>30 October 2006</td>
</tr>
<tr>
<td>Environment Agency</td>
<td>Colin Foan</td>
<td>16 November 2006</td>
</tr>
<tr>
<td>HM Treasury, Green Book</td>
<td>Joseph Lowe</td>
<td>7 November 2006</td>
</tr>
<tr>
<td>Joint Nature Conservation Committee (JNCC)</td>
<td>Emily McKenzie</td>
<td>31 October 2006</td>
</tr>
<tr>
<td>Natural England</td>
<td>Steve Preston</td>
<td>25 October 2006</td>
</tr>
<tr>
<td>NERA Economic Consulting (formerly HM Treasury)</td>
<td>Michael Spackman</td>
<td>8 November 2006</td>
</tr>
<tr>
<td>SEERAD</td>
<td>Robert Henderson</td>
<td>17 October 2006</td>
</tr>
<tr>
<td>Sustainable Development Commission</td>
<td>Victor Anderson</td>
<td>27 October 2006</td>
</tr>
</tbody>
</table>
References


Gibbons, M. 1999 'Science's new social contract with society', Nature 402(C81): 11-17.


Harris, J. M. 2003 'Sustainability and Sustainable Development', in International Society for Ecological Economics (ed) Internet Encyclopaedia of Ecological Economics.

Hinloopen, E. 1985 'De Regime Methode' Interfaculty Actuariat and Econometrics, Amsterdam: Free University Amsterdam.

Hinloopen, E. and Nijkamp, P. 1990 'Qualitative multiple criteria choice analysis', Quality and Quantity 24: 37-56.


Martí, N., Vidal, V., Mánuel, D. and Munda, G. 2000 'DIAFANIS. Evaluación Ambiental Integrada de propuestas de actuaciones de desarrollo socioeconómico y ambiental en el entorno del Parque Nacional de Aigüestortes y Estany de Sant Maurici (España).'


Millstone, E., Lobstein, T., Stirling, A., Mohebati, L. and the PorGrow National Teams 2006 'Policy options for responding to obesity: cross-national report of the PorGrow project', Brighton: SPRU - University of Sussex.


O'Brien, E. A. 2003 'Human values and their importance to the development of forestry policy in Britain: a literature review', *Forestry* 76(1).


Renn, O. 2006 'Participatory processes for designing environmental policies', *Land Use Policy* 23: 34-43.


— 2001 'Deliberative Monetary Valuation' 5th Nordic Environmental Research Conference, Aarhus, Denmark.


