

**Multi-Attribute Value Measurement
and
Economic
Paradigms in Environmental
Decision Making.**

Alison Ruth Joubert

**Thesis presented for the degree of
Doctor of Philosophy**

**in the Department of Statistical Sciences
University of Cape Town**

December 2002

**Supervisor: Professor Theodor Stewart, Department of Statistical Sciences,
University of Cape Town**

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

**For Louise, André, Liz, and Dave, and especially for my
parents who gave so much.**

Multi-Attribute Value Measurement and Economic Paradigms in Environmental Decision Making.

Alison Ruth Joubert, April 2002

ABSTRACT

The two environmental decision-making approaches of environmental economics (EE) valuation and multi-criteria decision analysis (MCDA) differ fundamentally in their underlying philosophies and approach; hence they are characterised as paradigms. The EE paradigm includes the idea that, if appropriate prices can be found and implemented for goods not normally traded on the market, then the market mechanism will efficiently distribute resources and decisions are therefore based on the concepts of individual willingness to pay and consumer sovereignty. That an efficient market is not necessarily equitable or sustainable has long been acknowledged, but EE adjustments are subject to theoretical and methodological problems. The MCDA paradigm is based on the idea that values and preferences should be examined and constructed through interaction between workshop participants and the analyst, given basic measurement theory axioms.

Various EE and MCDA methods have been devised for measuring value in different contexts, some of which were applied, in the context of environmental (particularly water resources) management, in six action research case studies. The EE methods were contingent behaviour valuation, the contingent valuation method, conjoint analysis and the travel cost method. The MCDA method was a version of the simple multi-attribute rating technique (called SMARTx). In the SMARTx cases, applying a group-value sharing model during a series of workshops, stakeholders rated the effect of alternatives on a number of environmental, social and economic attributes directly or using value functions and gave weights to criteria. Indirect compensatory values of one criterion in terms of another were determined. In the EE cases, survey respondents were asked their travel costs, preference for multi-attribute profiles and willingness to pay for alternatives. Total and average willingness to pay for an amenity, its attributes or changes in environmental quality were determined. The practical and theoretical implications of applying the different methods were examined and compared in terms of four metacriteria: resonance with and validity within the prevailing political and decision-context, general validity and reliability, ability to include equity and sustainability criteria and practicality.

Conclusions from the case studies were that:

- (a) the contingent valuation method may produce citizen rather than consumer values, is expensive to apply and produces unreliable results,
- (b) SMARTx produces group-based rather than individual values,
- (c) SMARTx has a representativeness problem due to the small numbers who can be directly involved in workshops,
- (d) the EE approach has an inherent pro-rich bias,
- (e) the EE approach is unable to adequately include equity and sustainability criteria,
- (f) the travel cost method is unreliable where source populations are dispersed and their tastes and incomes variable and skewed,
- (g) respondents in EE and SMARTx applications are able to give strength of preference information, and
- (h) SMARTx provides a valuable problem structuring framework and is generally regarded as simple and transparent.

There was considerable support in the literature for conclusion (a), and some support for (d), (e), (f) and (h). The strengths of SMARTx included its development of values through debate, and its ability to include a wide range of types of value not limited by translation to monetary scales. The strengths of EE were its comprehensive typology of values (which are nevertheless all anthropocentric) and its broad inclusivity through surveys.

Based on this work, a pragmatic SMARTx-based approach was proposed that was informed by the EE concepts of use and non-use value and additionally augmented by a survey to find individual priorities and issues of concern rather than willingness to pay. This combined approach should be representative and more able to deal with equity and sustainability and with the trade-offs between these and efficiency than applying either method on its own or other available hybrid approaches.

Table of Contents

ABSTRACT.....	i
Table of Figures.....	v
Tables.....	vi
Abbreviations and Acronyms.....	vii
Symbols.....	vii
Acknowledgements.....	viii
1. INTRODUCTION.....	1
1.1 Overall context, aim and limitations.....	4
1.2 Why 'paradigms'?.....	5
1.3 Layout of thesis.....	7
1.4 Outline and context of case studies in Part II.....	8
1.4.1 Sand River catchment planning - Chapter 6.....	8
1.4.2 Tourism value of the Crocodile river flowing through the Kruger National Park - Chapter 7.....	9
1.4.3 Other case studies - Chapter 8.....	9
1.5 Conclusion.....	11
1.6 Editorial notes.....	12
PART I - THEORY AND REVIEW.....	13
2. VALUE MEASUREMENT, VALUE AND UTILITY THEORY.....	15
2.1 Scales of measurement and binary relationships.....	15
2.1.1 Scales.....	15
2.1.2 Value and utility functions.....	16
2.1.3 Binary preference relationships.....	17
2.2 Value functions.....	17
2.2.1 Single attribute ordinal value functions.....	17
2.2.2 Single attribute cardinal value functions and value difference functions.....	18
2.2.3 Single attribute utility functions.....	18
2.2.4 Multi-attribute value functions.....	19
2.2.5 Independence.....	20
2.2.6 Value and weight elicitation.....	23
2.3 Summary and conclusion.....	24
3. A BRIEF HISTORY OF ECONOMICS AND THE DEVELOPMENT OF ENVIRONMENTAL ECONOMICS.....	25
3.1 The classical and neo-classical economists.....	25
3.2 Consumer choice, social choice, welfare economics and cardinal utility.....	29
3.3 Utility under uncertainty: 'measurable' or not?.....	32
3.4 Origins of environmental and resource economics.....	34
3.4.1 CBA and valuation tools used in environmental economics.....	36
3.5 Summary and conclusion.....	38
4. TYPES AND APPLICATIONS OF MCDA.....	41
4.1 General background and schools of MCDA.....	41
4.1.1 Multi-attribute decision-making.....	43
4.1.2 Multiple objective decision-making.....	45
4.1.3 Extensions to group decision-making.....	46
4.1.4 Extensions to uncertainty and imprecise judgements.....	54
4.2 Applications in environmental management.....	58
4.3 Summary.....	62
5. CONTINGENT VALUATION, CONJOINT ANALYSIS, TRAVEL COST AND SMART.....	63
5.1 Measuring welfare and multi-attribute value.....	64
5.2 Environmental economics valuation methods.....	66
5.2.1 Stated preference: Contingent Valuation (CVM).....	66
5.2.2 Stated preference: Contingent behaviour valuation (CBV).....	69
5.2.3 Stated preference: Conjoint analysis (CA).....	69
5.2.4 Stated preference approach and survey biases.....	71
5.2.5 Revealed preference methods – the travel cost method (TCM).....	72
5.3 The simple multi-attribute rating technique (SMART).....	73
5.3.1 Problem structuring.....	74
5.3.2 Evaluation and analysis.....	76

5.3.3	Group decision-making model for SMARTx	79
5.3.4	Comparison of MCDA methods - the use of SMARTx in environmental decision-making	80
5.4	<i>Summary and conclusions</i>	84
PART II - CASE STUDIES		87
6.	DEVELOPMENT AND ANALYSIS OF CATCHMENT LAND- AND WATER-USE ALTERNATIVES USING SMARTx	89
6.1	<i>Methods</i>	90
6.1.1	Group membership and group process	91
6.1.2	Problem structuring: Development of catchment land-use alternatives	91
6.1.3	Problem structuring: Criteria and value tree formation	94
6.1.4	Consequences and evaluation of scenarios	96
6.1.5	Evaluation: Aggregation of scores for each zone	98
6.1.6	Evaluation: Criteria weights	98
6.2	<i>Results</i>	100
6.2.1	Preferred scenarios for Zone A	100
6.2.2	Preferred scenario for Zone B	101
6.2.3	Preferred scenario for Zone C	101
6.2.4	Sensitivity to weight changes	102
6.3	<i>Discussion</i>	102
6.3.1	Costs and benefits of preferred scenarios	102
6.3.2	Implied trade-offs and indirect compensatory values	103
6.3.3	Sources of value	107
6.3.4	Practical issues	109
6.4	<i>Summary and conclusions</i>	110
Appendix 6A. Consequences of scenarios, scores and weights.		111
7.	ESTIMATING CHANGES IN THE TOURISM VALUE OF A RIVER WITH EE	115
7.1	<i>Methods</i>	115
7.1.1	The travel cost method (TCM)	117
7.1.2	Contingent behaviour valuation (CBV)	118
7.1.3	Conjoint analysis (CA)	118
7.2	<i>Results</i>	122
7.2.1	Travel cost, consumer surplus and contingent behaviour	122
7.2.2	Conjoint analysis	123
7.2.3	Relating the CA scores and attributes to travel cost	125
7.3	<i>Discussion</i>	126
7.4	<i>Summary and conclusions</i>	129
Appendix 7A. Comparison of conjoint models: value functions		131
Appendix 7B. Prediction of scenario scores with one survey version omitted from the regression		132
8.	OTHER SMARTx AND EE CASE STUDIES	135
8.1	<i>Maclear district forestry and land-use decisions</i>	135
8.1.1	Background	136
8.1.2	Methods	137
8.1.3	Results	141
8.1.4	Comments	144
8.2	<i>Evaluation of supply augmentation and demand management options for City of Cape Town</i>	144
8.2.1	Results	146
8.2.2	Comments	146
8.3	<i>Evaluation of contribution of the Big Five to tourism value of Hluhluwe National Park</i>	147
8.3.1	Results and comment	149
8.4	<i>Evaluation of different types of urban open space using CA, CVM and ranking and rating</i>	152
8.4.1	Results and comments	153
8.5	<i>Summary and Conclusions</i>	155
Appendix 8A. Scenario descriptions, consequences and scores for Maclear forestry case study		156
PART III - DISCUSSION AND SYNTHESIS		158
9.	COMPARISON OF THE METHODS APPLIED IN THE CASE STUDIES	159
9.1	<i>Social choice and environmental decision-making</i>	159
9.2	<i>Framework for and summary of methods</i>	162
9.2.1	Evaluation of alternatives with AD-SMARTx and SMARTx	163
9.2.2	Valuation of alternatives or attributes using EE	164
9.3	<i>Metacriteria for comparison</i>	164

9.3.1	Political and decision-making context.....	165
9.4	<i>Resonance with the prevailing political milieu and policy context.....</i>	<i>166</i>
9.4.1	Theoretical validity of EE.....	167
9.4.2	SMARTx: Are group-values also citizen values?.....	173
9.4.3	Summary discussion and proposal for Section 9.4.....	175
9.5	<i>General validity and reliability.....</i>	<i>176</i>
9.5.1	Consumer surplus and the travel cost method.....	176
9.5.2	Measuring utility - or the ordinal/cardinal debate lives on.....	177
9.5.3	Mutual preferential independence and difference independence.....	180
9.5.4	Survey, questioning and meeting biases.....	180
9.5.5	Weights and explicit trade-offs.....	181
9.5.6	Uncertainty.....	181
9.5.7	External validity and reliability.....	182
9.5.8	Summary discussion and proposals for Section 9.5.....	183
9.6	<i>Efficiency, equity and sustainability criteria.....</i>	<i>184</i>
9.6.1	Equity.....	185
9.6.2	Sustainability.....	189
9.6.3	Summary discussion and proposal for Section 9.6.....	193
9.7	<i>General practicality and provision of aid.....</i>	<i>193</i>
9.7.1	Costs in time and money.....	194
9.7.2	Simplicity, flexibility, ease of use, transparency and participation.....	194
9.7.3	Facilitation.....	197
9.7.4	Perceptions.....	197
9.7.5	SMARTx questionnaire.....	198
9.7.6	Summary discussion and proposals for Section 9.7.....	199
9.8	<i>Summary of performance and proposals.....</i>	<i>199</i>
10	SYNTHESIS AND RECOMMENDATIONS.....	203
10.1	<i>Decision and evaluation contexts.....</i>	<i>204</i>
10.2	<i>Summary of strengths, shortcomings and proposals.....</i>	<i>206</i>
10.3	<i>Social choice theory and environmental choice.....</i>	<i>207</i>
10.4	<i>Recommendations for practical aid for environmental decision-making.....</i>	<i>208</i>
10.4.1	Adjustments to SMARTx and the group-value sharing model.....	209
10.4.2	Problem structuring.....	211
10.4.3	Evaluation of consequences.....	212
10.4.4	Analysis and feedback.....	213
10.5	<i>Comparison with other combined approaches.....</i>	<i>213</i>
10.6	<i>Research issues.....</i>	<i>215</i>
10.7	<i>Conclusion.....</i>	<i>216</i>
	References.....	219

Table of Figures* captions are abbreviated

Figure 1.1. Layout of thesis.....	8
Figure 1.2. Location of case studies.	10
Figure 2.1. An illustration of difference independence modified from von Winterfeldt and Edwards (1986).....	22
Figure 2.2. An illustration of multilinear independence.	22
Figure 3.1. Main contributors and developments in economics and EE.	27
Figure 3.2. Indifference curves for two goods x_1 and x_2 , $V(c) >_1 V(b) >_2 V(c)$ but the size of $>_1$ is unknown.	28
Figure 3.3. Individual consumer surplus (left) and aggregate consumer surplus (right).	37
Figure 4.1. Development of MCDA. K&R=Keeney and Raiffa, VW&E=Von Winterfeldt and Edwards.....	42
Figure 5.1. The cumulative density function for rejecting the offer of q^1 at a cost of T.....	68
Figure 5.2. Determination of consumer surplus from travel cost.....	73
Figure 5.3. The SMARTx framework.	75
Figure 5.4. A value tree.	75
Figure 5.5. (a) A thermometer scale, (b) a non-linear value function.....	77
Figure 5.6. Trade-off between aggregate benefits and costs of different dam sites.....	78
Figure 5.7. Marginal rate of substitution (MRS) of an environmental criterion and a cost attribute.	79
Figure 5.8. Summary of group-value sharing model applied in SMARTx (see also 4.1.3 for the original version). 79	
Figure 6.1. Zones used for the development of alternative alternatives within the Sand River catchment.....	90
Figure 6.2. Value tree structure, criteria used and their associated scales.	96
Figure 6.3 Example of (a) 'thermometer scale' direct-scoring, (b) linear and non-linear value functions.....	97
Figure 6.4. Preferred scenarios from different points of view for each zone.	100
Figure 6.5. Relative contributions of criteria to overall scores of scenarios for Zone A.	101
Figure 6.6. Relative contributions of criteria to overall scores of scenarios for Zone B.....	101
Figure 6.7. Relative contributions of criteria to overall scores of scenarios for Zone C.	102
Figure 6.8. Relative contribution of criteria weight to the three criteria group weights in Zone A.....	103
Figure 6.9. The trade-off between SEr and OM moving from A to A' and from B to B'..	107
Figure 6.10. Sources of value for the catchment land-use scenarios (spp =species).....	108
Figure 7.1. Map of the study area.....	116
Figure 7.2. Links between the different sections of the questionnaire and the analyses.	117
Figure 7.3. Left: visits and travel cost from different origin zones. Right: derived demand curve for visits to KNP.123	
Figure 7.4. Frequency of unadjusted scores in five different scenario versions.....	123
Figure 7.5. Relationship between Z-score and Travel Cost for Model 1, 2 and 3.	126
Figure 8.1. Map of the Maclear area of the northern Eastern Cape.....	137
Figure 8.2. Value tree constructed from the criteria and categories defined during the workshops.....	140
Figure 8.3 (a) The linear value function for employment, and (b) non-linear value function for land-types preserved.	140
Figure 8.4. Weights of three criterion groups, with contributions by criteria.	141
Figure 8.5 Aggregate scores, showing contributions from lower level criteria.	142
Figure 8.6. Overall scores (solid line) and ranks (dotted line).	147
Figure 8.7. Sensitivity of supply option scores to weight on the criteria group 'political and consumer acceptance'147	
Figure 8.8. Visits and travel cost from different origin zones (used to develop the demand curve for visits to HUP).149	
Figure 8.9. Scores for different levels of the four attributes in the CA question using Model2..	151
Figure 8.10. Relationship between scores and average WTP and the scores.....	152
Figure 8.11. Average WTP for different amounts of open space.....	154
Figure 9.1. Inkomati / Crocodile catchment showing the study areas for the Sand and KNP studies.....	160
Figure 9.2. Social choice through the EE valuation and MCDA paradigms.....	161
Figure 9.3. Mean, median and standard deviation for changes in length of stay and WTP.....	171
Figure 9.4. A SMARTx value tree including criteria that together might indicate tourism WTP.	184
Figure 9.5 Frequency of responses to the questionnaire.....	199
Figure 10.1. Intellectual and methodological links between EE and MCDA.....	204
Figure 10.2. Flow chart for choosing an appropriate framework and methodologies.....	205
Figure 10.3. Approaches to social choice, showing critical weaknesses.	207
Figure 10.4. Proposed decision-making framework. EE contributions are shaded.	209
Figure 10.5. Inputs from SMARTx and a survey combined to inform social welfare decisions.	211

Tables* captions are abbreviated

Table 3.1. Values and techniques used in environmental economics or within CBAs.	38
Table 4.1. Simple classification of MCDA methods.....	43
Table 6.1. Participants in the SMARTx workshops.....	91
Table 6.2. Land-use scenarios for the Sand River catchment.....	94
Table 6.3. Weights for each sub-criterion and global weights.....	99
Table 6.4. Sets of weights suggested for the main criteria groups for the different zones..	100
Table 6.5. Benefits and costs of choosing Scenario 7.....	103
Table 6.6. Value of 1 point change in OM and inferred SEr value	105
Table 6.7. Trade-offs between all criteria and OM and the benefit of choosing Scenario 7 over Scenario 5.	105
Table 6.8. Compensating OM value for gains or losses in each criterion.....	106
Table 6.9. Value of changes from Scenario X to Scenario 7, with a <u>linear value function for OM</u> ..	107
Table 7.1. The five survey versions for the conjoint analysis showing attributes and attribute levels.....	119
Table 7.2. Survey versions with corresponding average scores, standard deviations and adjusted.	124
Table 7.3. Summary of the four conjoint analysis models' regression coefficients and regression statistics.....	125
Table 7.4. Predicted Z-scores for all scenarios from all models, compared to original (adjusted) scores.....	125
Table 7.5. Summary of economic values of goods and services supplied by the Crocodile River.	126
Table 8.1. Points of view, stakeholders and their representatives at the workshops.....	138
Table 8.2. Hectares of forestry owned, afforested, primary processing and constraints for the Maclear scenarios.	139
Table 8.3. Monetary value of 1 value point changes in each criterion (indirect compensatory value).	143
Table 8.4. Water supply and demand management options.....	145
Table 8.5. Criteria, criteria levels and associated scores.....	145
Table 8.6. The different conjoint analysis scenarios in the three survey versions.	149
Table 8.7. Average and adjusted average from the CA question.	150
Table 8.8. Regression results for conjoint analysis using equivalent of Model1 and Model2 of KNP study.	150
Table 8.9. Percentage importance of different attributes from direct questioning and from the CA question.	151
Table 9.1. Comments arising from the questionnaire. R = respondent code.....	198
Table 9.2. A 'score-card' for the SMARTx and EE derived from the discussion in this chapter.	200

Abbreviations and Acronyms

CA	-	Conjoint Analysis
CBA	-	Cost-Benefit Analysis
CBV	-	Contingent Behaviour Valuation
CVM	-	Contingent Valuation Method
DM	-	Decision-Maker
DWAF	-	Department of Water Affairs and Forestry
EE	-	Environmental Economics
EIA	-	Environmental Impact Assessment
IEM	-	Integrated Environmental Management
MAUT	-	Multi-Attribute Utility Theory
MAVT	-	Multi-Attribute Value Theory
MCDA	-	Multi-Criteria Decision Aid/Analysis
MCDM	-	Multi-Criteria Decision Making/Methods (MCDA and MCDM are the same)
NPV	-	Net Present Value
NWA	-	National Water Act
SMART	-	Simple Multi-attribute Rating Technique
SMARTx	-	The particular form of SMART applied in the case studies
TCM	-	Travel Cost Method
USA	-	United States of America
WTA	-	Willingness To Accept (also Willing To Accept)
WTP	-	Willingness To Pay (also Willing To Pay)

Symbols

$\{a,b\}$	-	refers to the strength of preference relationship for alternative a over b .
$\succsim$	-	Weakly preferred, preferred
$\succsim^*$	-	‘at least as strong as’, as in ‘the strength of preference of a over b is at least as strong as the strength of preference of c over d : $\{a,b\} \succsim^* \{c,d\}$
$\succ$	-	Strictly preferred, preferred
$\sim$	-	Indifferent
$\forall$	-	Across all, for all
$\in$	-	Element of
$\subset$	-	Subset of
$\not$	-	not (e.g. $\not R = \text{not } R$)
$\Leftrightarrow$	-	if and only if
$\Rightarrow$	-	implies that
$\exists$	-	there exists
f	-	function
Σ	-	Sum of (e.g. $\Sigma_{i=1}^n x_i = \text{Sum of all } x_i \text{ from } i=1 \text{ to } n$)
Π	-	Product of

Acknowledgements

I would like to thank my supervisor Professor Theo Stewart for having quiet faith in my ability to get on with it, and for believing that I actually did have the material for a PhD.

I would like to thank my colleagues in the Dept of Statistical Sciences for their support. I would particularly like to thank Phil Whittington for his ready ear, tea, and Penguins and for putting up with my illegal smoking for so long, and Leanne Scott for understanding so well where I was coming from, and for also always providing an understanding ear. Thanks to David Glassom for reading revisions of the last two chapters.

I would like to thank my colleague and friend Dr Jane Turpie, Dept of Zoology, University of Cape Town, for inspiration, support, and the opportunity to be in wonderful places while collaborating on various projects with her, thereby contributing much of the material used in this thesis. She also made insightful and valuable comments on the content and structure of my thesis.

Thanks to another colleague and friend Dr Sharon Pollard of AWARD, who wouldn't allow me to be vague about what MCDA was all about. She also unwittingly provided me with thesis material through asking me to be involved in the Save the Sand project.

I owe a huge debt also to Tony Leiman of the Dept of Economics, University of Cape for initiating me in the arcane art of economics and for always finding the time to answer more questions.

Thanks to those who have danced with me through these last few years, especially Nicola Visser and Malcolm Black and the others who showed so poignantly that there are no handicaps, and for reminding me about serenity, and to the dancers and teachers at Jazzart for maintaining some balance in my very unbalanced life.

I have in my four siblings, Louise, André, Liz, and Dave, a wonderful support team, and they have all in their own ways persuaded me that I can, will and should achieve anything I want to, and kept me buoyed with their own unique senses of humour and insights.

My biggest thanks to my parents, who should have been here to see me finally hand this in, and to allow me the chance to spoil them for a change.

Funding Acknowledgements

The Foundation for Research Development provided a bursary for my initial studies, and subsequent funding came through my employment as a researcher for several Water Research Commission funded projects. I would like to thank these organisations for their financial support.

1. Introduction

"Happily there is nothing in the laws of value which remains for the present or any future writer to clear up; the theory of the subject is complete." J.S. Mill 1848, in Black 1970 p. 11

This thesis is concerned with ways of making decisions about projects or policies which impact on the environment, and therefore, indirectly, with the allocation and management of natural resources. With the evolution of human society from hunter-gatherer to urban pleasure seeker, and with the concomitant increase in population, environmental impacts have accumulated, grown in spatial and temporal scales and intensity, and therefore become more widely felt, as healthy environments have become scarcer. The quality or 'health' of the environment has a direct bearing on human existence for a number of reasons. Firstly, we are all directly dependent on certain ecological functions for food, water and air, even though, in urban settings this dependence has become less obvious. With climate change, both oxygen and food may become scarcer, and with increasing pollution loads potable water may be at a premium. Secondly, many people derive pleasure from experiencing natural environments and from knowing that they exist. Thirdly, we directly utilise many species of plants and animals, which, through mixing with 'wild' stocks, may become more hardy. Fourthly, some species, yet to be discovered, may be useful to future generations, for example in disease prevention or cure. Many of these impacts are felt differently by rich and poor, and by urban and rural populations. For example, environmental degradation may directly affect subsistence rural populations' livelihoods, while impacting urban dwellers either through scarcity or price increases or both. The methods used in environmental decision-making therefore have a non-trivial role to play in our futures.

The business of making decisions about the environment is one which therefore now affords a high priority globally, as indicated by world bodies either dedicated to this issue (e.g. the International Union for Conservation of Nature and Natural Resources (IUCN) and the World Wide Fund for Nature (WWF)), or having large sections dedicated to improving their decision-making in this regard (e.g. World Bank) and by global treaties (e.g. the Kyoto Protocol, Agenda 21, Convention on International Trade in Endangered Species of wild flora and fauna (CITES)). Attendees at the World Summits on Sustainable Development (1992, 2002) have found themselves defining and redefining the concepts of 'sustainability' and 'sustainable development'. They remain ill-defined concepts that mean different things to different people, but increasingly they are associated with social issues such as poverty as much as with the protection of the natural environment. The Brundtland Report (1987 p. 43) defined sustainable development as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" and this remains the starting point of most debates on the issue. The translation of this or any other definition into policies and policies into specific actions necessarily will involve different understanding of and interpretation of what should be the dominant policy 'narrative': efficiency, equity or ethics including 'green' ethics (Harrison 2000). In the interim, decisions need to be made which at least take cognisance of the issues as understood by the relevant government or government departments. United Nations Secretary General Kofi Annan said of the 2002 World Summit on Sustainable Development: "This Summit will put us on a path that reduces poverty while

protecting the environment... We have to go out and take action” (UN website) and this includes decisions about which actions to take.

Analytical approaches which may be used to assist in the analysis of projects and policies with multiple and environmental impacts include psychometrics, economics and cost-benefit analysis, mathematical optimisation, risk assessment, decision analysis and environmental impact assessments amongst others (Parsons 1995, Bell *et al.* 1977). Two disciplines or paradigms that have evolved over the years, particularly the last four decades, are explored in this thesis. These are the disciplines of multi-criteria decision analysis (MCDA) and environmental economics (EE), a sub-discipline of economics. These differ in the degree to which they are intended as prescriptive or descriptive analyses, and the degree to which they are bound to any specific normative, moral or philosophical theory, or claim to be positive (i.e. believing that analyses can make use of objective facts separately from existing values, perceptions and theories (Parsons 1995)). The different approaches rely differently on preferences being revealed (e.g. in the market) or stated (e.g. in a survey), deliberated or constructed (e.g. in a meeting). They therefore also differ in the manner and degree to which they are participatory and the way in which they use individual preference or expert input. This difference is also linked to the underlying moral and associated theories. For example, the concept of economic value as applied in environmental decisions is associated with the theory of welfare economics (Kontoleon *et al.* 2001). The similarities and differences between the paradigms will be examined from theoretical and practical perspectives and will be illustrated with practical applications and reference to the large body of literature available.

I have referred to these as paradigms of environmental decision-making, and try to illustrate why in the chapters that follow. However, it is clear that they are two approaches which are linked through their use of similar basic concepts from utility and measurement theory and may be applied in the same decision contexts. A brief statement of measurement theory is therefore given in Chapter 2. The literature treats the links between EE and MCDA in a variety of ways. MCDA and EE are felt by some to have the same origins but to have subsequently followed different development paths so that “Though the differences [between CBA and MAUT¹] at the level of mathematics and conceptualisation are relatively minor, the differences in both intellectual flavour and actual procedures are vast.” (von Winterfeldt and Edwards 1986 p. 567). Others consider that they have quite separate origins but also “..differ fundamentally in goal and detail” (Watson 1981 p. 242). I agree with the former view, i.e. that their origins are essentially the same, and therefore explore the history of the concept of utility in economics (Chapter 3). As an initial step in examining the paradigmatic differences we may consider the stated aims of EE and MCDA.

Economics may be defined as the study of the production, consumption and distribution of scarce resources (Samuelson and Nordhaus 1985 p. 4), given the notion of each economic agent as a rational utility maximiser. ‘The market’ is seen as the ‘clearing house’ where quantities and prices are settled efficiently. In evaluating projects or policies, the aims of EE valuation are to find consumers’ total willingness to pay for a gain, or to avoid loss, in environmental quality (or willingness to accept compensation for a loss or to forego a benefit), where the gain or loss impacts their welfare or utility. EE methods were therefore developed to derive “..credible estimates of people’s

¹ CBA=Cost-benefit analysis, MAUT=Multi-attribute utility theory.

values in contexts where there are either no apparent markets or very imperfect markets (so called ‘missing’ or ‘incomplete markets’)” (Pearce and Seccombe-Hett 2000 p. 1420). The point of deriving the estimates is either to include them in a cost-benefit analysis (CBA) to evaluate the potential benefits and costs of proposed projects or policies, or in evaluating environmental damages so as to determine compensation to be paid.

Decision analysis synthesises ideas from economics, statistics, psychology, behavioural science and operations research in order to, behaviourally or axiomatically, describe, prescribe and construct decision-making processes. Specifically, the aims of MCDA are to decompose a complex decision problem into a set of smaller problems which are easier to handle (Goodwin and Wright 1998), to influence and improve decision-making (French 1988) and to assist in understanding what alternatives are best and why (Keeney 1992). In doing this, the agenda is explicitly prescriptive; to help people to do what economics assumes they do (von Winterfeldt and Edwards 1986).

An example of EE valuation used in CBA-type project evaluation was a 1995 household survey which provided the basis for reducing Los Angeles’ water rights by nearly a half in order to increase environmental flows to Mono Lake (Loomis 2000). A famous example of the use of EE to estimate compensation was the 1989 Exxon Valdez oil spill; subsequent environmental damage assessments using EE techniques estimated damage of between \$3 and \$15 billion due to various groups (Braden 2000), but an out-of-court settlement was reached for \$1 billion (Kontoleon *et al.* 2001). An well-known example of a MCDA application was an evaluation of potential sites for nuclear power stations in the United States of America (USA), where human health, safety, environmental, socio-economic and financial impacts were considered (Keeney and Nair 1976). A later analysis of five potential sites (Merkhofer and Keeney 1987), prompted the Department of Energy to further investigate three sites which had ranked first, third and fifth in the original study, although the fifth was clearly dominated (verified in a study of portfolios of options by Keeney 1987). As a result, the Department of Energy had various lawsuits filed against them, and the MCDA analysis was supported in the subsequent investigations (Dodgson *et al.* 2001).

Environmental impact assessment, cost benefit analysis, and multi-criteria decision analysis each include a variety of different techniques, and differ in various ways, although they may also be seen as a continuum. Both MCDA and CBA are sometimes viewed as particular methods of carrying out an environmental impact assessment, while in some countries the three are perceived as quite distinct. MCDA is sometimes viewed as a particular way of doing a CBA, or CBA is considered as a generic term covering MCDA, environmental impact assessment and risk assessment (e.g. Fischhoff 1977), while some MCDA practitioners view CBA as a special case of MCDA (e.g. Bell *et al.* 1977). In practice, particular methods are often applied with little thought to the associated philosophical, theoretical, technical or practical issues, but rather because academics are interested in examining new approaches, or local expertise in a particular approach exists, or a particular method has become entrenched either through legislation or habit. Clarification is therefore needed as to what the associated ‘moral theories’, normative assumptions and philosophies are, so as to be able to apply methods in appropriate contexts.

Certain EE valuation techniques and the philosophical basis of the CBA framework have been the subject of heated debate over the last couple of decades, within both the environmental and economics literature, and it is particularly the controversial nature of cases like the Exxon Valdez which have brought these methods into question. While MCDA is apparently less controversial, this might be because it has not been applied in such famous or notorious situations. However, in some quarters there is resistance to the acceptance and use of MCDA as a viable alternative or complement to EE.

1.1 Overall context, aim and limitations

The importance and potential impact of environmental decision-making at local, national and global scales means that it is important to carefully consider which methods are appropriate for different decision-making contexts. Despite their quite different sounding aims, as mentioned above, there are many situations where either of the two approaches could be applied: how do we decide on the appropriate approach and does it matter which is applied? The criticisms of both approaches and of the particular techniques associated with them, relate to a wide range of issues from the fundamental philosophies to the details of application. These criticisms need to be coherently examined to clarify which can be overcome or not. Of further interest is whether the two approaches can be combined in a way that improves on their separate use.

There are a number of issues which combine to make the differences between the two paradigms “vast” (*sensu* von Winterfeldt and Edwards 1986), some of which have been mentioned above. Although there is a substantial literature on many of these aspects, as well as some comparisons of MCDA and EE, these have usually only considered some of the points mentioned. I therefore hope to consolidate much of this debate, providing a holistic overview, as well as an examination of some details.

The aims of this thesis are therefore threefold. The first aim is to describe the two paradigms of MCDA and EE in terms of:

1. their origins and underlying philosophies, theories and assumptions, including the types of ‘rationality’ emphasised,
2. their use or adaptation of the concept of value and utility, specifically related to the values associated with environmental impacts and consequent ways of measuring value (is value revealed, stated, constructed), and techniques and procedures followed,
3. the use of individual’s, consumer’s or citizen’s preferences,
4. their prescriptive or descriptive natures, and
5. their practical approach to the process of decision-making, and usefulness in terms of the provision of decision-aid.

This description begins in Part I, which introduces the history, background and theory behind the methods, and continues in Part II in which the case-studies are described.

The second aim, is to compare the methods, based on the background in Part I and the applications in Part II, in terms of the following metacriteria:

- Resonance with and validity within the prevailing political milieu and decision context,
- General validity and reliability,
- Ability to include equity and sustainability criteria, and
- Practicalities.

The third aim is to propose a ‘working model’ of decision-aid for environmental decision-making based on the above comparison.

The case studies make use of some MCDA and EE methods in an action research setting. This meant that only a few approaches could be applied. The applications are limited to ‘subjective value approaches’ of EE (Section 3.4.1) and ‘value measurement approaches’ of MCDA (Section 4.1.1). These are similar in their reference to utility, and in the contexts in which they can be applied. Other EE (Chapter 3) and MCDA (Chapter 4) methods may well be appropriate in similar contexts or as complementary methods. The action research setting means that the primary purpose of the case studies was not to address particular research questions but to apply methods in real settings, with very real time and budget constraints, and a requirement for applicable results. Therefore there was not really room for experimentation with the particular methodology chosen, and the applications were not ‘purist’. There was also no possibility to use a number of methods for direct comparison. Therefore, for the most part, direct numerical comparisons are not possible, and the comparisons relate to the types of results produced, their accuracy and robustness, the procedures followed, qualitative feedback from those involved, and to the implications of the underlying theories on the application and for interpretation of results. Many practical lessons were learned relating to the process of decision-making. In fact, it is often in the process and the practicalities that the two paradigms are most obviously different. I have therefore tried to capture these practical aspects in the discussions.

1.2 Why ‘paradigms’?

I have applied the term ‘paradigm’, in the sense of Kuhn (Kuhn 1962), to the two approaches to environmental decision-making. He felt that the scientific community tended to work within a general theory (virtually an ideology), seeking to flesh out and further explain the theory with additional work. In so doing, anomalies may appear but are treated as anomalies rather than as problems which fundamentally call into question the current body of theory or paradigm. But in time, sufficient anomalies might accumulate that a ‘paradigm shift’ occurs, akin to a revolution, where a new paradigm becomes current. Additionally, Kuhn together with Popper, viewed science as being dominated by the paradigmatic viewpoint, which controls how observations and research are conducted. Kuhn suggested that for robustness, problems should be framed and analysed in competing paradigms (Parsons 1995).

My view is that, despite the common origins and germ theory of MCDA and EE, and the many common elements between them, the two are different paradigms, or rather, that people working in these two fields work within fundamentally different paradigmatic or ideological views.

For example, environmental economists work within a particular paradigm: that decisions should be based on consumer sovereignty and willingness to pay, with adjustments for market imperfections, and to deal with distributive problems. Despite anomalies that have been apparent, and strong criticisms that have been made from within the economics fraternity, they continue to work within this paradigm, in part, they would say, for lack of a satisfactory and practical alternative. It would appear that a ‘paradigm shift’ is underway in the EE field, in part reflected in the burgeoning field of ecological economics (see for example the *Ecological Economics* journal, and other books and papers by its editors and regular contributors). Another indication of this paradigm shift is that despite the decades of ‘separate development’ there has recently been acknowledgement in some EE circles that MCDA may offer a viable alternative to traditional EE in certain contexts (e.g. Munasinghe 1993, Pearce and Seccombe-Hett 2000).

On the other hand, MCDA does not appear to be in the same paradigmatic flux. This may be for a number of reasons. Firstly, MCDA is the younger of the two paradigms and struggles still to properly define itself as a field. It may not be considered to be a cohesive paradigm at all (T Stewart pers. comm.) and has been characterised as “pre-paradigmatic” (Lootsma 1996 p. 37). The term MCDA implies different things to different people even within the discipline, but for the purposes of this thesis includes the full range of methods mentioned in Chapter 4 as well as the ‘soft’ problem structuring techniques (usually called ‘soft operations research’ e.g. Rosenhead 1989). This inclusivity is probably a positive aspect of MCDA, but does mean that the field is ultimately less definable as a paradigm than EE. Secondly, MCDA may be taught in a variety of departments and faculties (economics, management science, conservation biology, statistics etc.) and therefore has a less obvious paradigmatic home than EE. However, those working within the field of applied MCDA, do operate from a paradigmatic viewpoint: that problems need to be structured and preferences constructed on the basis of (perhaps approximations to) various rationality and measurement theory axioms, through interaction between the decision-maker and the analyst.

As mentioned above, some within the EE paradigm have suggested that the MCDA framework offers a viable alternative to EE, while perhaps most are either unaware of the existence of MCDA or regard it as inferior because it is ‘subjective’. Many decision-makers share the latter view and feel that EE offers a more objective and rational analysis. Some within the MCDA paradigm have considered the use of MCDA as an alternative to CBA or EE valuation (e.g. Watson 1981, Gregory *et al.* 1992), usually with the conclusion that it is superior in the context of projects with intangible or qualitative environmental and social impacts. Sometimes the economic approach is seen as part of the general MCDA paradigm (e.g. Keeney and Raiffa 1976). Thus the views emanating from within each paradigm are largely consistent, with a few dissenting voices. On the practical side, some of the methods more conventionally housed within the MCDA paradigm, have been adopted and developed in EE applications, and have become rather econometrically sophisticated. An example is conjoint analysis which has been adapted to find consumers’ willingness to pay for alternatives and marginal rates of substitution for attributes of alternatives (i.e. it

has been adapted to the EE paradigm). For decades, economists and EE practitioners working within the EE paradigm, have been grappling with philosophical questions such as the concept of economic value as relevant to environmental impacts, and the implications of particular methods in terms of distribution, equity, and ethics. MCDA theoreticians and practitioners have only more recently begun to consider the equity and ethical consequences of the use of a particular methodology (e.g. Rauschmayer 2001, Wenstøp and Seip 2001), although from the first MCDA conference the question of the inclusion of morality within MCDA was raised (Churchman 1972 in Wenstøp undated). Thus on both the practical and philosophical levels, there is both parallel development of the paradigms and convergence.

1.3 Layout of thesis

The thesis is divided into three parts, the first (**Part I**: Chapters 2 to 5) includes an outline and review of existing theory (Figure 1.1). Chapter 2 briefly outlines utility and measurement theory as applicable to this thesis, followed by a discussion of the origins and development of this theory in economics, and the later development of EE in Chapter 3. Chapter 4 introduces the different fields encompassed in the term MCDA and discusses some current theory and applications. Chapter 5 explains in more detail the evaluation methods that were used in the case studies. These were contingent behaviour valuation, the contingent valuation method, the travel cost method (three EE methods), conjoint analysis (used within both EE and MCDA), and the simple multi-attribute rating technique (an MCDA method). Extensions of economic and MCDA theory to collective or group choice are given within Chapters 3, 4 and 5. Part I, therefore, forms the basis of the chapters that follow by tracing the common origins of MCDA and economic theory in measurement and utility theory, and by describing the different evaluation or valuation methods used within these two approaches. It also provides the basis for examining and interpreting the results of the case studies and the implications of the paradigms in Part III.

The case studies are introduced in Part 1.4 below and are discussed in detail in Chapters 6 to 8, which forms **Part II**. Two case studies are reported in more detail (Chapters 6 and 7) while several other case studies, which used the same methods or variations, are reported more briefly in Chapter 8.

The final part, **Part III**, is a discussion and synthesis of the preceding parts. The relevant features of the methods, in terms of philosophy, theory and practicalities are discussed in Chapter 9 in the light of the lessons learnt in the applications and the earlier background, and are evaluated in terms of the four metacriteria mentioned in Section 1.1. Finally, Chapter 10 summarises the key points and proposes a decision-making framework for environmental decision-making (Figure 1.1). Initial ideas for the last two chapters were presented at the 28th Societas Internationalis Limnologiae (SIL) congress in February 2001 in Melbourne, Australia and at the 4th International Conference on Operations Research in Development (ICORD) in May 2001 Kruger National Park, South Africa. Some sections originated in a paper co-authored by A Leiman of the School of Economics, University of Cape Town for presentation at a conference 'Ecology, Society, Economy', May, 1996, St Quentin-en-Yveline, France and an earlier version was published in *Ecological Economics* (Joubert *et al.* 1997).

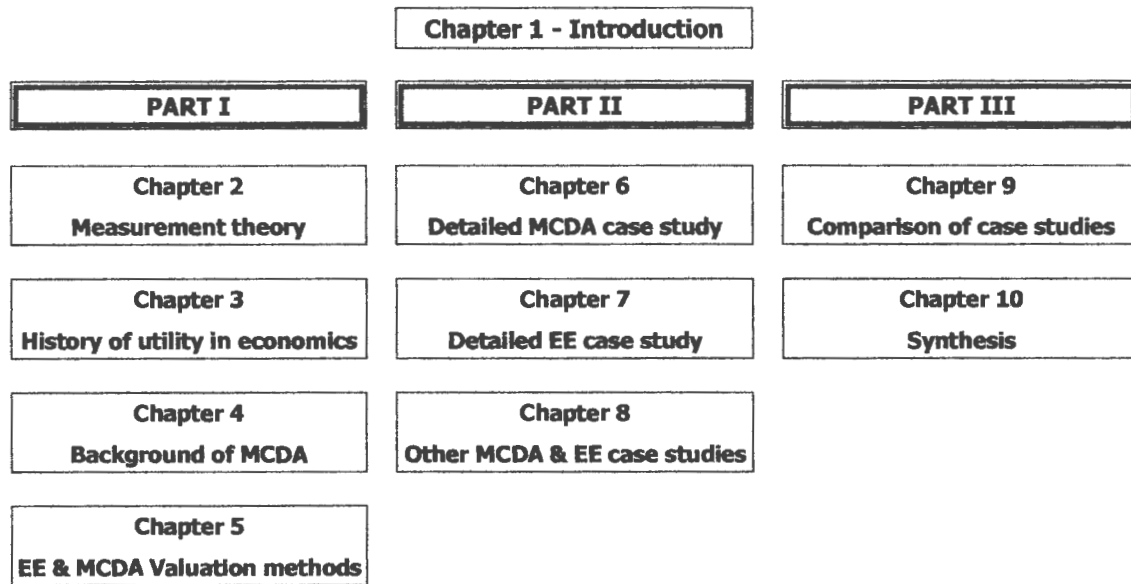


Figure 1.1. Layout of thesis.

1.4 Outline and context of case studies in Part II

An MCDA application is described in Chapter 6 and an EE application in Chapter 7. These two case studies were both concerned with land- and water-use or catchment planning, and arose in the context of the promulgation of a new National Water Act (NWA) (Act No. 36, RSA 1998), which fundamentally changed the approach to water resource management in South Africa.

Additional case studies are briefly outlined in Chapter 8. In these studies the same methods were applied as in the previous chapters as well as variations (e.g. open-ended contingent valuation instead of contingent behaviour valuation). Experiences gained in the latter case studies have informed and influenced the discussions in Chapter 9 and Chapter 10, but because the essential features are covered in the main case studies, they did not warrant detailed repetition here and only the main points are captured in Chapter 8.

1.4.1 Sand River catchment planning - Chapter 6

The Sand River catchment (Figure 1.2) project was run as a pilot project to investigate approaches to catchment planning within an integrated catchment management framework, and in anticipation of the new NWA. The overall project was commissioned by the Dept. of Water Affairs and Forestry and the Dept. of Agriculture, undertaken by the Association for Water and Rural Development (AWARD) and funded primarily by the Sabie-Sand Game Reserve. AWARD invited me to run the decision-aid part of the project, which I did as part of a Water Research Commission funded project (Stewart *et al.* 2001). This decision-aid concentrated on land-use and associated water-use implications of land-use alternatives, while the broader project also considered bulk-supply issues, water conservation strategies, catchment management agency structuring, education and training, amongst others. The decision-aid consisted of:

- providing an overall MCDA framework,

- facilitating four MCDA workshops with the project team of specialist who broadly represented ecological, social and economic issues in the catchment,
- assistance with the creation of land- and water-use alternatives,
- using MCDA tools for alternative evaluation (i.e. using so-called thermometer scales and swing weights),
- the development of a database for the analyses, and
- analysis of trade-offs and sensitivity.

The output consisted of a report which formed a chapter of the overall Sand River project write up (Pollard *et al.* 1998) and of Stewart *et al.* (2001). This output was used to make overall recommendations regarding land-use in the catchment. These recommendations are being carried through into Phase II of the Sand River project, while the preferred alternatives are currently being translated into 'visions' and management classes for the river as required by the NWA. The MCDA work was also published in abridged form in Joubert and Pollard (2000), reported at the Southern African Society of Aquatic Sciences conference (SASAQS June 1999), and at the Integrated Management of River Ecosystems conference (August 1999).

1.4.2 Tourism value of the Crocodile river flowing through the Kruger National Park - Chapter 7

The context for this case, was the requirement in the NWA that a management class (broadly indicating levels of protection or development) should be chosen for sections of each river. A Water Research Commission project (Turpie *et al.* 2000, Mander *et al.* 2001) recommended the inclusion of economics in this process, and I participated as part of a Water Research Commission project (Stewart *et al.* 2001). One of the economic values identified was tourism value. For the purposes of exploration of methods, therefore the work described here was limited to assessing the tourism value of the main river and its tributaries flowing through a protected area (the Kruger National Park - Figure 1.2). The output was a report on tourism value (also published in Turpie and Joubert 2001), which is to be integrated with the other values being estimated. An honours student in the Dept. of Statistical Sciences, University of Cape Town examined different models for analysis of the conjoint data (Cohen 2001) and I have extended this examination in this chapter. I would like to acknowledge the efforts of the University of Cape Town conservation biology masters students who participated in this project, namely: Sharon Bosma, Daniel Chongo, John Foord, Sarah Frazee, Mathew Hemming, Alina Lengyel, Nonfo Mosesane, Jean Mwicigi, Paul Ndag'ang'a, Genevieve Pence, Bianca Preusker, Domitilla Raimondo, Samuel Soto, Lochran Traill and Ruth Wiseman.

1.4.3 Other case studies - Chapter 8

I have been involved in a number of other applications of MCDA and EE over the last few years. Four of these are summarised in Chapter 8 in varying degrees of detail. Two of the MCDA applications (Maclear forestry and City of Cape Town water supply) and one of the EE applications (urban open space) also had a land- and water-use management theme. The other EE application related to the valuation of attributes of a game park.

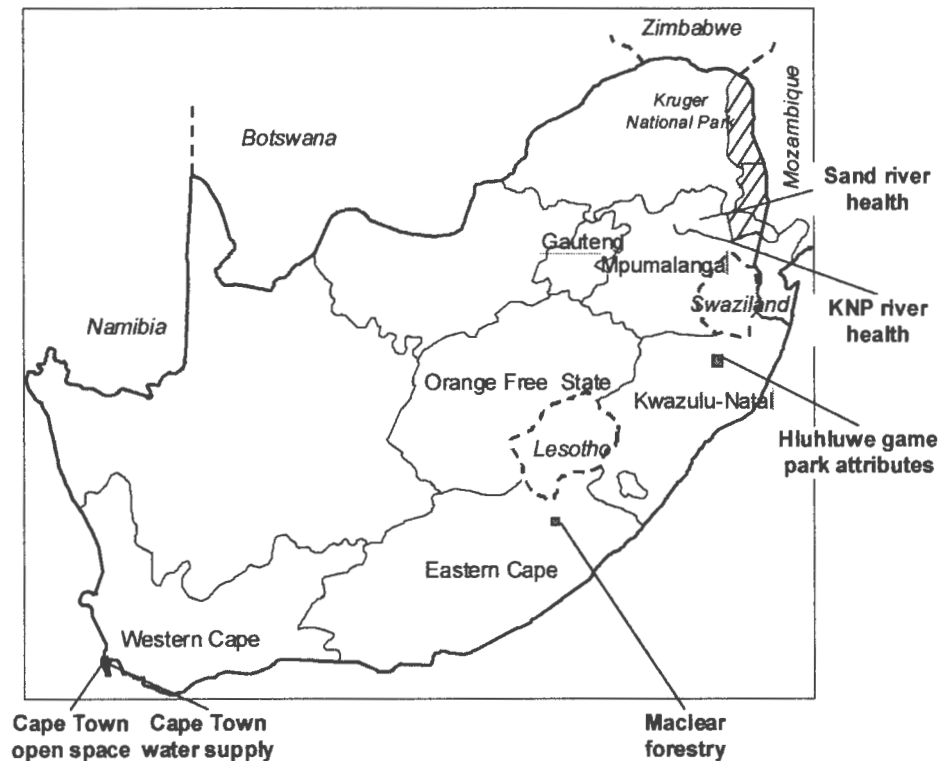


Figure 1.2. Location of case studies.

MCDA applications

Maclear district forestry and land-use

The University of Stellenbosch invited Professor Stewart and me to run an MCDA exercise to look for ‘appropriate’ levels of afforestation in the Maclear magisterial district (Figure 1.2). There was no direct client, although indirectly, the clients were the forestry company and the Dept. of Water Affairs and Forestry. The decision-aid consisted of:

- providing an overall MCDA framework,
- facilitating four MCDA workshops with representatives of various interests in the area,
- assistance with the creation of land- and water-use alternatives,
- using MCDA techniques for alternative evaluation (using thermometer scales and swing weights),
- the development of a database for the analyses, and
- analysis of trade-offs and sensitivity.

The output consisted of a report that was sent to the participants. The project was reported in brief at the South African Statistical Association conference (November 1997), formed a chapter of a Water Research Commission Report (Stewart *et al.* 2001), and a chapter of a book (Stewart and Joubert in Beinat and Nijkamp 1998). The general approach was also reported to a meeting of the Forestry Review Panel in the Eastern Cape (1997) as a possible strategic level planning tool, within which licensing decisions could be made. In the context of the NWA and a new collaborative project it is probable that MCDA methods will ultimately be used to assist with afforestation licensing decisions.

Evaluation of water supply augmentation and demand management options

We were invited to provide support for the use of MCDA to help to choose between different water supply augmentation schemes (e.g. dams, desalination) and demand management options (e.g. pressure control, tariffs) for the City of Cape Town (Figure 1.2). Several ‘thematic’ workshops were organised which we co-facilitated, during which MCDA was used to give scores to options and weight criteria. The results were used to inform decisions made by the municipality of Cape Town regarding future options for Cape Town. The results were updated a year later when three new supply schemes needed evaluation within the same framework; this was relatively simply and quickly achieved. The results were reported to the City of Cape Town in Eberhard and Joubert (2000 and 2001).

Environmental economics applications

Valuation of the contribution of park attributes to the tourism value of Hluhluwe Game Park

As part of a University of Cape Town 2001 conservation biology masters student project, we investigated the contribution of different nature reserve attributes, including the ‘Big Five’, and each of the Big Five species to the tourism value of Hluhluwe Game Reserve in Kwazulu-Natal (Figure 1.2). The survey instrument made use of scoring, ranking, the travel cost method, conjoint analysis, and open-ended contingent valuation. I would like to acknowledge Dr J Turpie who invited me to join this project and the efforts of the students, particularly George Amutete, Andrea Angel, Anthony Cizek, Emily Kisamo, Anthony Kuria, Sophie McCallum, Jo Shaw and Rowena Smuts. The analysis is still in progress.

Evaluation of different types of urban open space

The City of Cape Town commissioned Dr J Turpie and me to take part in a study to investigate the use of environmental and resource economics to assist in decision-making around the open space system in Cape Town (Figure 1.2). Several different types of open space were compared and evaluated using scoring, ranking, the travel cost method, conjoint analysis, and open-ended contingent valuation. The analyses were reported to the City of Cape Town in Turpie *et al.* (2001).

1.5 Conclusion

There are several reasons for undertaking this study including the current lack of clarity about the way and extent to which MCDA and EE differ and where they may be appropriate in application. Without yet fully justifying it, I have claimed that MCDA and EE currently constitute two different paradigms. Initially, one notes that they differ in their underlying moral philosophies, their descriptive/prescriptive/positive/normative basis, their reliance on individual or other revealed, stated, or constructed values, and the way in which they are participatory. However, they are linked by many similarities in the concepts which are used, primarily that of utility, and the types of decision-making settings in which they might be applied. The similarities and differences have a direct bearing on the appropriateness of their use in environmental decision-making.

In Part I, the foundations for the comparison are laid by outlining measurement theory, examining the origins of utility theory in economics, the development of EE and MCDA, and the details of the MCDA evaluation and EE

valuation methods. The practical applications follow in Part II. Part III includes a discussion of the case studies from the point of view of the theoretical and practical points which they highlighted, and finally, a synthesis of all of the issues, together with suggestions for a practical approach to environmental decision-making based on the discussions.

1.6 Editorial notes

Note that, where I have quoted directly from American authors or journals I have changed the spelling to the British English version. Apologies to any that this might offend. I have used “double quotation marks” for direct quotes, and ‘single quotation marks’ for other usage. I have used the personal pronoun ‘I’ where I think it helps to clarify a sentence – I have it on good authority that this is acceptable scientific form (King and Roland 1968).

Unfortunately the fields of MCDA and EE have their own extensive jargons which are not necessarily self-evident to people not within each field and acronyms also abound. I have repeated the full meaning of each acronym where it has not been used for some time before, and a list of acronyms is given on page vii, as well as a list of mathematical symbols.

PART I - THEORY AND REVIEW

2. Value measurement, value and utility theory

This chapter presents a general statement of the definitions and assumptions of value measurement, value and utility theory as a foundation for the remainder of the thesis. The concept of utility is central to EE and much of MCDA, although the interpretations are rather different, particularly in relation to how utility is measured. Measurement theory is therefore of fundamental importance in either field. MCDA, in particular, focuses on value measurement from the point of view of the axioms and theoretical assumptions that validate a particular approach, and therefore a statement of measurement theory is more-or-less equivalent to a statement of the basics of utility-based MCDA. EE, on the other hand, focuses more on the implications of value measurement for economic theory (e.g. the relationship between a particular utility function and the resulting demand function).

In essence, measurement theory provides a way of comparing objects or alternatives. Value measurement relates to the analysis of relations (e.g. longer than) between objects and the operations that may be valid on those relations (e.g. length of two objects measured together). The relations and operations (structure) are expressed as an algebraic structure with various conditions (e.g. transitivity, commutativity). A numerical structure with the same algebraic structure is then identified with the same properties (e.g. the relations $>$, $=$, and the operations $+$, $\times$). A function is then constructed which assigns numbers to the objects such that the relations and operations of the empirical and algebraic structure coincide (von Winterfeldt and Edwards 1986). Therefore, a representation theorem links objects, relations, operations, and values in a numerical system. A particular theorem rests on properties or axioms relating to the objects and relations which allow a particular numerical relation system. Within a normative framework the axioms express the minimum level of rationality required - therefore they are 'primitive', 'intuitive' or 'basic' assumptions (Bogetoft and Pruzan 1991).

In what follows, measurement scales, binary preference relations and the associated terminology are first discussed, leading to the definitions of single and multi-attribute value functions, with a brief mention of value and weight elicitation approaches.

2.1 Scales of measurement and binary relationships

2.1.1 Scales

The characteristics of objects can be measured on different types of scales. There are four scales of measurement which are used to either apply directly to an object as a representation of its characteristics, or to represent the value we place on its characteristics. *Nominal* (or categorical) *scales* place objects into categories (e.g. male or female) which are mutually exclusive. Only equivalence relations can be stated and no arithmetic operations may be performed on these. *Ordinal scales* allow objects to be discretely or continuously ranked either as a partial ordinal ranking (e.g. objects are placed into each of the categories high, medium and low), or as a complete ordinal ranking

(each object is placed in its own unique rank, typically 1st, 2nd, 3rd etc.). *Interval scales* allow comparison of intervals or differences between points on a scale, but have no absolute origin (e.g. temperature, time). *Ratio scales* are special cases of interval scales as they have an origin or absolute zero, and thus it is meaningful to compare ratios of numbers (e.g. a stone will be twice as heavy as another, whether measured in kilograms or pounds). The term *cardinal* is often used to refer to interval and ratio scales; in other words any scale which includes more information than does an ordinal scale.

Arithmetic transformations may be performed on the numbers representing order within these scales; those which preserve the same ordering as the original numbers are called the ‘admissible transformations’. Thus, an ordinal scale will preserve the same ordering under all strictly monotone increasing functions (i.e. any function g such that if $x \geq y$ then $g(x) \geq g(y)$, e.g. rank order) but linear combinations of ordinal scale values are not applicable. An interval scale will preserve the correct ordering under any positive affine transformation (i.e. $g(x) = \beta + \alpha(x)$, e.g. changing temperature measured in Celsius to Fahrenheit). Note that French (1988) and others call such transformations positive affine while Roberts (1979) calls them positive linear. Admissible transformations for a ratio scale are the similarity transformations (i.e. $g(x) = \alpha x$, e.g. changing grams to kilograms) (Roberts 1979).

2.1.2 Value and utility functions

Value, preference or utility functions allow the translation of the characteristics of an object to the value of that characteristic to an individual onto one of these scales (usually ordinal or interval). In this context, they are a mathematical representation of an individual’s preference structure. Which scale is used depends on the information which the individual is assumed to be able to supply, the rationality and other assumptions made. This thesis is concerned mainly with two types of value functions; those which reflect an ordinal scale (i.e. the rank indicated by the number is relevant and the size of the number has no further meaning), and those which allow measurement of degree of preference on an interval scale.

In the MCDA literature, the terms *utility* and *utility function*, are generally used when risk or uncertainty is involved, or in other words when the probabilities of various occurrences are referred to explicitly. Utilities which include probability are therefore also referred to as ‘von Neumann-Morgenstern utilities’ in reference to those who axiomatised this theory and approach (see Sections 2.2.3 and 3.3), and ‘utility theory’ is generally associated with the theories around maximisation of expected utility, more explicitly referred to as multi-attribute utility theory (MAUT). In economics literature the term utility is used more broadly, retaining some of its original Benthamite meaning (see Chapter 3), and may refer to ordinal, cardinal, riskless or risky values, while the term cardinal may imply interval or ratio scales. In the MCDA literature, the terms *value* and *value function*, are generally used when risk or uncertainty are not explicitly included in the function, and may be associated with multi-attribute value theory, though value functions are used elsewhere than in this theory. In economics, the term value often specifically means monetary value.

In the remainder of this thesis, utility U and value V functions are used in a general sense, relatively interchangeably, and not specifically excluding or including risk, except where the context clearly indicates that risk is included (therefore referred to as utility U) or excluded (therefore referred to as value V). This means that I switch from reference to value in the MCDA sections to reference to utility in the EE sections.

2.1.3 Binary preference relationships

The existence of a value function presupposes some substantive rationality on the part of an individual (or group), generally that either preference or indifference can be specified (alternatives are comparable) and that preference or indifference relations are transitive. More specifically, where A is the set of alternatives, $a, b, c \in A$, the binary ordering relations (R) of preference (P or $\succ$), indifference (I or $\sim$) or weak preference ($\succeq$) may be specified. These relationships may have one or more of the following properties:

Transitivity:	if aRb and bRc then aRc .
Comparability:	either aRb or bRa or both (a, b can be compared, same as 'completeness')
Negative transitivity:	if $a \not R b$ and $b \not R c$ then $a \not R c$.
Asymmetry:	if aRb then $b \not R a$. Or aRb and bRa not both true
Symmetry:	if aRb then bRa
Reflexivity:	$\forall A, aRa$

Relationships which display these properties in various combinations form various types of 'orders' (i.e. objects can be ordered to varying degrees). Thus, a relationship that is transitive is an **order**, a relationship that is comparable and transitive is a **weak order**, a relationship that is transitive and asymmetric is a **strict order**, a relationship which is reflexive, symmetric and transitive is an **equivalence** relation. The binary ordering relation of weak preference $\succeq$ is comparable and transitive, thus providing a weak order, strict preference $\succ$ is comparable, transitive, and asymmetric providing a strict order, and $\sim$ is reflexive, symmetric and transitive providing an equivalence relation (Beinat 1995, Roberts 1979).

Thus, which of these properties is satisfied determines the type of ordering relations. In general terms, if R is comparable and transitive then there exists a real-valued function V on A such that if $a \succeq b$ then $V(a) \geq V(b)$.

2.2 Value functions

2.2.1 Single attribute ordinal value functions

Given two objects, a and b , $\in A$ and the relationship, weak preference, $\succeq$ which is comparable and transitive, a real valued function V exists such that $a \succeq b \Leftrightarrow V(a) \geq V(b)$. Degrees of preference (e.g. comparison of $V(a) - V(b)$ and $V(c) - V(d)$) are not necessarily specified by the values generated by the function V , the rank order indicated by its

value is the only information supplied. The real valued function is not unique, as any strictly monotone increasing function will preserve the same rank order. Therefore, if $V(\cdot)$ is an ordinal scale and $\phi(\cdot)$ a strictly increasing transformation then $V(a) \geq V(b)$ and $\phi(V(a)) \geq \phi(V(b))$ are both valid ordering statements. For example, if $V(a) = y$ is a value function then $G(y) = 2^y$ and $G(V(a)) = \log(V(a))$ are both permissible as they preserve the rank order of the original preference. Arithmetic operations such as linear combinations of ranks or average ranks are not numerically meaningful (although this is fairly common practice).

2.2.2 Single attribute cardinal value functions and value difference functions

A 'cardinal value function', implies that the relative numeric values indicate relative worth in some sense. Unless a natural scale with an absolute zero exists, relative worth has to be indicated by the relative value of differences between objects. Thus 'cardinal' value functions are in fact given on an interval scale, where we can choose the zero and unit of measure. Given four objects, a, b, c and d , the strength of preference for a over b (represented by $\{a, b\}$) can be compared to the strength of preference of c over d $\{c, d\}$. Using the same notions as applied to comparisons of objects (e.g. $a \succeq b$), comparisons of preferences over objects can be made, using the notation $\succeq^*$ to define the relationship 'at least as strong as' between strengths of preferences. Given that $a \succ b$, and $c \succ d$, then a real valued function V can be defined such that $\{a, b\} \succeq^* \{c, d\} \Leftrightarrow V(a) - V(b) \geq V(c) - V(d)$. Thus, if the strength of preference of a over b is as least as strong as the strength of preference of c over d , a real valued function can be defined such that differences between values are quantitatively meaningful and maintain the given relationship. Thus, for clarity, these are often termed value difference functions. The function is unique up to a positive affine transformations of the form $g(V(a)) = \alpha(a) + \beta$, $\alpha > 0$ as the strength of preferences indicated by V will be maintained.

In general, for single-attribute value functions to be constructed one also needs the following conditions to apply (von Winterfeldt and Edwards 1986):

1. Transitivity of preferences: if $a \succeq b$ and $b \succeq c$ then $a \succeq c$
2. Comparability of alternatives: either $a \succeq b$ or $b \succeq a$ or both (judgements can be made about preferences)
3. 'Summation': if $a \succeq b$ and $b \succeq c$ then $a \succeq' c$ where $\succeq'$ is at least as strong as $\succeq$. Or $\{a, c\} \succeq^* \{a, b\}$ and $\{a, c\} \succeq^* \{b, c\}$.
4. 'Cancellation': if $\{a, b\} \succeq^* \{a', b'\}$ and $\{b, c\} \succeq^* \{b', c'\}$ then $\{a, c\} \succeq^* \{a', c'\}$ (given that $a \succeq b \succeq c$).
5. Solvability: an object can be found to satisfy a particular strength of preference statement (the set of alternatives is rich enough)
6. Archimedean assumption: there are no infinitely positive or negative values.

2.2.3 Single attribute utility functions

Utility functions, as demonstrated by von Neumann and Morgenstern, can be derived by observing or eliciting choice information about lotteries, and normatively rests on the maximisation of expected utility. A set of axioms similar to

1 to 6 above underlie these utility functions (e.g. see Keeney and Raiffa 1976). A reference lottery is defined with a 'best prize' A, and a worst prize B, such that four other prizes C, D, C', and D' are clearly intermediate. A lottery with probability P of winning A (and probability 1-P of winning B) is represented by [P:A,B]. Comparing the 'sure thing' prize of C to the lottery [P:A,B], there will be some probability P_c at which point the gambler will be indifferent between C and $[P_c:A,B]$, i.e. $C \sim [P_c:A,B]$, and a probability P_d at which point $D \sim [P_d:A,B]$. Similarly, assessing prizes C' and D', probabilities $P_{c'}$ and $P_{d'}$ can be found. These probabilities (probability equivalents) can be found from questioning or observation. From these four numbers we can predict which of the two lotteries [P:C,D] and [P':C',D'] will be chosen or preferred. Obviously, the decision context needs to inform how the 'lotteries' are expressed, but in essence there is always a choice between a 'sure thing' and a lottery, from which the utility function is inferred.

Expected utility or risk-based MAUT is dealt with superficially here and elsewhere in this thesis; some justification (by no means comprehensive or conclusive) for my ignoring it in the practical environmental management decision contexts is given in Section 5.3.4 and Chapter 10. The remainder of this chapter and thesis therefore concentrates on value functions or 'risk-free' MAUT, usually referred to as multi-attribute value theory (MAVT).

2.2.4 Multi-attribute value functions

The neo-classical utilitarians (Section 3.1), assumed that the 'utilities' or values of separate commodities $a, b, c, \dots$ could simply be added to give 'total utility' V . It was later realised that this implied that the value of the objects a and b were independent of each other which is not generally true, and so the generalised form $V = \phi(a, b, c, \dots)$ was introduced. This is, however, a not very useful practical form for preference elicitation, and the development of measurement theory and utility theory (a subset of measurement theory) has involved making explicit the requirements for various forms of the V function (e.g. additive, multiplicative). In other words we want to know the conditions under which the individual values, v_j for separate commodities in a bundle or separate attributes x_j of an alternative, can be aggregated in specific ways to obtain V (or how V can be disaggregated to obtain v_j). This also has relevance for the 'cardinality' of V and v_j . The discussions which follow refer equally to the comparison of the value of a number of objects in combination (e.g. V of a commodity bundle) or a number of objects that differ in a number of attributes, x_j .

Note that in economics the term 'marginal' usually refers to the incremental increase or decrease in value (utility) with an incremental change in its quantity. In MCDA, the term marginal value (utility) function often refers to a single attribute value (utility) function v_j which contributes to V . For the remainder of this thesis, the term marginal is used in its economics sense, and the term marginal value function, when referring to single attributes, is replaced with the more unwieldy, but more explicit, 'single attribute value function' or with 'partial value function'.

Under specific conditions a value function relating to more than one attribute may be an additive, multiplicative or multilinear function of the single attribute values (von Winterfeldt and Edwards 1986). These conditions relate to the types of independence which exist between attributes. In other words, in the two attribute case, to what extent are

preferences for objects with different levels of attribute x_1 independent of the level of attribute x_2 . In brief (expanded below), mutual preferential independence allows the ordinal additive form, difference independence allows the cardinal additive form, multiplicative independence allows the multiplicative form, and multilinear independence allows the multilinear form (Beinat 1995, von Winterfeldt and Edwards 1986). Given the six requirements for single attribute value functions (Section 2.2.2), the additional independence assumptions then specify the existence and form of the multi-attribute value function. The forms of independence are expanded on below.

2.2.5 Independence

The following notation applies. The set A of alternatives, $a, b \in A$, and the set X of attributes are defined, where $x_j \in X$ is a particular attribute and $\mathbf{x}^{(-j)} \in X$ is the vector of attributes excluding x_j . The notation x_j' refers to different levels of the attribute x_j . The relation $\succeq$ implies weak preference, and $(x_j, \mathbf{x}^{(-j)})$ refers to an alternative with a particular combination of levels of the attributes. The notation $\{a, b\}$ or $\{x, x'\}$ refers to the strength of preference for alternative a over b or attribute level x over x' so that $\{x_j, x_k\} \succeq^* \{x'_j, x_k\}$ means that the strength of preference for x_j over x_k is at least as strong as the strength of preference for x'_j over x_k . A single attribute value function may be expressed as $v_j(x_j)$ or $v_j(x_j(a))$, the former being more general, the latter referring specifically to the value of alternative a . The v_j is therefore the value to an individual or group of an attribute x_j . Therefore x_j refers to an attribute measurable on a 'natural' scale and v_j to a 'criterion' value, or the relative value of x_j . The term attribute is therefore associated with the characteristics of an object and criterion with the value of that attribute. The attributes levels are ordered so that $x \succeq x' \succeq x'' \succeq x'''$.

Preferential, mutual preferential independence and additive value functions

Suppose that two alternatives a and b differ only in one attribute x_j (a has x_j and b has x'_j). Attribute x_j is preferentially independent of attributes $\mathbf{x}^{(-j)}$ if $\forall x_j, x'_j, \mathbf{x}^{(-j)}, \mathbf{x}'^{(-j)} \in X$ the relationship $a \succeq b$ or $(x_j, \mathbf{x}^{(-j)}) \succeq (x'_j, \mathbf{x}^{(-j)})$ for some $\mathbf{x}^{(-j)}$, implies that $a' \succeq b'$ or $(x_j, \mathbf{x}'^{(-j)}) \succeq (x'_j, \mathbf{x}'^{(-j)})$ for every level of $\mathbf{x}'^{(-j)}$ (note that this refers to the alternative $(x_j, \mathbf{x}'^{(-j)})$ and not to the strength of preference $\{x_j, x'_j\}$). Thus, preference for an alternative that differs only on the level of one attribute, does not depend on the particular level at which other attributes may be fixed. Mutual preferential independence between the attributes $x_j, \mathbf{x}^{(-j)}$ will hold when attribute x_j is preferentially independent of $\mathbf{x}^{(-j)}$ while $\mathbf{x}^{(-j)}$ is also preferentially independent of x_j . If mutual preference independence holds, then an ordinal multi-attribute value function:

$$O(a) = \sum_{j=1}^n v_j(x_j) \quad (2.1)$$

may be constructed such that $a \succeq b \Leftrightarrow \sum_{j=1}^n v_j(x_j(a)) \geq \sum_{j=1}^n v_j(x_j(b))$. As yet no strength of preference statements have been made, and so O is an ordinal value function. This ordinal value function is unique up to a strictly increasing monotone transformation, so O is equivalent to W if $\alpha > 0$ and β exist such that $W(\cdot) = \alpha O(\cdot) + \beta$. More generally, the description above extends to all vectors of attributes $\mathbf{x}_j$ being preferentially independent of $\mathbf{x}^{(-j)}$.

Difference independence and interval scale additive multi-attribute value functions

Difference independence is also referred to as additive difference independence (von Winterfeldt and Edwards 1986).

Suppose that two objects a , b differ only on one attribute x_j , (a has x_j and b has x_j' , $x_j \succsim x_j'$) while all other attributes $\mathbf{x}^{(-j)}$ are at the same levels in both alternatives. If difference independence holds then the value difference between $V(a) - V(b)$ depends only on the levels of x_j and x_j' as the other terms are equal and ‘cancel out’. So $V(a) - V(b)$ is equal to $w_j[(v_j(x_j)) - (v_j(x_j'))]$. For clarification, suppose that two other objects c and d have the *same* levels x_j and x_j' as for objects a and b respectively, but different levels $\mathbf{x}'^{(-j)}$. The value difference $V(c) - V(d)$ is equal to $w_j[(v_j(x_j)) - (v_j(x_j'))]$ as before. Therefore $V(a) - V(b) = V(c) - V(d)$ and does not depend on the levels of $\mathbf{x}^{(-j)}$ (Figure 2.1).

The same argument may be followed but referring to the strength of preference over attribute levels rather than over alternatives, then $\{x_j, \mathbf{x}^{(-j)}\} \succsim \{x_j', \mathbf{x}^{(-j)}\}$ (therefore, $V(a) \geq V(b)$). If difference independence holds, then if all the other attributes were to change to different equal levels, $\mathbf{x}'^{(-j)}$, the strength of preference relation would stay the same, thus the strength of preference $\{(x_j, \mathbf{x}^{(-j)}) \succsim (x_j', \mathbf{x}^{(-j)})\}$ is the same as $\{(x_j, \mathbf{x}'^{(-j)}) \succsim (x_j', \mathbf{x}'^{(-j)})\}$ and $V(x_j, \mathbf{x}^{(-j)}) - V(x_j', \mathbf{x}^{(-j)}) = V(x_j, \mathbf{x}'^{(-j)}) - V(x_j', \mathbf{x}'^{(-j)})$ as before. For example, if comparing two different land-use alternatives, the difference in value in terms of effect on aquatic ecology between the two alternatives should not depend on other attributes such as the number of people employed.

Mutual preferential independence *and* difference independence allows the representation of preferences over multiple attributes through a sum of single attribute value functions so that:

$$V(a) = \sum_{j=1}^n v_j x_j(a) \quad \text{or} \quad V(a) = \sum_{j=1}^n w_j v_j(x_j(a)), \quad (2.2)$$

where the w_j are scaling constants for each attribute x_j (this weighted form has been called the ‘canonical form’ (Beinat 1995)). Difference independence allows the assessment of v_j without reference to interactions with other attributes, and the resulting V is interval scaled. Therefore, the shape of the v_j and the weight w_j do not change with different levels of $\mathbf{x}^{(-j)}$. Additive value functions are therefore unique up to positive affine transformations. Only the additive form of value function is used in the practical cases in this thesis, but the other forms are briefly mentioned below.

Multiplicative and multilinear multi-attribute value functions

As difference independence is quite restrictive, less demanding forms of independence are multiplicative and multilinear independence, the former being more restrictive than the latter.

The three attribute *multiplicative* model is $V(a) = w_1 v_1(x_1) + w_2 v_2(x_2) + w_3 v_3(x_3) + w w_1 w_2 v_1(x_1) v_2(x_2) + w w_1 w_3 v_1(x_1) v_3(x_3) + w w_2 w_3 v_2(x_2) v_3(x_3) + w^2 w_1 w_2 w_3 v_1(x_1) v_2(x_2) v_3(x_3)$ clearly allowing for interactions between attributes, where all interactions depend on the single parameter w which is pre-determined by the w_j and so preference for x_1 can not depend more on x_2 than it does on x_3 (von Winterfeldt and Edwards 1986). The compact form is $1 + wV = \prod_{j=1}^n [1 + w w_j v_j(x_j)]$. This model requires that the ordering and relative size of strengths of preferences on one attribute do not depend on the levels of any of the other attributes (although the *absolute* strength

of preference might change). In other words, if all x_j levels change, the difference $v(a') - v(b')$ should still be preferred to $v(c') - v(d')$ as was $v(a) - v(b)$ preferred to $v(c) - v(d)$ (the two attribute multilinear case illustrated in Figure 2.2 is the same as the two attribute multiplicative case).

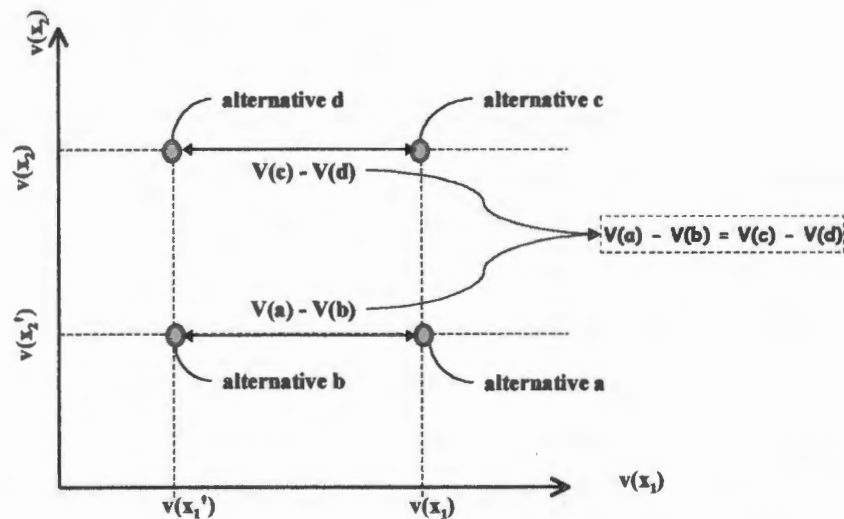


Figure 2.1. An illustration of difference independence modified from von Winterfeldt and Edwards (1986).

The three attribute *multilinear* model is $V(a) = w_1v_1(x_1) + w_2v_2(x_2) + w_3v_3(x_3) + w_{12}v_1(x_1)v_2(x_2) + w_{13}v_1(x_1)v_3(x_3) + w_{23}v_2(x_2)v_3(x_3) + w_{123}v_1(x_1)v_2(x_2)v_3(x_3)$, where there is no restriction on the interaction terms (preferences for x_1 can depend more on x_2 than they do on x_3). This model requires that the ordering and relative size of strengths of preferences on one attribute do not depend on the level of one attribute (although the absolute strength of preference might change), for *each* attribute taken *separately* (therefore there is a different w_{ij} for each of the terms in the model). So the absolute strength of preference might change differently for different attributes (Figure 2.2).

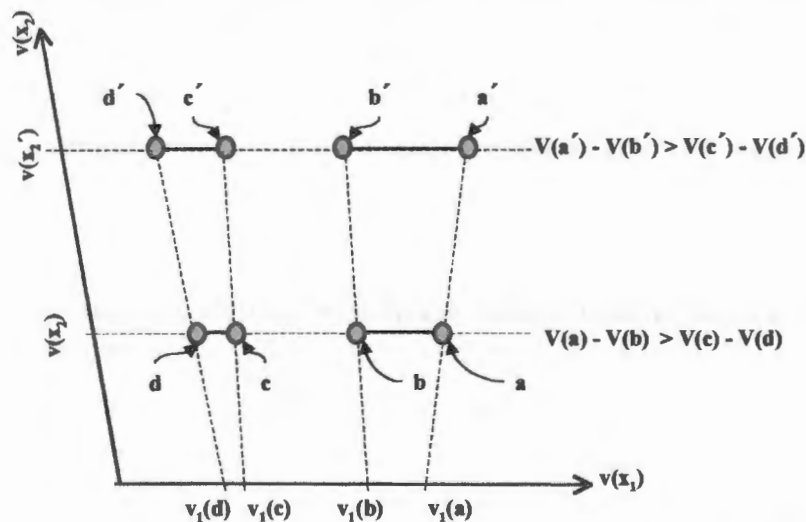


Figure 2.2. An illustration of multilinear independence modified from von Winterfeldt and Edwards (1986). Two attribute multiplicative independence is the same, but would need to maintain the same preference with all $x^{(j)}$ levels changing.

2.2.6 Value and weight elicitation

In practice many procedures exist for eliciting the values and value functions described depending on the information available. We consider situations in which an ‘analyst’ interacts with a decision-maker or group of decision-makers. Assuming that all value functions of interest are interval scaled, ‘scores’ may be given directly, in which case the analyst needs to emphasise the *differences* or gaps between alternatives. For attributes for which only qualitative information is available, ‘scoring systems’ might be formally constructed where a score is associated with a level of attribute x_j . For attributes with a natural measurement scale (e.g. number of jobs) a value function may be associated with the natural scale which indicates the changing value to the decision-maker or changing attribute levels. These may be mathematical functions of the attribute (e.g. logarithms) or constructed in a piecewise linear fashion (e.g. in the Visual Interactive Sensitivity Analysis-VISA software, 1995). Proxy measures are also used, in which case, care needs to be taken with the underlying cause and effect relationships (e.g. pollution levels as a proxy for health effects – there may be thresholds and non-linearities) (Beinat 1995). In economics, price is often used as a linear proxy measure of value as in the EE cases described in Chapters 7 and 8.

In the models above, weights w_j have been included which indicate the ‘importance’ of the different attributes. These weights have a specific meaning which is not captured purely by the abstract or intrinsic importance of the attribute. In the additive model, if the best level, v_j^* and worst level v_j^o are pegged to the same number, say 10 and 0 respectively, for all single attribute value functions, then the weights should shrink or stretch each v_j scale, so as to make the strength of preference for $v_j(x_j^*)$ over $v_j(x_j^o)$ the same as the strength of preference for $v_k(x_k^*)$ over $v_k(x_k^o)$ where x^* and x^o are the best and worst levels of the attributes, respectively. In practice a number of weight elicitation procedures are used which do not formally assess the relative strength of preference, but in an intuitive sense assess the importance of the attribute, e.g. direct numerical rating or ranking of weights. In the latter case various methods can be used to derive a numerical weight (e.g. rank reciprocal, rank sum, rank exponent). In order to emphasise the real meaning of the weights, attempts have been made to avoid the term weight altogether and refer to them as scaling constants (e.g. Keeney and Raiffa 1976). The swing weighting and cross-attribute strength of preference procedures try to formalise the concept of scaling the v_j . In swing weighting (many variations exist) the analyst may construct a hypothetical alternative where all attributes x_j are at a specific low level (e.g. worst or middle). The decision-maker is then asked, if only one attribute could be improved to its best level, which one she would choose, then which she would choose to move second, etc. The first weight is arbitrarily given a weight of 100, and the subsequent swings are given a relative percentage. The weights are then normalised to sum to one. Thus, the specific ‘swing’ from v_j^o to v_j^* is measured rather than the intrinsic ‘importance’ of v_j , and the weights reflect both the difference between worst and best, and how important that difference is. The value difference approach to weight elicitation is similar to the swing weighting procedure but more formally looks for the level at which two hypothetical alternatives become equivalent. In other words, if the alternatives (x_j^*, x_j^o) and (x_j^o, x_j^*) are compared, the decision-maker needs to decide which is better. If the first is preferred, then x_j^* is changed to x_j' until the decision-maker finds the two equivalent. This reduces to the statement $w_j v_j(x_j') = w_k v_k(x_k^*)$, and therefore $w_k/w_j = v_j(x_j')$. The procedure is repeated for all pairs of attributes. Various of the approaches to value elicitation were applied together with swing weighting in the MCDA case studies described in Chapters 6 and 8.

2.3 Summary and conclusion

Measurement theory provides a theoretical basis for comparing objects or alternatives. Value or utility functions are a central feature of this theory, and measure our value or preference for alternatives. Utility functions (in MCDA) include risk or probability, while value functions are 'risk-free'. Value and utility functions may be on ordinal, interval or ratio scales. The scale depends on the amount of information available or the degree to which preferences can be given, and the satisfaction of the axioms of transitivity, comparability, 'summation', 'cancellation', solvability and the Archimedean assumption. The overall value of objects or alternatives with many attributes may be formed by combining single attribute values (once the axioms for single value functions are satisfied). The form of the multi-attribute function (additive, multiplicative or multilinear) depends on the independence between the different attributes, in other words, the degree to which preferences can be given without reference to the other attributes.

The outline of measurement theory in this chapter, although superficial by many standards, covers most of the aspects which are essential to the theory and practice of much of MCDA and, apart from the section on value and weight elicitation (Section 2.2.6), for the rationality assumptions underlying much of economics. Although coverage in economics texts is usually briefer, "from its earliest stages, then, the analysis of value has been one of the key subjects in economic thought" (Black 1970 p. 8), and the field of EE has arisen through the struggles of economists to appropriately measure environmental values.

Chapter 3 continues the theme of value measurement by giving a brief history of economic thought as it relates to the concept of utility, its measurement and its ordinal or cardinal nature, from Bentham to the present day. This history predates the development of measurement theory as a field of research, but is crucial in understanding the use of the concept of utility in economics and in EE as described in the remainder of the thesis. The development of EE is also traced, with an introduction to some of the valuation (i.e. measurement) tools. A statement of measurement theory is synonymous with a statement of the axioms of many MCDA methods. Therefore this chapter lays the background for the discussion of MCDA in Chapters 4. However, some of the MCDA methods to be mentioned do not make reference to value measurement at all. Chapter 4 mentions the many influences on, and development of, the field of MCDA, together with some recent developments and applications. Chapters 2, 3 and 4 are therefore introductory and the material is covered in as simple (but hopefully not simplistic) a way as possible. Various subtleties are ignored, and the mathematical translations are deliberately the most straightforward possible. Extensions of MCDA to group decision-making are given in Sections 4.1.3 and 5.3.3. Some of these rely on the consideration of social choice theory, to which a brief introduction is given in Section 3.2.

3. A brief history of economics and the development of environmental economics

"But I have planted the tree of utility. I have planted it deep, and spread it wide."
Bentham in Stigler 1950 p. 307.

"Only those who consider general welfare as the algebraic sum of individual utilities require that utility be measurable in a cardinal sense." Samuelson 1938 p. 65.

The concept of utility underlies much of economic theory and the EE valuation paradigm. Although developments in economic theory have meant that reference to utility is not always necessary and often regarded as undesirable, most economic approaches to environmental decision-making, make use of the concept of utility in their development. The debates relating to the concept of utility within economics which are most relevant to this thesis, revolve around:

- the interpretation of utility as being either 'cardinal' or ordinal,
- whether the appropriate means of discovering utility is through introspection/interrogation or through observation and how utility can be measured in practice,
- whether the 'cardinal' nature of some types of utility originates from operations used in developing the function, rather than any intrinsic degree of preference implication, and
- whether cardinal utility is necessary for welfare economics; i.e. to make justifiable social choices.

These issues have not been resolved despite a century of wrangling and are of direct theoretical and practical relevance when considering which value measurement method might be appropriate in a particular context, within either the EE or MCDA paradigms.

In order to introduce the above issues, this chapter therefore gives a brief history tracing the development of the concepts of utility and of marginal utility and the 'measurability' of utility in economics. This is followed by an outline of the development of EE, which introduces some valuation techniques. The aim therefore, is to lay the foundation for the later comparison of the two paradigms, as its economics history has provided much of the 'flavour' of the EE paradigm. In contrast, although the MCDA paradigm has some of this history in common, as will be seen in Chapter 4 and later, the flavour has changed considerably.

3.1 The classical and neo-classical economists

The classical economists (Smith to Mill) were concerned with the wealth of nations and its distribution. Smith (1723-1790) considered the value in use versus value in exchange paradox, which one could view as a precursor to the debate on value and its measurement which has continued since then. Incidental to the central concern about

wealth and its distribution, was the need for valuation, as some measuring rod was needed to determine the value of the heterogeneous goods constituting wealth.

Bentham (1748-1832) placed choice and preference within a comprehensive utilitarian framework (see Figure 3.1 for a summary of the main contributors to utility theory from the point of view of this thesis). The leader of the Utilitarians and the Radical Reformers, he considered that the ‘greatest happiness of the greatest number’ (Bentham 1789) should be the basis of social decision-making, and advocated the measurement of utility for the creation of a more rational law system. The utility of something considered in isolation was the aggregate of its pleasure and pain giving properties, the four dimensions of which were intensity, duration, certainty, and propinquity/remoteness. In addition, when considering the context, one also should consider the fecundity (chance of one pleasure leading to another), purity (chance of a pleasure not being followed by a pain) and extent (number of people affected) of a pleasure or pain (Jevons 1871, Staley 1989). Concerned with the inequality of income distribution, Bentham listed 32 individual characteristics (e.g. age, sex, ‘firmness of mind’) that needed to be taken into account in ‘pleasure’ and ‘pain’ calculation. He felt this would allow interpersonal comparisons of utilities but recognised the problem of assuming the “..addibility of the happiness of different subjects..” (Bentham in Stigler 1950 p. 309) and felt that money would be a reasonable approximation of utility. J.S. Mill (1806-1873) and Ricardo (1772-1823) continued, with developments or adaptations, in the utilitarian framework. In particular, Ricardo developed his labour theory of value (Smith also felt that labour was the source and measure of value; Black 1970) and from other analysis deduced that one could give absolute value to goods while the ‘true’ neo-classicists used relative values.

The shift from the classical to neo-classical economics was partly a shift to a concern with micro-economic issues, and included the view that “..value depends entirely on utility”, rather than labour (Jevons 1871² in Black 1970 p. 77), and that utility was a) the aggregate of the feelings of pleasure and pain prevented or gained by consuming a commodity and b) dependent on the “exchange ratios determined by the “psychology of the parties making the exchange” (Black 1970 p. 11). Under the influence of Jevons (1835-1882), Menger (1840-1921) and Walras (1834-1910), utility theory became more generally accepted in economics, forming the basis of much economic theory. The cardinal nature of utility was generally accepted, and the effect of utility on demand was the focus of study. The neo-classical interest in utility was therefore in order to determine demand. If utility is an increasing function of quantity and quantity purchased is a decreasing function of price, the quantity purchased can be found once the utility function is known (Edwards 1954).

Jevons³ distinguished between the total utility U of consuming a good x , and the increment of utility du , contributed by an increment in the good dx - or marginal utility (MU) du/dx . He explored the law of decreasing marginal utility and thus began the ‘marginal revolution’ (although the notion of the marginal utility of money had earlier been explored by Bernoulli, it was at this time that it became a generally accepted law). Jevons, Walras, Menger, and

² Jevons felt that his development of utility theory had destroyed the Ricardian labour theory of value (Staley 1989)

³ Jevons dealt with Smith’s paradox of value by saying that value in use is U , while the “urgency of the desire for more” is MU (Jevons 1871 in Black 1970 p. 129). The value in exchange is the ratio of exchange of one good for another, which is equal to the ratio of MUs.

later Marshall (1842-1925), to greater or lesser extents, subscribed both to the notion of cardinal (or 'measurable') utility, which could be found through introspection, and to the general applicability of the law of diminishing marginal utility. However, they felt that the actual tools to measure utility were not yet available, but suggested that money could be used as a surrogate measure. In other words, they felt that the prices people paid indicated the utility to them of the goods in question: "...it is from the quantitative effects of the feelings that we must estimate their comparative amounts" (Jevons 1871 in Black 1970 p. 83). Jevons also argued that one could not make interpersonal comparisons of utility although he did so in practice (Stigler 1950). Marshall considered that Bernoulli's logarithmic utility function of money (Section 3.3) might be applicable to income classes where the utility gain for a ten times richer class would be about 1/10 of the utility of the poorer class for the same gain (Stigler 1950). However, it was later shown that this particular shape was a) restrictive on income elasticities, b) implied there would be no gambling and c) did not correspond with observable economic behaviour (Stigler 1950).

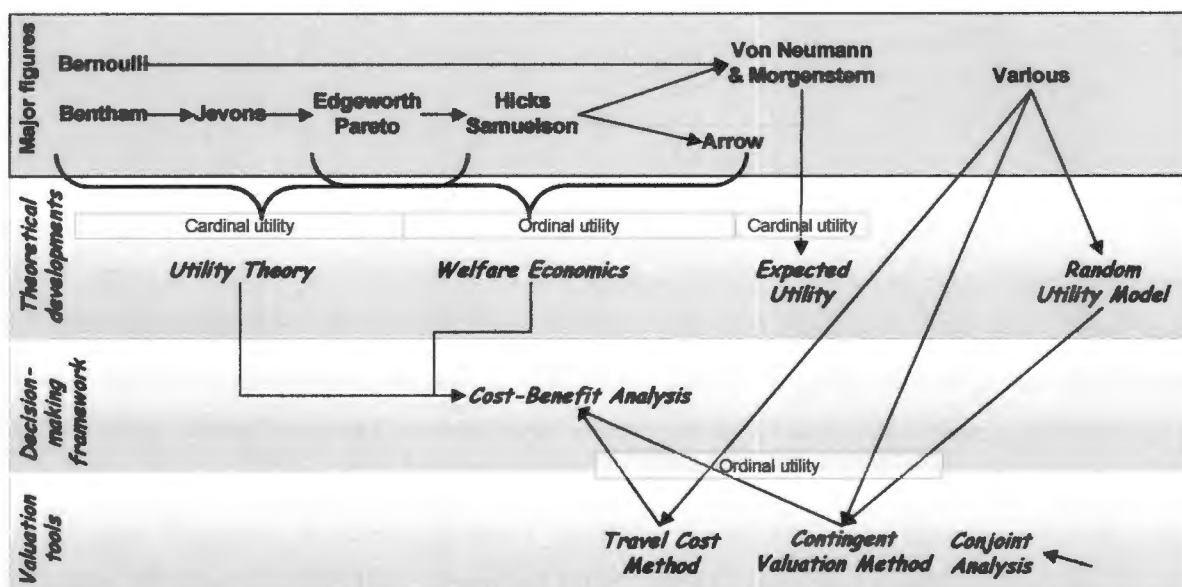


Figure 3.1. Main contributors and developments in economics and EE.

Assessing the effect of utility on demand, Walras deduced the equilibrium utility-demand relationship. Thus, a central result of this phase was the conclusion that, with perfect competition and information, the ratio of marginal utilities of goods x and y is proportional to the ratio of their prices: $MU_x/MU_y = P_x/P_y$, and that only the ratios of marginal utilities were needed for demand analysis (Stigler 1950). A more important aspect, for this chapter, was that Jevons, Walras and Menger assumed that utility was a simple function of quantity, and that the total utility of a number of commodities was the sum of the separate utilities: $TU = g(x_1) + g(x_2) + g(x_3) + \dots$. This would imply that the utilities of different goods were independent of each other.

The waning of the idea of the cardinally measurable utility and its centrality in economic theory began with the generalisation of the utility function and the development of indifference curves. Thus, in this context, the next major development was the generalisation of the additive utility form by Edgeworth (1845-1926). Concerned with the non-independent nature of the utility of commodities, he generalised the U function: $U = \phi(x_1, x_2, x_3, \dots)$. This meant that U could no longer be drawn in two dimensions, and so, still using the concept of measurable utility, he developed

indifference curves in 1881. Indifference curves showed the combinations of goods x_1 and x_2 (consumption bundles) which yielded equal satisfaction (i.e. contours where $U(x_1, x_2) = \text{constant}$, Figure 3.2), and are observable or inferable from observing riskless choices. The shape of indifference curves (convex to the origin) stems from the assumption of a diminishing marginal rate of substitution (MRS) when moving from left to right along the curve (which in turn stems from the law of diminishing MU). The MRS of good x_1 for good x_2 at any point (x_2, x_1) is given by the (negative reciprocal of the) slope of the indifference curve at that point. The slope is the measure of the goods' relative MUs (the ratio of the MU_{x_1} to the MU_{x_2}). In Figure 3.2 $U(b) > U(a)$, and $U(c) > U(b)$ but there is no strength of preference implication.

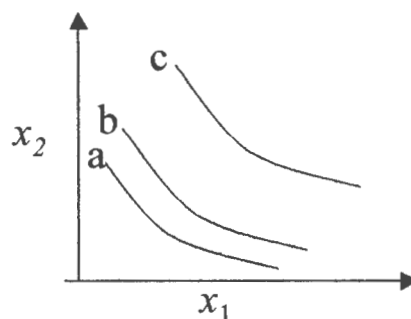


Figure 3.2. Indifference curves for two goods x_1 and x_2 , $V(c) >_1 V(b) >_2 V(c)$ but the size of $>_1$ is unknown.

The next development was Pareto's questioning of the use of cardinal utility in the 1890s: he felt that people could say that they preferred a to b , but not how much more they preferred a to b , therefore that preferences could only be indicated on an ordinal scale (Edwards 1954). In 1906 he showed that, in fact, indifference curves could be derived from ordinal information and that the same conclusions for demand could be derived from indifference curves as had previously been derived from cardinal utility. One still needed to assume that the consumer has a weak ordering, but only to find the combination of goods among which she is indifferent. Pareto was somewhat inconsistent in his use of ordinal utility in that he (a) still referred to marginal utilities, (b) felt that people could tell if $u(a) - u(b)$ was greater or larger than $u(c) - u(d)$, which implies cardinal utility, and (c) used additive utility (Edwards 1954). From the assertion that one only needed (or had available) ordinal utility, flowed Pareto's principle or criterion for welfare choices - as long as one person were better off, and no-one were worse off, social welfare would be improved (see the next section).

In other respects, Marshall and Pareto continued in the same vein as Jevons. Marshall continued with the idea of an additive U function, assuming, with Jevons, that money could be used as a proxy measure of utility. Pareto also continued to use the additive utility function, although acknowledging that it was unrealistic, saying that for small changes in quantities of substitutes or complements the effect would be small. Dupuit (1804-1866) introduced the idea of consumer surplus (see Section 3.4.1) in 1844. Pigou showed (1892) that interpersonal comparisons would need to be made in order to add consumer surpluses and Marshall eventually acknowledged this, but Marshall, Pareto and Pigou all continued with the additive form (of consumer surplus and U) as the only practical route⁴ (see Section 3.4.1) (Ross 1999, Stigler 1950). Others were also critical of the assumption of independence of goods, and

of interpersonal comparability of utilities, but relaxed these criticisms, as they were looking for practical results that could be used in income taxation or welfare analysis (Stigler 1950). By the turn of the twentieth century (and as taught in economics courses to date) only ordinal utility was needed or used.

Considering the inconsistencies prevalent, Hicks and Allen (1934) tried to “..purge the theory of consumer choice of its last introspective elements..” (Edwards 1954 p. 385). In other words, they were trying to develop a theory of choice that would not require cardinal utility and interpersonal comparability. However, this still required introspection to the extent that indifference curves could be thought of as arising from “a sort of imaginary questionnaire” (Edwards 1954 p. 385), but consumer demand conclusions could be derived without reference to even ordinal utility.

The ‘final blow’ to the perceived necessity of cardinal utility came when Samuelson showed that the theory of consumer choice could be derived (preference could be inferred) from the observation of choices made by consumers in the market rather than from an ‘imaginary questionnaire’, direct questioning, or introspection (the traditional approach). Samuelson’s concern was to “..bleach the concept of utility of all psychological content..” (Ross 1999 p. 173). The theory of revealed preference, simply put, states that for any set of prices and income (budget constraints), the commodity bundle chosen by a consumer is revealed as preferred to all cheaper commodity bundles (and with transitivity of choice among commodity bundles, utility functions can be derived from demand equations). Thus, cardinal utility was felt to be unnecessary for consumer choice theory, and because it was so restrictive, also undesirable.

3.2 Consumer choice, social choice, welfare economics and cardinal utility

The development of welfare economics in the early 1900s led to a search for alternatives to the Pareto criterion. Welfare economics as systematised in the 1960s, provides a framework within which normative judgements about alternatives can be made so that one can be said to be better than another in some way (Perman *et al.* 1996). Recall (above) that the Pareto criterion was that an economic outcome is allocatively efficient if the utility of one person cannot be improved without reducing the utility of another (i.e. if it is on the utility possibility frontier). This criterion does not require measurable utility, nor interpersonal comparisons, as it does not matter how much a person’s utility is increased or decreased, just that it is.

Hicks, Kaldor and Scitovsky pursued the refinement of the Pareto criterion (given that it would so rarely be satisfied) within the context of non-comparable individual utilities. Kaldor (1939) proposed that if those who gained could compensate those who lost and still gain, then the change would improve overall welfare. Hicks (1939) proposed that the losers should not be able to ‘bribe’ the gainers into not undertaking the move. Scitovsky (1941) proposed (in

⁴ Pigou and Edgeworth were also interested in the effects on individual utility of other’s utility. Edgeworth tried to measure this by looking at the effect of party size on wine consumption (Stigler 1950).

order to avoid reversals and incomparability results from Hicks or Kaldor) that a change would be acceptable if both the Hicks and Kaldor tests could be satisfied. When assessing a single project, the Hicks-Kaldor criterion (which leaves the question about whether compensation actually is paid open) is passed when the benefits are greater than the costs. The actual size of the benefit-cost ratio is immaterial. However, the ratio is calculated using cardinal values (money), and this makes the assumption that money has the same value (or marginal utility) for all, losers and winners. So, although the cardinal property of the ratio is not relevant, its calculation is based on the comparability of money (and by implication of utility). When comparing several projects, all may pass the Hicks-Kaldor test, but one needs to assess which produces greater benefits. Whether using net present value, benefit cost ratio or internal rate of return, the measure now indicates a rank order, still based in the comparability of money. Therefore, the Hicks-Kaldor criterion did not solve the problem of making choices between projects within welfare economics; ordinal utility might be sufficient for individual riskless choice but not social choice. Note that the developments by Pareto, Hicks and Allen were still based on the assumption that a weak order could be established, as this was implicit in the construction of indifference curves. Hicks' argument (1941), parallel to the trickle down and free-trade arguments, was that the if pursued consistently, even if compensation were not paid, in the long run, everyone would benefit, as those who lose now, may benefit later, and so on. As Little (1957) pointed out, this is no guarantee that this will happen (HET website).

A brief excursion is needed, to more formally introduce the theory behind collective choice and social welfare functions, a branch of rational choice theory (Hargreaves Heap *et al.* 1992). Much of the discussion above referred to individual choice (and therefore to consumer theory), where the usual assumption is that individuals' choices and demand are (or should be) based on the maximising of individual utility (even though, as we have seen the appeal to utility was expunged, the rationality assumptions remain central to discussions). The extension to collective or social choice is largely based on individual preferences (and therefore the same rationality assumptions). The question, when applied to projects with broader social and environmental consequences, is how to make choices that either are based on individual preferences or utility or that are in some other way appropriate because an improvement in 'social welfare' is indicated. Social welfare can be loosely defined as the overall well-being or utility of society (Douglas 1982). The rules applied to find the appropriate decision or aggregation are called collective choice rules (voting is an example). For simplicity, this discussion will assume a predominantly democratic political system and will be based on utilities that indicate preference (which may mean mental satisfaction, desires, values or choices (Sen 1997)) in some way. A more extensive and in depth discussion (e.g. other political systems, and methods based on game theory), and evaluation in terms of measures of social justice is beyond the scope of this thesis. Issues arising from this section will form an important component of the discussions in Chapters 9 and 10. In addition, this section provides the background for extensions to group decision-making in MCDA in Sections 4.1.3 and 5.3.3.

If we have found the utilities of individuals U^i for particular (multicriteria) states of nature $\mathbf{x}$ (e.g. projects, policies, consumption bundles) then the aggregation function G :

$$W(x_j) = G(U^1(\mathbf{x}), U^2(\mathbf{x}), \dots, U^n(\mathbf{x})) \quad (3.1)$$

is a collective choice rule which will produce a group preference. If we consider only the orderings O^i of alternatives which the utility functions U^i provide for each individual, then P is the set of all such orderings $P = \{O^1, O^2, \dots, O^n\}$. Then G is a collective choice rule that produces a preference relation R . If we require that G produces a group ordering as well, that is, it is an *ordinal social welfare function*, then Arrow has shown that, under some requirements, no such function exists (Arrow 1950). However, fewer restrictions on the information that can be used from equation (3.1) (e.g. cardinal utility functions rather than ordinal), may produce results that are more useful in application. The question remains as to what form the aggregation function G can or should take. The general form of equation (3.1) is referred to as the Bergsen-Samuelson social welfare function (Hargreaves Heap *et al.* 1992) and is ‘welfarist’ (Sen 1977), in that it depends on utility rather than quantity of goods (as in the version by Samuelson as described below).

In the absence of an operational alternative, the social welfare function is often taken to be additive (Bell *et al.* 1977), which corresponds to the classical utilitarian view, or with weights, is a *weighted utilitarian social welfare function*, $W = \sum_{i=1}^n \alpha^i U^i(x)$. Both forms have the obvious problem of interpersonal comparability of utilities, and the weights are either found by invoking a ‘supra-decision-maker’ who decides on the relative sizes of U^i , or by requiring the function to satisfy some other social justice criterion. Other proposed social welfare functions are the so-called ‘Rawlsian’ which seeks to maximise the minimum utility: $\max W = \min (U^1, U^2, \dots, U^n)$ (i.e. aims to improve the lot of the poorest in society) and the ‘Bernoulli-Nash’ which is multiplicative $W = \prod_{i=1}^n (U^i)^{\alpha^i}$ (where α^i are the weights assigned to the individual or household) (HET website). Other approaches or aspects (e.g. social processes which might determine the weights) are discussed in connection with group decision-making methods (Sections 4.1.3 and 5.3.3).

Samuelson also tried to ‘remove’ the need for cardinal utility from welfare economics. With the understanding that individual utilities were not comparable, he showed that the social welfare function $W = G(U^1, U^2, \dots, U^n)$ could be represented as a function of the goods consumed i.e. $U = F(X, Y, \dots)$. Then individual utility functions U^i could be translated into demand functions, and given prices and incomes, a general demand function could be derived from individual demand (using Hick’s composite commodity theorem), in turn translatable into a family of “social indifference curves” dependent on the totals of each of the goods alone (Samuelson 1956, Ross 1999). This combines individual utilities in a way which would show if the Pareto criterion were satisfied (Edwards 1954), and was based on individual preferences rather than dictatorially based. Samuelson required that decision-makers were rational to the extent that they had transitive preferences, from which ordinal value functions could be constructed. However, the derivation of Samuelson’s function $U = F(X, Y, \dots)$ requires that “...within [society] there ... take place an *optimal reallocation of income so as to keep each member’s dollar expenditure of equal ethical worth*, then there can be derived for the whole [society] a set of well-behaved indifference contours relating the totals of what it consumes: and the [society] can be said to *act as if* it maximises such a group preference function” (Samuelson 1956 in Ross 1999 p. 226, my emphasis). This social welfare function would need to take “...into account the deservingness or ethical worths of the consumption levels of each of the members” (Samuelson 1956 in Ross 1999 p. 217). The initial endowment would need to be continuously readjusted, given some desired endpoint.

It would appear from this ‘big if’ of continuous endowments and comparative “ethical worths” that Samuelson has not avoided interpersonal comparisons at all.

In general, Samuelson and others during this phase were striving to remove normative elements from economics, and transform it to a more positive discipline: i.e. economists should be concerned with predicting what will happen but not what should. However, Arrow (1950) showed that none of the above criteria nor the social welfare function satisfactorily answers welfare questions, without interpersonally comparable (and therefore cardinal) utility, although he felt there was no practical way of defining such utility. Therefore, some still hoped for the “impersonal amorality” of an interpersonally comparable cardinal utility measure (Edwards 1954 p. 390). Little (1957), for example, still argued that individual utilities were comparable, or at least empirically or intuitively evident e.g. in the difference in the value of a dollar to a poor or rich man. Being most interested in distributional issues, he proposed that once the Kaldor test were passed, the change should be examined to see if it improved the distribution of income (Freeman 1993, Jakobsson and Dragun 1996). Sen later (1970) showed that even relaxing some of Arrow’s conditions, the Pareto criterion was not compatible with mild liberal requirements of individual freedom of choice.

There was still much confusion during this transition period, as to what was meant by ‘measurability’, cardinality, and what was implied in the use of either cardinal or ordinal utilities. Hicks and Allen were the first to use the terms ‘ordinal’ and ‘cardinal’ (Fishburn 1976), as previously the term ‘measurable’ was used synonymously with ‘cardinal’, and they helped to clear up the confusion in terminology and implications. In the meantime, the idea of utility (value) difference functions was also introduced, i.e. utility functions based on the idea that people could indicate whether the utility difference between two goods was greater or less than the utility difference between two other goods (i.e. answer whether $u(a)-u(b) > \text{or} < u(c)-u(d)$). This clarified that the ‘measurability’ or cardinal nature of utility was in a relative rather than an absolute sense (Fishburn 1976). However, confusion over ‘ordinal’ and ‘cardinal’ once again reigned when, in 1947, von Neumann and Morgenstern axiomatised the concept of expected utility (see below). The ensuing confusion was cleared by Steven’s (1946) classification of scale types, by Stigler (1950), Ellsberg (1954), and subsequent development of measurement theory by for example Pfanzagl and Krantz (see Fishburn 1976).

3.3 Utility under uncertainty: ‘measurable’ or not?

We have up to now been concerned with consumer behaviour and social choice under certainty, and the bulk of this thesis is concerned with ‘riskless’ utility. Utility under uncertainty is included in this chapter for completeness, and for occasional reference from later chapters, but primarily for the re-emergence of the measurability question which arose as a result of its development.

The theory of risky choices, may be considered to begin with Bernoulli's St Petersburg paradox⁵ (1738). The solution was to consider the maximisation of *expected utility* rather than the expected monetary return. Associated with the expected utility of the monetary return was the assumption of the diminishing marginal utility of money (he used the terms 'moral expectation' or 'expected moral worth' for what we would term expected utility). Bernoulli felt that the gains yielded MU in inverse proportion to wealth: $dU = k \, dx/x$, thus proposing that U is log function of wealth $U = k \log(x/c)$ (where c is the subsistence wealth level, thus when wealth = c , $U=0$) (Stigler 1950).

The neo-classicists generally can be regarded as believing that utility under uncertainty could be found in the same way as 'riskless' utility, i.e. from asking or observing strength of preference between outcomes. The theory of utility under uncertainty was developed by Ramsey (1931), formally axiomatised by von Neumann and Morgenstern (1947) and later by Savage (1954). The theory builds on Bernoulli's notion of expected utility. Von Neumann and Morgenstern used the description of choices under uncertainty to describe behaviour in a way that was useful from a game theoretic point of view for duopoly (which, in turn became useful in the study of competitive economies and general equilibrium, Staley 1989). The basic procedure was to ask for ordinal preferences (or observe the choice) between sure outcomes and prospects (see Section 2.2.3), and thus to build up a utility function which, it turns out, is measurable on an interval scale (Edwards 1954, Ellsberg 1954), i.e. a "complete weak ordering of risky choices implies the existence of utility measurable on an interval scale" (Edwards 1954 p. 411). Von Neumann and Morgenstern therefore provided axioms for deriving interval scale utilities for individual behaviour under uncertainty, based only on observed or asked for risk-behaviour and ordinal preferences. This fitted with the prevailing economic thrust by avoiding introspective Jevonsian utility and favouring the developing revealed preference theory as a more 'objective' basis for economic theory and predictions.

Although not immediately apparent at the time, and still confusing now, the concepts and operations involved in von Neumann and Morgenstern's cardinal utility and the measurability of the Jevons era are fundamentally different (Baumol 1977a), the results obtained are not comparable, nor can von Neumann and Morgenstern utility measurement be seen as a way of estimating 'Jevonsian' utility (Ellsberg 1954). The two approaches are not measuring the same thing, and are derived for different purposes and, for example, differences in von Neumann Morgenstern utilities are not comparable. In fact, according to Ellsberg, "...only the ordinal features of the index are relevant. The only numerical operation performed is that of forming mathematical expectations, which is related to risk-behaviour: it makes no sense for example to add von Neumann and Morgenstern utilities... It would be of no aid whatsoever in formalising consumer behaviour under certainty... nor would it seem to be of any relevance in welfare evaluations (whereas a Jevonsian index might be)." (Ellsberg 1954 p. 555, 556). The underlying judgements required have no real relation to earlier notions of measurable utility as a reflection of a psychological state (Edwards 1954, Ellsberg 1954).

⁵ The St. Petersburg Paradox: A fair coin is tossed until a head appears. If the head appears on the first throw you receive \$2, if on the second throw, \$4, if on the third throw, \$8, etc. The question is how much you would be prepared to pay to participate in this gamble? The expected returns are: $\$2 \times 0.5 + \$4 \times 0.5 + \$8 \times 0.5 + \text{etc} = 1 + 1 + 1 + \dots$ i.e. your expected returns are infinite. Therefore, looking at the expected monetary value, you should be prepared to pay a large amount (limitless) to take part. But, given that there is a 50% chance that your return will be only \$2 and an 87.5% chance that it will be \$8 or less, this would be an irrational amount to pay.

A side-effects of expected utility as developed by von Neumann and Morgenstern was the “..rehabilitation of cardinal or interval-scale utility..” (von Winterfeldt and Edwards 1986) after its decline under Hicks and Samuelson. However, this rehabilitation was taken up in the field of MCDA rather than in economics (except for game theory) and EE (Figure 3.1).

3.4 Origins of environmental and resource economics

Mainstream EE analyses are rooted in neo-classical economic theory. This implies acceptance of a particular set of assumptions and values whose implications may be far reaching in practical applications, but which are often ignored in practice. A brief history of the development of EE follows touching on the various schools of thought associated with EE from neo-classical economics to ‘deep ecology’. The terms resource economics and environmental economics are often used interchangeably or as a unit as ‘environmental resource economics’. For the remainder of this thesis the term environmental economics is broadly used to refer to the set of theories and tools used to:

1. Promote the efficient use of renewable and non-renewable natural resources.
2. Decide on ‘appropriate’ levels of environmental degradation (e.g. pollution) or loss, and prescribe economic incentives to control these, and
3. Ascribe monetary costs and benefits to environmental impacts in order to help in the making of choices between projects, development or policy alternatives.

This thesis concentrates on the third aspect above. The classical economists of the eighteenth and early nineteenth century, whose concerns were essentially macro-economic, viewed the long-term prospects of growth as limited by the availability of resources, particularly land. Later, Mill’s contention was that technical progress could counteract the natural limits to growth, but that ultimately a steady state should be reached, and that economic and population growth should be limited in order to maximise social welfare⁶. Marx (and Ricardo) believed in the labour theory of value, which meant that prices were seen as a measure of labour costs. In contrast, the neo-classicists (late nineteenth century until the 1950s), shifted the perspective to micro-economics and the optimum allocation of given resources, and they perceived price to be a function of a commodity's scarcity. Thus, analyses of supply and demand, particularly using marginal analysis (which was not previously used), became the norm. Neo-classical capitalist economics states that the most allocatively efficient outcome will result from a market economy as long as conditions of perfect competition exist and there are no externalities (i.e. all social costs are reflected in prices). It cannot, however, predict the distributive outcome and this will depend on the original distribution. Since the 1950s Keynesian economics (a response to mass unemployment during the Depression), with its emphasis on government spending has grown alongside its neo-classical roots, adding macro-economic theory to the neo-classical micro-economics.

⁶ He actively promoted birth control and was arrested for his pains (Staley 1989).

Since the neo-classical and welfare paradigms replaced the classical and until the 1970s most economists believed (and some still do) that economic growth was possible indefinitely through an efficiently functioning price system (Perman *et al.* 1996). Technological change and substitution would compensate for the depletion of resources. With increasing environmental concerns since the 1970s, EE has emerged as an economics sub-discipline which acknowledges environmental constraints and externalities. Neo-classical economics views man as a selfish, rational, utility maximiser who seeks to satisfy substitutable wants or preferences. The economic value of all things (marketable goods, unpriced environmental goods, concern for future generations) is determined according to the amount of personal utility yielded and the preferences of individuals are revealed by the choices which they make. Analyses within traditional EE are therefore based on consumer (individual) preferences as expressed through prices paid or prices which the consumer is hypothetically willing to pay (for a gain) or accept (for a loss). There is much debate about whether individual preferences do or should reflect the selfish utility maximiser met earlier or a mixture of selfish (private) and selfless (public or communal) preferences, and whether the elicitation methods determine which of these is being measured (see Section 9.4 for more on this debate in the light of the case studies). The aggregate of individual preferences in any case is then taken to reflect societal preferences and social desirability is determined through the Pareto or some other criterion (Section 3.2)

Welfare economics sees this rational economic behaviour as socially desirable and government's role as to intervene where market failures exist and therefore collective welfare is not being maximised. Welfare economics is essentially normative and rests on "...clearly stated value judgements about economic organisation, income distribution, or tax policy.." (Samuelson and Nordhaus 1985 p. 682), as opposed to its purely descriptive or positivist neo-classical framework. Therefore, considering that the allocative efficiency achieved through market mechanisms (or through government intervention) will not be equitable unless equitable distribution existed originally, government action may be necessary to bring about a more equitable solution (by moving along the Pareto frontier). Improvements in technology may bring about improvements for everyone (i.e. push outwards the utility possibility frontier), while wars, strikes, or inappropriate government intervention may have the opposite effect.

EE is thus derived from neo-classical and welfare economics: how to counteract market failures, how to price all goods, and include all utility where projects have environmental effects. The term EE is associated with a range of philosophies from those who believe it should retain a positivist, descriptive function (with normative judgements being made only at macro-economic level) to those who believe analyses need to include normative judgements. The range also includes the 'techno-optimists' who believe that growth can continue forever through technical innovation and substitution and, although peripherally, the 'eco-pessimists' who believe that only zero-growth options are sustainable and compatible with environmental protection. Associated with the 'eco-pessimists' is the so called 'deep ecology' view that ecosystems and species have intrinsic value and rights not only determinable by their utility to humans. There are those who believe that macro-economic policy interventions to constrain economic growth are necessary because of physical and social limits. The latter view means that although EE normally falls under micro-economic studies, 'green accounting' indices instead of GNP or NNP are being promoted within macro-economic studies. However, EE has primarily been seen as a means of 'getting the prices right', under the assumption that the

market will 'sort things out'. Recently the field of ecological economics (incorporating the more eco-pessimist and deep-ecology views) has emerged, whose followers would prefer to see the creation of a completely new paradigm not rooted in the neo-classical framework. Ecological economists use interacting ecological-economic models and/or measures of values other than prices (e.g. energy), include MCDA within their 'toolbox', and hope to promote a more unifying theory for an economics of sustainability than that provided by neo-classical economics.

3.4.1 CBA and valuation tools used in environmental economics

The use of the environmental or welfare economic tool of cost-benefit analysis (CBA) was catalysed by the Flood Control Act of 1936 in the USA, although much of the welfare economics and theoretical foundations underlying CBA were developed later (Pearce 1983) particularly by Mishan (e.g. Mishan 1972). Thus, the approach developed before the theory. With the increasing emphasis on the value of the environment, CBA has been extended to try to include the values of goods not normally traded on markets (externalities) through the use of, for example, contingent valuation, travel cost, and hedonic pricing methods (Table 3.1 and e.g. Pearce and Turner 1990). These methods are mentioned briefly in this section and those used in the case studies are described in detail in Chapter 5.

The benefits and costs in a cost-benefit analysis are valued on the basis of individual willingness to pay (WTP) (or willingness to accept (WTA) compensation) where costs are associated with opportunity costs. Environmental benefits are associated with WTP for a benefit derived from the environment and/or WTA compensation for degradation of the environment and thus reduction of the benefit. WTP is the price paid for a good plus the consumer surplus (Figure 3.3). Consumer surplus arises as a direct result of the law of diminishing marginal utility. The price paid for each unit consumed is what the last unit was worth, where the initial units would in fact be worth more than later units (for example, units of water to a thirsty person) and so with each additional unit a larger and larger surplus is enjoyed. The concept was developed by Dupuit (1804-1866) in 1844⁷, and the term consumer surplus coined by Marshall. It is defined as the "excess of the price which a consumer would be willing to pay rather than do without the good over the price actually paid" (Staley 1989 p. 187). There are in fact five different measures of consumer surplus, the original (Marshallian) assuming that income is held constant with movement along the demand curve, and four others developed by Hicks (e.g. 1943). Hicks' measures were developed, because more correctly, utility (not income) should be kept constant, which means the demand curve must be adjusted. The four new measures are: compensating variation and surplus, and equivalent variation and surplus. Compensating measures relate to maintaining the level of utility at the level prior to change, and equivalent at the level subsequent to change. Surplus measures are used where the consumer cannot adjust the quantity of good consumed, and variation measures where consumers can adjust consumption in response to a price change (Staley 1989, Freeman 1993, Jakobsson and Dragun 1996). Extending from the individual to the market (i.e. from consumer's surplus to consumers' surplus), the market demand curve is the horizontal summation of the individual demand curves. Therefore, individual surpluses are also added (Figure 3.3) although this is really only valid if individuals' marginal utilities of money are equal (Baumol 1977b in Ross 1999 p. 263, Dixon *et al.* 1986 p. 25, Staley 1989 p. 187) (see

⁷ Incidentally, Dupuit had tried to develop a CBA approach to evaluating public works as early as 1844.

also Chapters 7 and 9 on consumers' surplus). Theory predicts that the differences between the different measures of surplus as reflected by WTP and WTA will be small (Pearce 1983), but empirical evidence seems to contradict this (Dixon *et al.* 1986, Jakobsson and Dragun 1996). The differences may be due to income effects, the effects of perceived or legal rights and the fact that WTP is constrained by income (Dixon *et al.* 1986).

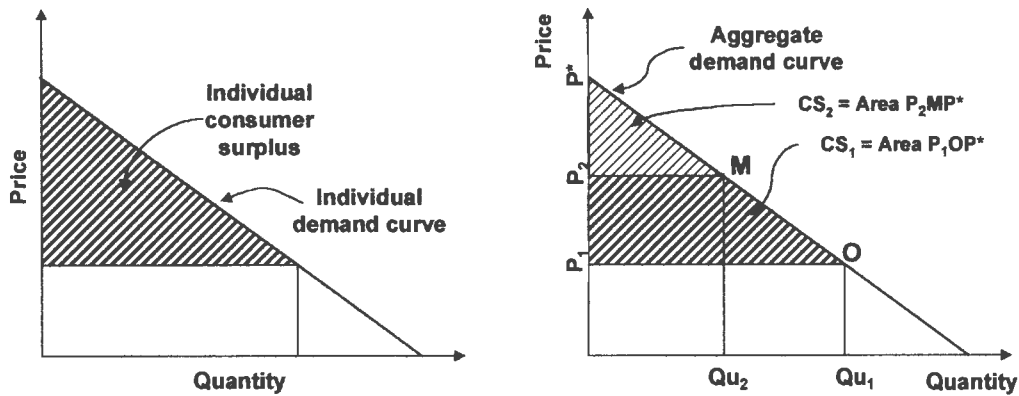


Figure 3.3. Individual consumer surplus (left) and aggregate consumer surplus (right) where in the two person case, total consumer surplus = $CS_1 + CS_2$.

Given that $WTP^i = \text{Price paid} + \text{Consumer surplus for individual } i$, for choosing between two projects a and b the cost-benefit equation is (Pearce 1983):

$$\text{Net benefit} = WTP(a) - WTP(b),$$

or where only one project a is considered, the opportunity costs of that project are used instead:

$$\text{Net benefit} = WTP(a) - C(a) \quad \text{or} \quad \text{Net benefit}(a) = \sum_{i=1}^n B_i(a) - \sum_{j=1}^m C_j(a),$$

where $C(a)$ = opportunity costs of a , which is equivalent to the foregone output of b (what an input of x in a could have produced in b). 'Shadow prices' are usually used to estimate these i.e. the prices if the world obeyed efficient pricing principles. As gains and losses accrue over time and we value gains in the future differently to immediate gains, discounting is used to find net present value (NPV):

$$NPV = \sum_{t=0}^T \frac{B_t}{(1+r)^t} - \sum_{t=0}^T \frac{C_t}{(1+r)^t}, \quad (3.2)$$

where r is the discount rate and t time. Distributional weights may also be used to try to include equity considerations:

$$NPV(a) = \alpha_A B_A + \alpha_B B_B + \alpha_C B_C + \dots - \alpha_n C_n,$$

where for example, α_A is the weight for income group A which may be derived from, for example: $\alpha_A = \bar{Y}/Y_A$ where $\bar{Y}$ is the average affected population income and Y_A that of group A (Pearce 1983). Decision criteria include NPV, or internal rate of return (the discount rate at which benefits will equal costs) and the benefit/ cost ratio.

In practical terms, market prices may exist and represent social values, or they may exist but need adjustment in order to represent social values (shadow pricing), or no markets exist and proxy values or hypothetical markets are needed (UNEP 1995). There are many market based techniques such as change in productivity, loss of earnings, opportunity foregone, and replacement costs (relocation costs and shadow projects are variants of the latter), and various non-market valuation techniques have been developed (Table 3.1).

Table 3.1. Values and techniques used in environmental economics or within CBAs.

Total Economic Value			
Use Value			Non-use values Existence, option* value
Direct use		Indirect use e.g. Ecosystem functions	
Consumptive e.g. Harvesting	Non-consumptive e.g. Recreation, tourism, aesthetics		
Market values=revealed preference			
Market/shadow/surrogate prices Changes in productivity (production functions) Replacement costs etc.		Replacement costs etc.	
Surrogate markets=revealed preference			
Travel cost method Hedonic pricing		Hedonic pricing	
Hypothetical markets=stated preference			
Contingent valuation Conjoint analysis			

*Option value is sometimes regarded as a use value (e.g. UNEP 1995).

The types of values included in EE analyses have been divided into different categories, and the techniques used to find them are shown in (Table 3.1). Values are generally divided into 'use' and 'non-use' values, and further divided into direct, indirect, consumptive, non-consumptive, option, existence and bequest values. All of the valuation methods could be used within a CBA framework, but are often used outside any actual decision-making context. The methods are also classed as either stated preference (e.g. contingent and conjoint analysis) or revealed preference (the market-based methods) approaches. As the name suggests, revealed preference approaches rely on WTP being revealed in the market directly (e.g. market prices for a product) or indirectly (property prices), while stated preference approaches construct a hypothetical market in a questionnaire. For example, with the hedonic pricing technique distance from an amenity is regressed against house prices to determine the value of the amenity, where house price is a proxy for amenity value. The travel cost method ascribes value to an amenity by developing a demand function which allows consumer surplus to be calculated from numbers of visitors and the distances travelled. The contingent valuation method asks respondents what they would be willing to pay for a change in amenity or ecosystem quality. For example, if a nature reserve were threatened by a development, a representative sample of visitors and non-visitors would be surveyed and asked how much they would be willing to pay towards, for example, a dedicated fund to help to conserve the area. The approaches which use market values have been called "objective valuation approaches" and those that use surrogate or hypothetical markets "subjective valuation approaches" (Dixon *et al.* 1986 p. 30). This thesis is mostly concerned with the latter as they are directly linked to the concept of individual utility as opposed to the objective valuation methods, and therefore can more easily be compared to utility function or value function based MCDA. The travel cost and contingent methods are described in more detail in Chapter 5, applied in Chapter 7 and Sections 8.2 and 8.3 and discussed in Chapters 9 and 10.

3.5 Summary and conclusion

This chapter has in the first instance summarised the development of the concept of utility in economics, traced its ordinal and cardinal nature and the abandonment of cardinal utility, and the application of utility within social choice and welfare economics. The ordinal/cardinal debate, far from being over a century later, is just as relevant as before.

The main reason for this is that ordinal social welfare functions or criteria may be of little practical use (e.g. Pareto improvements are rarely achieved) and the implication is therefore that interpersonally comparable cardinal utility is still needed (as shown by Arrow). Secondly, this chapter introduced EE as a discipline emerging from neo-classical and welfare economics. CBA, an applied EE decision-making framework, was then briefly introduced, together with various EE valuation tools which are used either within a CBA or in some other decision-making framework, or outside of any particular decision-making context. Following the discussion of the field of MCDA in the following chapter, three specific EE valuation tools are detailed in Chapter 5, together with one MCDA approach.

In effect, EE and CBA are “applied utilitarianism” (Ross 1999, p. 370); a collection of theories and methods developed to help us to choose between projects which have effects on individual utility. Philosophers and economists have long debated the political, philosophical, theoretical and methodological dimensions of social welfare criteria and social choice rules and their interactions. This chapter has therefore tried to highlight some aspects of this debate which are relevant to the extension of the domain of the debate to include the environmental consequences of choices. This extension arose with the relatively recent recognition that individual utility and social welfare are affected by the quality of the environment and that individuals (and society) have preferences for different environmental qualities.

The fundamental economic assumptions of the EE paradigm were laid out in this chapter. These are that WTP is an indicator of individual utility and that aggregate individual utility, as reflected by WTP, indicates social welfare. The practical means of discovering willingness to pay are discussed in Chapter 5.

The debates in the final two chapters (Part III) refer frequently to the issues raised in this chapter both to inform the discussion of the case studies (Part II) and to place the discussions in a broader context.

4. Types and applications of MCDA

"The most trivial lesson of human engineering is to design tasks mercifully, so that they will be easy to perform. Decision analysts do this less often than they should."

Von Winterfeldt and Edwards 1986 p.xi.

The basics of measurement theory as relevant to utility-based MCDA were outlined in Chapter 2. This was followed in Chapter 3 by a discussion of the development of the concept of utility within economics in the 19th century, the axiomatisation of risky choices by von Neumann and Morgenstern (1947) and the development of environmental economics. The history, after von Neumann and Morgenstern, is continued in this chapter in the context of the development of the broad field of decision analysis (Figure 4.1). The main 'fields' of MCDA (Sections 4.1.1 and 4.1.2), approaches to group decision-making (Section 4.1.3) and dealing with uncertainty and imprecision (Section 4.1.4) are introduced. Practical applications of various MCDA methods to environmental decision-making (mainly from the non-MCDA literature) are described in Section 4.2. Throughout, more emphasis is placed on value function (or utility) based methods, as value function approaches were used in the MCDA case studies in this thesis, and utility forms a commonality between EE and these MCDA methods. The aim of this chapter is therefore to provide the background and context for the description in Chapter 5 of the particular methods used, for the case studies in Part II and for the philosophical considerations of the later part of the thesis (Part III). Chapter 5 describes the EE and MCDA methods applied in the case studies, and a brief comparison of different MCDA methods in support of the choice of the MCDA method actually applied.

4.1 General background and schools of MCDA

Decision analysis is generally considered as part of the operations research field which emerged during World War 2. Developments in linear and goal programming, economics (including various concepts of utility), Bayesian statistics, measurement theory (e.g. Scott and Suppes 1958) and psychology fed the development of decision analysis (Figure 4.1) which began to emerge as a field during the late 1950s and early 1960s at the Harvard Business School. For example, Schlaifer published "*Probability and Statistics for Business Decisions*" in 1959 and in 1961 Raiffa and Schlaifer published "*Applied Statistical Decision Theory*" (von Winterfeldt and Edwards 1986) and Aumann (1964) considered the use of utility functions in multi-attribute problems (particularly in relation to the completeness axiom of von Neumann and Morgenstern's expected utility (Goicoechea *et al.* 1982)). The theory began to be more applied thanks to the stimulus of interest from the USA military. Howard, the first to use the term decision analysis, published "*Decision Analysis: Applied Decision Theory*" in 1966. He emphasised the use of monetary variables as measures of value, making his version closer to cost-benefit analysis (CBA) than any other decision analytic technique (but his version does not rely on observed willingness to pay). Then in 1968, Raiffa published "*Decision Analysis*", developing the main ideas of multi-attribute utility theory (MAUT), expanded in Keeney and Raiffa (1976) which also included multi-attribute value theory (MAVT). Simon also made important contributions to the

development of the MCDA field in his discussion of, for example, goal hierarchies, criteria weights and means-end relationships as early as 1945 (e.g. Simon 1945 in Wenstøp unpubl.).

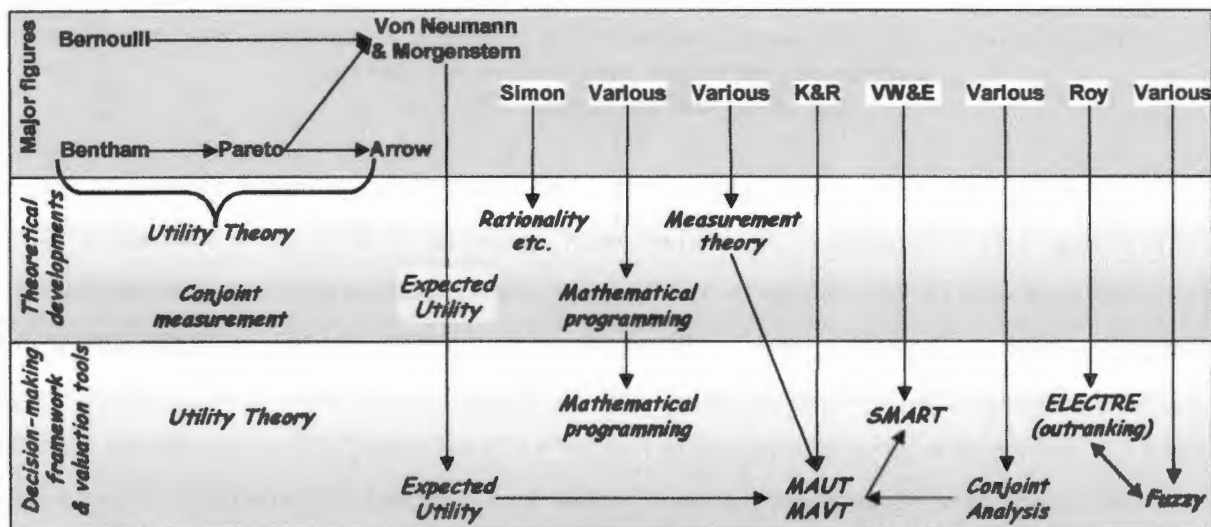


Figure 4.1. Development of MCDA. K&R=Keeney and Raiffa, VW&E=Von Winterfeldt and Edwards.

The development of decision analysis was associated with a shift from a primarily descriptive and positive economics to an emphasis on developing prescriptive decision-aids (generally now referred to as MCDA, the term appearing in the 1970s) on the one hand and descriptive models of decision-making behaviour (behavioural decision theory) on the other. Although for some decades into the development of decision analysis, there was confusion about these two distinct aspects, the debate seems to have largely been clarified (e.g. Edwards 1992). MCDA is regarded as prescriptive in that, if the fundamental axioms (Section 2.2) are considered acceptable or desirable attributes of rational decision-making, then the alternative with maximum value or utility or evidence in its favour should be preferred. However, practitioners and theoreticians supporting a variety of different types of MCDA (some of which do not refer to this axiomatic basis) emphasise that the primary purpose of MCDA is to facilitate communication, learning and information processing (Belton and Stewart 2002, p. 4) and the idea that "...there is only the provision of a framework in which to think and communicate" (French 1989 p. 1).

These various influences have led to the development of distinct MCDA schools and the accumulation of a multitude of approaches. These are conventionally divided into the discrete *choice* methods or multi-attribute decision-making (MADM), and those which are used to *construct* or design an optimal alternative (or a set of alternatives) or multiple objective decision-making (MODM) (Malczewski 1999). The MODM or optimisation methods, traditionally associated with continuous attributes, and/or an infinite (very large) set of alternatives, can also be applied to discrete alternatives. More recently, fuzzy methods have been developed to deal with imprecision and applied within both MADM and MODM contexts. A simple taxonomy of methods is given in Table 4.1, and some of these methods are outlined in this chapter. Other taxonomies exist, such as those that relate to the type and/or timing of information flow. These include: analyst to decision-maker or decision-maker to analyst information flows (Cohon 1978), and 'prior' or 'progressive articulation' of preferences (Goicoechea *et al.* 1982, Bogetoft and Pruzan 1991). This thesis is concerned with discrete choice or ordering of projects with effects that are not continuously nor necessarily

quantitatively measurable. It is therefore primarily concerned with MADM methods, and specifically with value measurement theory based methods. The discussions and applications therefore refer mainly to these methods, while other MADM and MODM methods are introduced in this chapter for completeness.

Table 4.1. Simple classification of MCDA methods

Alternative generating, optimisation MODM	Discrete choice, classification and ordering MADM		
	Value measurement theory	Other additive models	Other
Goal programming, integer programming, reference point methods, compromise programming	MAUT, MAVT e.g. SMART	e.g. AHP	Outranking e.g. ELECTRE, PROMETHEE

4.1.1 Multi-attribute decision-making

Multi-attribute utility and value theory approaches and SMART

Multi-attribute theory and practice is now divided between that incorporating uncertainty through probabilistic approaches and decision trees etc. (MAUT) as formalised by Keeney and Raiffa (1976) and that either ignoring uncertainty (MAVT) or incorporating it in more *ad hoc* ways. The fundamental ideas were outlined in Chapter 2 and Section 3.3.

Most important in the development of MCDA, from the point of view of this thesis, was Edwards' publication of his first paper on SMART (simple multi-attribute rating technique, Edwards 1971), an attempt to make Raiffa's MAVT/MAUT ideas more practical (von Winterfeldt and Edwards 1986, p. 567). Frustrated with the complexity of MAVT, Edwards developed SMART which "ignored measurement theory and non-additivities and instead relied on simple additive models, numerical estimation techniques for eliciting single-attribute values, and ratio estimation of weights" (von Winterfeldt and Edwards 1986 p. 278). Although Edwards (1977) suggested that usually linear value functions were sufficient, obtainable or desirable, the method does not imply linear value functions. SMART has evolved since first proposed and similarly simple approaches are generally used but with more theoretical validity. The original SMART was extended to include swing weights and therefore referred to as SMARTS (although SMART usually implies SMARTS these days). A variation was developed which made use of the rank orders of weights rather than eliciting swing weights (referred to as SMARTER, Edwards and Barron 1994). More emphasis is now also placed on the problem structuring stages. SMART (combined with alternative development) is the only MCDA technique applied in the cases described in this thesis and it is described in more detail in Section 5.3 together with a brief justification for its exclusive use. A full description of SMART is delayed until this stage so as to describe all the methods that were applied in the case studies in one chapter.

One approach to value function development or elicitation (see Section 2.2.6) in MAVT is conjoint measurement and extensions such as conjoint scaling. Another extension, conjoint analysis, which was developed in market research, is the holistic rating or ranking of multi-attribute profiles and is applied within both EE and MCDA and thus forms a bridge between the two. Conjoint analysis is described in Chapter 5 (within the context of stated preference EE methods) and was applied in the EE case studies in Chapter 7 and Sections 8.3 and 8.4.

Various classifications of MAVT techniques have been proposed, the main division being variously into compositional/decompositional (Green and Krieger 1993), holistic/decomposed (Beinat 1995, p. 49), or aggregation/disaggregation (Jacquet-Lagrèze 1990) techniques. Some of the procedures found in the literature were classified into these categories by Beinat (1995):

- Compositional or Decomposed or Aggregation: Single attribute value functions v_j and weights w_j are assessed separately, and overall value V is built by combining the parts (additively if difference independence holds). Specific approaches designed for real decision-aid applications include SMART, MACBETH, and COMPAIRS (see Section 4.1.4 for the latter two). In other words depending on ones' point of view, the overall value function is 'composed' from single attribute utilities (Green and Krieger 1993), the 'decomposed' single attribute utilities are evaluated (Beinat 1995), or the overall value function is formed by 'aggregation' of the single attribute utilities (Jacquet-Lagrèze 1990).
- Decompositional or Holistic or Disaggregation: Assessment is made of the value V of a multi-attribute profiles, allowing subsequent estimation of weights and single attribute value functions (v_j) by regression or optimisation by linear programming. Conjoint analysis (see Chapter 5) is one of these methods.
- Combined: Some methods make use of both approaches iteratively (e.g. PREFCALC, Jacquet-Lagrèze 1990), where, for example, initial overall judgements of V may define value functions and weights, which may then be refined through more detailed assessment.

Common to all of these methods in application are normative assumptions about rationality, a requirement for checks on preferential independence, and an acceptance that all criteria are fully compensatory.

Analytic Hierarchy Process

The analytic hierarchy process is also an additive model, which estimates scores and weights by pairwise comparisons, and eigenvector analysis. Briefly, the original AHP used a nominal 9 point scale of pairwise preferences of alternatives which were represented linguistically to the user (absolute preference for a vs. b , through indifference, to absolute preference for b vs. a). These statements of preference were then translated to the values (1/9, 1/7, 1/5, 1/3, 1, 3, 5, 7, 9) within the preference matrix (Saaty 1980). The use of linguistic or semantic statements of preference is appealing and AHP is a popular method that has been widely applied in practice, including in contexts very similar to those of the MCDA cases here (e.g. Ridgley and Rijsberman 1992 and 1994). A number of researchers have built on the ideas of AHP in various ways (Section 4.1.4). However, AHP has been subjected to criticism for various reasons (e.g. Stewart *et al.* 1993), the criticisms relating to the 'fundamental scale' used to quantify human judgement, the use of the eigenvector to estimate the impact scores, and the use of an arithmetic aggregation rule. Disregarding these criticisms, the relative simplicity of the cognitive tasks and the additive structure mean that it might well have been applied in the case studies described here. However, as discussed in Section 5.3.4, we felt that the SMART approach applied was more transparent and less tedious to users.

Outranking approaches

The outranking approach (or 'French school') avoids some of the issues arising in utility theoretic or MAVT approaches (e.g. the assumption of commensurability of alternatives and perfectly compensatory criteria) by calculating 'concordance' and 'discordance' scores. These summarise how often an alternative is preferred to another across all (weighted) criteria or dis-preferred, respectively. Outranking emphasises the fact that we do not have infinite discriminatory powers, and requires the establishment of 'indifference thresholds' q_j for each attribute or criterion. These are increments that are large enough to enable the decision-maker to say $x_j' \succsim x_j''$ (or $a \succsim b$ on criterion x_j), thus level x_j' is preferred to level x_j'' if $x_j' - x_j'' > q_j$. The indifference threshold may be defined so as to vary with x (e.g. as x_j increases the increment q_j might increase). For each criterion it is established in this way whether alternative a outperforms alternative b . Strict preference thresholds are also sometimes used together with the indifference thresholds. A concordance score, which may be a sum of the weights of the criteria for which it has been established that alternative a is preferred to b , can then be developed for each alternative. Discordance is established in a similar way, but now assessing on which criteria an alternative is outperformed. Often the concept of a veto threshold t_j is introduced: if the outperformance of a is sufficiently large on one criterion x_k , such that $x_k(b) - x_k(a) \geq t_j$, this can veto the preference for a over b . The weights used in outranking methods are no longer scaling constants as in MAVT but refer more to the intrinsic importance of the criteria. Examples of outranking approaches are the series of ELECTRE approaches (e.g. Roy 1990), and PROMETHEE (e.g. Brans and Mareschal 1990).

An example of an application of outranking to environmental decision-making is Abu Taleb and Mareschal's (1995) use of PROMETHEE V to choose among various water resource development options. The 42 options were grouped into technical, managerial, pricing, and regulatory options and the 18 objectives / criteria were ranked and weighted. The criteria in PROMETHEE are classified into six types of 'generalised' criteria whereby a preference function $P_j(a,b)$ measures the decision-maker's preference for alternative a over alternative b on the interval $[0,1]$ with respect to criterion j such that if $P_j(a,b) = 0$ the decision-maker is indifferent and if $P_j(a,b) = 1$ there is strict preference of a over b . A multi-criteria preference indicating the overall preference for a over b is then a weighted sum of the preference for a over b for each criterion. The preferences for a over all other options were summed, the preferences of all options over a were summed and the difference between these gave the 'net flow' from which the overall ranks were derived.

4.1.2 Multiple objective decision-making

It is beyond the scope and focus of this thesis to describe the various MODM methods in detail nor all the ways in which they can and have been applied. This brief discussion is included for completeness. In World War 2 the field of operations research was developed for analysis of military strategies, and linear programming developed from this through the work of Dantzig (1951) and Charnes and Cooper (1961) and work on water resource management (Maass *et al.* 1962). The intense work on water resource problems in the USA thus formed a common developing ground for both CBA and mathematical programming MCDA approaches. The general setting is that of elements, attributes or decision variables Z , of a set of alternatives A . An alternative is defined by a function of these attributes, usually assumed additive: $a = \sum_{j=1}^n \alpha_j z_j$. The task is to find a that maximises all z_j (where all attributes are

defined in the maximising sense), or has all $z_j \geq g_j$ where g_j are goals or reference points which may be defined for each z_j , subject to various constraints. However, all z_j cannot normally be maximised, or all goals achieved, simultaneously.

Various approaches can be followed to exploring this problem (e.g. Bell *et al.* 1977, Belton and Stewart 2002). One can informally explore the set of solutions with each z_j maximised or each $z_j \geq g_j$. Weights can be defined so that $a = \sum_{j=1}^n w_j z_j$ and each set of weights produces a point on the efficient frontier. The effect of changing the w_j can then be explored. More formally, one can define deviational variables $\delta_j = z_j - g_j$, i.e. distance of the achieved level of each z_j from the goal or reference point g_j for each z_j . Some aggregate measure of these deviations, or a norm, is defined and minimised, e.g. $\min \sum w_j |\delta_j|$, or $\min [\sum w_j (\delta_j)^2]^{1/2}$. One that is considered to be robust (Stewart *et al.* 2001) is the modified Tchebycheff norm: $\min [\max(w_j \delta_j) + \epsilon \sum_{j=1}^n w_j \delta_j]$. With ϵ a small positive number (e.g. 0.01), this norm seeks a balance between the importance placed on the biggest deviation and the average deviation. Lexicographic methods are also often used: each objective is maximised successively, and the level achieved used as a constraint in the next iteration. Some methods (called methods with progressive articulation of preferences) interact with the decision-maker to obtain feedback about the presented solution, which is then modified iteratively until the decision-maker is satisfied with the solution (e.g. Goicoechea *et al.* 1982) (these methods therefore find local optima).

MODM methods are particularly relevant at a technical screening stage, but less so when more intangible or qualitative issues need to be included. They can be used to construct or design projects or plans, to select an optimal alternative from an infinite or discrete set, and to generate value functions from partial information (e.g. Belton and Stewart 2002). Considerable research and application in this field has led to the development of interactive approaches, reference point approaches, and approaches for non-linear goal programming. MODM approaches are extensively used in water resources management (e.g. see the *Journal of Water Resources Planning and Management*).

4.1.3 Extensions to group decision-making

As individual preferences and rationality are extended into social choice within economics, one practical application of which is CBA, so individual rationality in terms of multiple criteria can be extended to group or 'social' MCDA. Decision-making around projects and plans with environmental effects is by its nature an interdisciplinary (although there may ultimately be one person with the responsibility of final assent or otherwise) and multicriteria task. EE and MCDA are two responses to this task. The inclusion of individual values in social choices in economics occurs through the invocation of collective choice rules and social welfare functions (Section 3.2). The theory of MCDA is usually expressed in terms of individual decision-making (e.g. value measurement theory in Chapter 2). Extensions of MCDA to group decision-making (i.e. via appropriate collective choice rules and social welfare functions) are introduced here, including the basis of the group decision-making model adopted in the case studies. This latter is explained in more detail in Section 5.3, together with SMART, the MCDA method applied in the case studies.

We are concerned here with prescriptive and/or normative decision-making, and therefore with ethical and moral criteria (rather than with game theory), as well as with issues of process and practicality. We consider the identification of solutions that maximise group utility, value or welfare in some sense. This means there is a requirement for explicit or implicit aggregation of utility, value, preferences or welfare and aggregation functions or processes (i.e. collective choice rules) need to be established. Social welfare W might be defined as a function of individual value: $W=G(V^1, V^2, \dots, V^n)$, where G is an aggregation rule (e.g. weighted summation), and the aim is to choose the project or plan which maximises W . As we saw in Section 3.2, no ordinal social welfare function could be found which would provide a group preference ordering from the set of individual orderings (O^i rather than V^i being used) under certain conditions. However, Sen showed (1970) that if we allow individual utilities or values V^i (rather than orders) to be used, and allow (some form of) interpersonal comparison of utility, we may develop a group decision rule G (or social welfare function) to find the complete social ordering (overall value) of alternatives $\mathbf{a}$:

$$W(x_j) = G(V^1(\mathbf{a}), V^2(\mathbf{a}), \dots, V^n(\mathbf{a}))$$

This equation is the same as equation (3.1) except that we have reverted to the use of value and V rather than utility and U . Note that here the V^i are the individual's (*multicriteria*) values arising from equation (2.2) or (2.1). Our focus will remain on risk-free value measurement-based MCDA (MAVT) and in particular on additive aggregation to find V^i using equation (2.2), and the additive form to find W :

$$W(a) = \sum_{i=1}^n k^i V^i(a) \quad (4.1)$$

(where the i superscripts refer to individuals i , whereas j subscripts refer to criteria).

Belton and Pictet (1997) categorised group MCDA decision-models into those that use sharing, aggregating and comparison of the elements (criteria, evaluations, alternatives and aggregation rule). Following Goicoechea *et al.* (1982) we consider three types of models for finding G . We have slightly renamed these models to avoid confusion with terminology in the rest of this thesis, and to conform more closely to the categorisation of Belton and Pictet (1997):

- Group value-aggregation models,
- Group value-sharing models, and
- Informal models.

Note that, the group value-aggregation models are those which are most closely linked to conventional social choice theory, and can therefore be considered (in certain situations) to be social welfare functions, whereas the group-value sharing model is couched in terms of group decision making rather than social choice.

Group value-aggregation models

The group value-aggregation models require explicit comparisons of value in some form, by estimating the scaling constants or weights k^i . This means estimating the 'importance' or 'experience' or 'power' or 'relative need' of the individual in the group. Which of these would apply would depend on the decision-context and would have to either be agreed to by the group, or decided by a 'supra-decision-maker' (Kersten undated). Essentially, an additional step is added to the traditional individual MAVT process, where the k^i are found. Bodily (1979) proposed a method

whereby an individual's k^i are applied to the remainder of the group, and the resulting V is substituted for his own V^i in equation (4.1). Harsanyi (1955 in Hargreaves Heap *et al.* 1992) proposes that one consider the situation where there is an equal chance of being an individual i in social state a . Applying expected utility theory there is a utility number associated with each consequence (individual i in state a). Preference between social states is then consistent with maximisation of expected utility, and because of the equal probability of being any of the individuals this is the same as the maximisation of the sum of utilities (Hargreaves Heap *et al.* 1992). Brock (1980) developed a Nash-Harsanyi bargaining version which is based on relative needs so that utility gains are equal to relative need. The solution should be equitable, efficient and unique (Goicoechea *et al.* 1982) and is the same as that found by solving the Nash-Harsanyi product $\prod_{i=1}^n (U^i - d^i)$ (where d^i is the *status quo* which is “interpersonally calibrated” (Goicoechea *et al.* 1982)), and therefore the k^i do not have to be calculated, only the U^i and d^i need to be found (a weighted version does, however, also exist).

In keeping with the focus of this thesis, we assume that criteria values v_j (which form the V^i for each individual) are measurable on an interval scale whether we want the V^i to be ordinal or quasi-interval (equation (2.1)) or interval (equation (2.2)). Remember from Section 2.2.5, that for V to be on an interval scale, the criteria v_j need to be additive difference independent. To allow for strength of preference comparisons between individual's multicriteria V^i and their summation, the V^i therefore need to be on at least an interval scale. Therefore, to aggregate the multicriteria V^i we require slightly stricter independence conditions of the original criteria v_j (additive difference independence) than if we were only dealing with one individual and did not need V to be more than ordinal.

There seem to have been few attempts formally to extend social choice theory (in particular in the group value-aggregation mode) to multicriteria decision-making and we have concentrated on the MAVT weighted summation model. Arrow and Reynaud (1986) considered an extension to ordinal models after examining the outranking approach. Van den Honert and Lootsma (1996) examined the application of the value-aggregation model to finding group weights in the context of the multiplicative AHP (Section 4.1.4). They showed that the resultant choice will be Pareto optimal (in the sense that if all members have $A \succ B$, then the group will have $A \succ B$). An approach which informally combined aspects of the group value-aggregation model and the informal techniques was applied in a practical setting by Edwards (1977) using SMART, where individuals gave their own weights, then after feedback from the group these were updated. Much of the literature examines variations of ‘power coefficients’. For example, Barzilai and Lootsma (1997) proposed that ‘power coefficients’ applied within multiplicative AHP or SMART should indicate the relative power (e.g. size of the constituency, or country that is represented) of the group member or ‘expert weighting’.

A group value-sharing model

Criticisms of group value-aggregation models exist from different points of view. They rely on explicit comparison of *individual* utility. In a small group, the former requirement may lead to hostility within the group. Where decisions taken by a small group will affect broader society, it may lead to inequities and biases. The approach also does not

conform to conventional decision-making processes. Rouse and Sheridan (1974 in Goicoechea *et al.* 1982) suggest that conventional group decision-making is 'goal-oriented' in the sense that, in a group process, a goal or common objective (at least in some, perhaps fuzzy, mother-earth-and-apple pie sense) is established, pre-determined or imposed (by the ultimate decision-maker such as the government). Alternatives are examined according to different points of view within this common goal, and the synergistic interactions of a group are the main reason for having meetings (workshops, decision conferences). This is in contrast to the group value-aggregation model where, first individual values are considered, and then the group or a supra decision-maker makes rules about aggregation (i.e. estimates weights). A simple approach to dealing with some of these criticisms was formalised by Krzysztofowicz (1979 in Goicoechea *et al.* 1982) and was called a *group-utility model* by Goicoechea *et al.* (1982) which I have renamed for consistency the *group-value sharing model*. This model is described here in more detail than the group value-aggregation model, as it forms the basis of the model applied in the SMART case-studies and of the proposed approach in Chapter 10. Krzysztofowicz's model is described by Goicoechea *et al.* (1982) in terms of utility functions, but I have adjusted this in terms of value functions and made other adjustments to the notation to be consistent with the rest of this thesis.

In Krzysztofowicz's model (1979) individual values and interpersonal comparisons are not found or formed explicitly. Rather the group value functions and aggregations are found directly. Through the group interaction, the group transforms individual judgements into group judgement. In essence, each partial value function $v_j(x_j)$ is instead a group partial value function for that criterion $g_j(x_j)$ and forms part of W directly. To distinguish between this 'group welfare function' and the others we will refer to it as GV . Krzysztofowicz defines (G_M, d) as the Group Decision Maker, where G_M is the set of the members of group, and d is the decision rule applied to make the individual to group translations. Krzysztofowicz proposes the following assumptions (from Goicoechea *et al.* 1982):

1. The group value function GV can be disaggregated on the basis of attributes x_j or criteria g_j . There is an operation H such that $GV = H(g_1, g_2, \dots, g_n)$, where g_j is group partial value function on x_j (i.e. a group-based single attribute value function),
2. Individual value judgements are non-deceptive expressions of preference,
3. The group value judgement is an expression of group preference from the application of d . This may not agree with any individual value judgement. The decision rule may involve processes such as debate or voting. The decision rule may be implicit or explicit,
4. Each member and G_M behaves in accordance with value measurement axioms.

With the acceptance of assumptions 1 to 4 a group value function GV can be found. A fifth additional simplifying assumption is made in Goicoechea *et al.* (1982), that the G_M determines (not necessarily disjunct) subgroups, G_M^k . The subgroup memberships correspond to expertise in a field, and the number of G_M^k s corresponds on a one to one basis to the number of criteria. The g_j will be the G_M^k 's marginal value functions on x_j , and are accepted as G_M 's marginal value function on x_j . Note that the experts may be from various disciplines and have different beliefs and base judgements on different information. Not also that we did not require a one to one mapping of experts to criteria in our applications in the case studies and had not only many-to-one but also more complicated mappings.

The procedure begins with assigning members of the group to G_M^k with the responsibility of assessing g_j . Two sets of judgements are made; first to determine each g_j and then to determine H (the aggregation operator). We will limit ourselves again to the additive form:

$$GV = \sum_{j=1}^n k_j g_j, \quad (4.2)$$

where the assessment of the parameters k_j is accomplished by the group as a whole. The k_j are assessments of trade-offs between the criteria g_j (equivalent to the w_j in the multicriteria model of equation (2.2)) and are not (directly) comparisons of personal value. Thus the main difference between this approach and an individual multicriteria model is in the, implicit or explicit, group decision rule d that transforms individual judgements into group judgements. Krzysztofowicz feels that it does not matter what d is, nor do we need to know what it is, it only matters that the group conforms to it.

The formalisation by Krzysztofowicz (1979) is appealing in many ways, in that it bypasses explicit interpersonal comparisons, is pragmatic in following conventional goal oriented group decision-processes and consequently Krzysztofowicz reported that it produced high levels of consensus and task integration, and participants reacted favourably to the approach.

Informal models

There are various other approaches to group decision-making which are designed to assist in the process rather than with any aspirations to social choice theoretical vigour; these include the Nominal Group Technique and the Delphi Technique. The Nominal Group Technique is designed for small groups. Initially ideas are generated (perhaps in response to a directed question) individually in writing (e.g. around a table). These are then verbalised and discussed in series. The participants vote for, rate or rank the importance of the items. The votes or ranks are tallied, communicated and this process may be iterated. The Delphi Technique works for small or large groups and the group does not necessarily need to meet face to face. An initial question is drawn up and sent to participants, who vote on, rate or rank the issues (as for the first step of the Nominal Group Technique). Results are analysed, and the next question developed with clarification for the respondents/participants. Another iteration produces a preliminary vote that is reviewed and a final vote is obtained (Goicoechea *et al.* 1982).

Another approach mentioned by Goicoechea *et al.* (1982) which is interesting from the point of view of the case studies and the proposed approach in Chapter 10, is the "Iterative Open Planning Process" developed by Ortolano (1974). This has two main ideas: that the public is involved in all stages of the process (by whatever means are appropriate; public meetings, focus groups, representatives in smaller meetings, questionnaires), and that all stages of planning occur more-or-less simultaneously (mirroring the way people think about problems). We feel that the former idea has potential for incorporation into a structured environmental decision-making process, but that, although the latter might reflect people's cognitive processes, it is likely to be rather unwieldy in practice, although probably worth further empirical testing.

Group decision-making and conflict resolution

Clearly, where group decision-making fatally breaks down there may be a need for formal multicriteria conflict resolution approaches. Although a proper discussion of conflict resolution is outside the focus of this thesis, three approaches are mentioned below, which were applied in the context of outranking, goal programming and AHP-like pairwise comparisons. These give the flavour of approaches considered.

Van Huylenbroek (1995) combined the preference function approach of ELECTRE and PROMETHEE (both outranking approaches) with the conflict analysis test of ORESTE in a model called the conflict analysis model (CAM). The CAM model which can be applied to both ordinal and cardinal data, combined the notions of indifference, incomparability, weak and strong preference (from ELECTRE), different types of preference functions (from PROMETHEE), and the conflict test (from ORESTE). Alternatives are given evaluation scores $e_j(a)$ based on each criterion j . The decision-maker has to identify a preference function (i.e. value function) from a choice of six which most closely reflects how these scores are reflected on a $[0,1]$ interval. The difference in evaluation scores on each criterion are translated to preference scores $\hat{e}_j(a,b)$ by the preference function identified: i.e. $\hat{e}_j(a,b) = f\{e_j(a) - e_j(b)\}$. Then overall dominance, the preference indicator $P(a,b)$, is found from a weighted sum of the preference scores. The preference indicator $P(a,b)$ is then used to identify degree of conflict in the third, conflict analysis stage of CAM. Threshold values are identified which allow the specification of strong and weak preference, indifference or incomparability. For example if $|P(a,b) - P(b,a)| > \beta$, this indicates preference. Two more thresholds then determine whether this is strong or weak preference. Although referred to as a conflict analysis tool, this aspect is not expanded, and in fact the result is simply an indication of strength of preference, or dominance and so the extra effort does not seem warranted.

An approach which is perhaps more straightforward and useful is that of Lewis and Butler (1993) who proposed a three stage interactive framework for multi-objective, multiple decision-maker situations. The procedure combines the SIMOLP (Simplified Interactive Multiple Objective Linear Programming) and/or Tchebycheff MOLP optimisation methods with a preference ranking tool and a 'minimise regret' consensus ranking heuristic. The first stage makes use of either or both of the MOLP optimisation procedures. In the second stage, the non-dominated alternatives from the first stage are ranked by the individuals. Ordinal preference rankings will not normally allow the interpersonal comparisons of individual preference needed to find group consensus or preference. Two processes allow them to bypass this problem: the format of the decision process itself, and the use of Cook and Kress's (1985) preferences scale. Firstly, the format suggested is a combination of nominal and interacting groups, where the talk-analyse-talk iteration, may encourage those present to make internal interpersonal comparisons of utility, or at least to fix the 'end-points' of their internal scales (this might correspond to Krzysztofowicz's function d). Secondly, in using Cook and Kress's (1985) preference scale, the decision-maker orders P alternatives into q positions (where $P < q$). The relative positions of the alternatives indicates intensity of preference and the ordinal rankings are converted to cardinal scale on a $[0,1]$ interval. During the third stage, the minimum regret heuristic of Beck and Lin (1983) is used to calculate the consensus ranking. If an alternative i is preferred to an alternative j , a decision-maker will experience some regret if alternative j is ranked above alternative i , but agreement if alternative i is placed above j .

Alternatives are placed in the consensus ranking based on the maximum difference in the total decision-maker agreement and total decision-maker regret. If the decision-makers agree to the preferred solution the process stops, else the three stages are reiterated. Thus, they combine aspects of the group value-aggregation, group value-sharing and informal models as well as a minimise regret heuristic. The 83 students who tested the process in 20 groups found that the nominal interacting groups were very important to the process, and that the choices were fair reflections of the aggregation of individual preferences.

Lootsma (1989) examined the pair-wise comparison of concessions for conflict resolution using linguistic statements of degree of acceptance of a particular 'deal' (very strong liking, strong liking, weak liking, indifference, weak aversion, strong aversion, very strong aversion). These are converted to numerical values on a geometric scale, although the subsequent comparisons by a mediator may be scale independent. The concessions are thus presented as perceived by the parties concerned rather than, for example their monetary value. The advantage of this approach over MCDA alone is that, for situations of real conflict, a 'rate of exchange' of concessions is established which could be useful to a mediator.

Losa *et al.* (2001) discussed the use of multivariate ordination techniques combined with SMART in conflict resolution and negotiation. Examination of the two-dimensional representation of SMART scores for alternatives, and their relative positions to various stakeholder groups (on the same graphic) allows areas of possible compromise or consensus to be identified. This might either encourage the creation of new alternatives which better satisfy the groups or lead to the formation of stakeholder 'coalitions' to outweigh alternatives which are preferred by other groups.

A characteristic of decisions regarding projects with environmental impacts is that there are likely to be fairly high levels of conflict, often simplistically perceived to be development versus conservation conflicts. Many environmental decision-making articles therefore discuss consensus seeking and conflict resolution. In comparison to the methods proposed in the MCDA literature mentioned above practical applications seem to favour less complex approaches as illustrated in the four applications to wildlife management conflicts discussed below.

Maguire and Boiney (1994) used expected utility within a conflict resolution framework in their hypothetical, but realistic example of northern white rhinoceros conservation in the Democratic Republic of the Congo (DRC) (where at the time of writing only about 28 individuals remained in the wild). They consider alternatives such as CBA, 'initial decision analysis' and adaptive environmental management, but reject these in favour of MAUT. This framework allowed for the updating of utilities and probabilities (using a Bayesian approach) with new information which facilitated compromise and encouraged the creation of new alternatives. The separation of the utility and probability models avoided the confounding of judgements about facts and judgements about values. The two extreme positions held by Western zoos and DRC wildlife officials respectively were: remove all remaining animals to captivity to ensure safety from poaching or, allow no more removals but concentrate resources on more vigilant anti-poaching activities. Expected utilities and probabilities of outcomes of the different approaches were elicited

from both groups and compared. A distributive bargaining alternative (phased removal to captivity) was not favoured by either group, even with sensitivity analyses. However, a contingency based alternative whereby particular actions taken contingent on observation of a particular event, provided a compromise agreeable to both: removal of the wild population if it dropped below a critical threshold.

Ralls and Starfield (1995) discussed the use of two decision analysis approaches to choose between eight possible management strategies applied to the problem of mobbing deaths in the endangered Hawaiian monk seal about which different interest groups have conflicting objectives. Both decision approaches required estimates of various probabilities (such as the probability of reaching a population of more than 400 after 20 years) derived from simulation models. The first approach, borrows from goal programming the idea of satisfying each goal sequentially (i.e. satisfy the most 'important' goal first, and then the next etc.). In this approach the goals do not have to be independent, and in fact their lack of independence could be advantageous as the goals of one group can be met part-way in the higher level goals and then re-addressed later in lower level goals once the needs of others have been addressed. The other approach was a variation of SMART and the probabilities of various outcomes are the attributes by which the alternative management strategies are measured (the square of the probabilities could have been used instead for a more risk averse view). The results of the SMART-like approach agreed well with those of the lexicographic approach and both approaches were found to (a) provide an explicit, well documented, reproducible decision process, (b) be easy to use and understand, (c) promote and focus discussion on objectives and priorities, (d) facilitate a structured examination of multiple objectives and trade-offs, and (e) be flexible, non-prescriptive in the terms of actual workshop process and therefore suitable for workshops, (f) were robust.

McDaniels (1995) used multiple objective decision analysis (in fact MAUT in our terminology) to conduct an *ex post* analysis of a specific fisheries management decision involving conflicting long-term objectives for stocks of sockeye salmon. In the salmon fishery, delaying the start of the fish season by a day, based on incoming information relating to the peak of the run, can mean an increase in potential benefits of millions of dollars. The broad level objectives were: long term stock health, short and long-term economic benefits, social acceptability and opportunities for learning. Surrogates were used (stock size in the year 2010) for some attributes while for others dollar values and constructed scales were used. The attributes were found to be (for the range of values in the present case) preference independent, utility independent and additively independent. Trade-offs between each pair of attributes were determined directly in order to determine scaling constants. Objectives other than those conventionally used in fisheries modelling were found to be useful for in-season management.

'Multiple accounts', GIS and production models were used by Brown *et al.* (1994) in order to provide decision support to resolve forestry and wildlife conservation conflicts between commercial forestry operations and the conservation of the caribou herd in the Mount Revelstoke and Glacier National Parks, Canada. The area of conflict was divided into forestry cutblock areas and variables relating to harvesting (log value for different species, roadbuilding, maintenance and hauling costs) were determined. Costs differed depending on slope and altitude etc. and were modelled using standard forestry models to obtain net present value (NPV) for each cutblock. The cutblock

areas were categorised into areas of high, medium and low habitat suitability for caribou, based on elevation, slope, slope position, dominant species, height class and crown closure class. Five different scenarios (type of harvesting regime and extent of protected area being some of the discriminating features of these scenarios) were simulated over the next 120 year harvesting cycle. The efficient frontier between NPV and caribou habitat (being termed 'multiple accounts' in this study) determined that the establishment of a park and the preservation of old-growth scenarios were inefficient. The areas of highest potential conflict (high caribou value and high NPV of timber) could be identified by the GIS overlay system, and thus attention could be concentrated on these areas. Cocks and Ive (1996) used similar 'conflict indicator maps' in another forest land allocation problem.

4.1.4 Extensions to uncertainty and imprecise judgements

Projects and policies with environmental impacts are characterised by the necessity of making decisions with low levels of information and high levels of uncertainty. Decision-makers may wish to minimise the potential for loss of future opportunities or to 'avoid regret', which in turn implicitly means adopting the 'precautionary principle' often cited in environmental decision-making (Vatn and Bromley 1994). The types of imprecision, risk and uncertainty may be listed as:

1. external uncertainty: uncertainty in long term environmental, political, social and economic conditions,
2. internal uncertainty: uncertainty in information about the likelihood of impacts and their extent (e.g. through lack of data),
3. imprecision: uncertainty in preferences and values (e.g. of different levels of impact), scores, and weights, and
4. the dynamic relationships among these.

Varis *et al.* (1994) look at how uncertainty and subjectivity in model structure, in objectives and in information can be structured and quantified. They divide decision-making into three components: (a) objectives, goals, targets, criteria and constraints (the objective functions), (b) elements in the system that are at least partially under control (decision variables), and (c) the information conditioning the above components (chance variables). External uncertainty and some aspects of internal uncertainty may be modelled with non-numerical techniques (rule-based systems, production systems) or numerically (e.g. Bayesian approaches, classical probabilistic approaches, fuzzy logic). Internal uncertainty may also be dealt with through utility theory and analyses of attitudes to risk. There are various other approaches for assessing value within MCDA where there are different levels of uncertainty, precision and information. Methods that use linguistic scales or only require ordinal inputs are less demanding on the respondent, while fuzzy approaches directly deal with our inability to be precise about uncertain outcomes. When faced with a lack of detailed, quantitative information, decision-makers may also rely on expert opinion, heuristics, set environmental standards, or use environmental indices.

Uncertainty

MAUT, which addresses uncertainty through the analysis of expected values is the only MCDA approach that formally includes uncertainty, although others may use expectation and /or variance as criteria (Stewart 1995), or more informally include 'risk' in criteria such as 'the risk of exotic vegetation invasion' (see Chapter 6).

Faucheux and Froger (1995) distinguish two types of uncertainty: risk (weak uncertainty) and strong uncertainty. They consider the use of Bayesian, MAUT or MAVT approaches inadequate for the 'strong uncertainty' characteristic of projects with environmental consequences. They feel that stochastic environmental decision-making approaches which rely on Bayesian theory are based on a substantive rationality hypothesis (the rationality of a decision is considered independently of the way in which it is made - rationality refers to the results of the choice), and that they require the listing of all consequences and their probabilities. They regard the use of expected utility analysis as similarly unsatisfactory because the range and distribution of future environmental effects are not known and / or unknowable. They further state that we do not act 'as if' these are known or knowable, but do not give any basis for this statement.

That MAUT requires knowledge of the full multivariate distribution of outcomes is problematic. Stewart (1995) used Monte Carlo simulation to analyse the effect of reducing this requirement in discrete multi-criteria problems with uncertainty. Although decision-makers violate many of the assumptions of rational decision-making, the complexities introduced into methods in order to account for this do not necessarily change the decisions which may result. The analysis showed, among other things, that uncertainties may in effect be represented by 3 to 5 scenarios: this result may open the way to the development of a solution to the issues raised by Faucheux and Froger (1995). The problem becomes more manageable and sensitivity analyses can help to assess the range of possible values for alternatives under different assumptions.

Interestingly, the approach proposed by Faucheux and Froger (1995), while not particularly addressing uncertainty, combines an MCDA approach similar to SMART with the ecological economics (see Section 3.4) innovation of using energy as a common numeraire rather than money. They feel that their approach has procedural rationality (the rationality of a decision is in terms of the way on which it was made - rationality refers to the decision-making process itself) in addition to the substantive rationality mentioned above. Their approach entailed: (a) The creation of a hierarchical tree of goals and sub-goals, where the goals may be unmeasurable but sub-goals are measurable (i.e. a value tree in our terminology), (b) the use of the satisficing principle (as opposed to optimising) to choose between options taking into account social, economic and ecological considerations.

Additional issues arise with regard to the use of MAUT that may be particularly relevant in developing countries. First, expected utility analysis does not take account of the notion of regret (Loomes and Sugden, 1982), which may be an important consideration in situations of irreversibility of environmental impacts. Second, focusing on hypothetical lotteries rather than the actual problem and alternatives, may be counter-productive and inefficient (Tochner 1977), and eliciting risk values more subject to framing biases than other values (Fischhoff *et al.* 1981). Third, MAUT may be inappropriate as a workshop tool where participants have widely different disciplinary backgrounds and levels of numeracy as might often be the case in developing world settings such as the cases described in this thesis. Last, in third world economies where much activity is outside the formal economy and involves close and reciprocal links with the environment, assessing the impacts of a project are not straightforward.

This is particularly the case as the state of “endemic unpredictability” which Rosenhead (1989, p. 195) proposes exists in the first world must surely be more evident in developing countries, and this is exacerbated by high levels of intra- and inter-annual climatic variation as seen in much of Africa. Decisions may be made on the basis of present or mean conditions without due regard for effects of high variance.

However, decision trees or influence diagrams (Gregory *et al.* 1992) may be useful in the construction of links between present decisions and possible future events as visual and cognitive aids, and in the determination of expected values through probabilities where appropriate. Additionally, combining several different approaches to the exploration of options may provide a reasonable basis for making decisions which ‘avoid regret’. Some of the ‘soft’ operations research techniques such as strategic choice and ‘robustness analysis’ (Rosenhead 1989) or the ideas of adaptive management suggested by Walters (1986) also offer some ways of managing uncertainty. Bayesian concepts of prior and posterior probabilities or other information might also be used, either on their own or in combination with other methods (including adaptive management). In robustness analysis, a set of decision ‘packages’ are identified and the pathways resulting from these and sequential decisions explored. Decision packages which produce acceptable or desirable results in a number of different scenarios, or which are less likely to result in undesirable outcomes, are preferred to others as they increase intertemporal flexibility of decisions.

Imprecision: Fuzzy approaches

Imprecision in preferences (which may be due to lack of information) is the main emphasis in fuzzy approaches, which have been combined in various ways with AHP, outranking, SMART and MODM methods. Fuzzy information can be included in decision models in three ways: by using linguistic variables, by using fuzzy numbers (Munda *et al.* 1994b) and fuzzy set operations. For example, degree of membership in the set ‘acceptability’ A , might be in one of three fuzzy categories A_1 (acceptable), A_2 (moderately acceptable), A_3 (unacceptable). A membership function $\mu_{A_i}(x_j)$ associates the level of an attribute x_j to these categories. A fuzzy set operation, e.g. $\min(\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_n}(x_n))$ might aggregate assessments for different attributes to give an overall assessment (Cornelissen *et al.* 2001).

In their example of a fuzzy approach Munda *et al.* (1994b), considered a function $u_A(x)$ that specifies the grade of membership of x in the set A . In their practical example, criteria levels were found for each of three alternatives, on natural scales (e.g. money) or qualitative scales (excellent, good, moderate). Then each of these was given a membership degree - on the interval $[0,1]$ - that indicated membership of “each action to the interval of feasible and acceptable values defined on each criterion”. In practice this was not dissimilar to rating alternatives on a $[0,1]$ preference interval scale for each criterion. They then continued by specifying the ‘degree of truth’ of three linguistic comparisons of each pair of alternatives: a_1 is better than a_2 , a_1 and a_2 are indifferent, a_1 is worse than a_2 etc. The process continued to obtain a weak order ranking. Smith (1994) also discussed in some detail the theory of fuzzy sets and numbers and used two examples: the siting of a hazardous waste disposal site and the building of a transport route, but do not illustrate the use of linguistic scales. Their explanation and approach is similar to Munda *et al.* (1994b). A slightly different approach was followed by Xiang *et al.* (1992) who looked at the application of fuzzy-

sets to a land-use planning problem. A continuous real interval $[0,1]$ was used as the preference measurement system, with descriptive labels for “control points” on this interval i.e. 0=absolutely unsuitable, 0.3 = unsuitable, 0.5 = moderate, 0.7 = suitable, 1.0 =perfect. A similar scale was used to derive the weights of the attributes: 0=not important, 0.3=less important, 0.5=moderate, 0.7=important, 1.0=very important. They have third similar scale that measures the “tolerance” of the respondent on each attribute. This may relate to what Munda *et al.* (1994b) referred to as a measure of “incertitude” but is not further explained.

Imprecision: Linguistic and semantic scales

The different scales of measurement (Chapter 2) may be associated with the levels of information available as input into the problem assessment. More imprecise preference statements may be associated with rank ordering of alternatives or scenarios, while greater levels of precision may be associated with interval and ratio statements of preference. There have been various initiatives over the last few years to formulate decision-aiding procedures which make use of linguistic or semantic statements of degree of preference which are then translated to a numerical scale. A forerunner of these approaches was the AHP. In response to criticisms of the original AHP, Lootsma looked at scale sensitivity in AHP and proposed some adjustments which are included in the REMBRANDT system (see Lootsma (1993) and Olson *et al.* (1995) for details). In REMBRANDT, the statements of preference remained essentially the same, but were subsequently treated differently to the AHP method. They were converted to a numerical value on a geometric scale and then logarithmic regression was used to calculate the ‘single criterion impact scores’ $v_j(A_j)$, $j=1,\dots,n$. Lootsma (1989) also used verbal judgement in a conflict resolution setting to assess degree of acceptance of a particular ‘deal’.

Bana e Costa and Vansnick (1994) and Bana e Costa *et al.* (1995) pursued a different approach in MACBETH (Measuring Attractiveness by a Categorically Based Evaluation Technique), where cardinal criteria functions are constructed from semantic category judgements. Pairs of alternatives (a,b) are compared and respondents are asked to state the difference in attractiveness between them. Six categories of difference of attractiveness $(C_1,\dots,C_6)$ are defined, the categories being intervals on a real line (very weak difference in attractiveness (C_1) , weak (C_2) , moderate (C_3) , strong (C_4) , very strong (C_5) , extreme (C_6)). A key difference between this idea and that in AHP is that the judgements concern differences of attractiveness rather than ratios of priorities or importance. Also, in the original AHP, a fixed number was associated *a priori* with each of the categories of ratios. In MACBETH, the categories were intervals on a real line, and the thresholds demarcating these categories $(s_i, i=0,1,\dots,5)$ were determined simultaneous to finding the numerical value function (i.e. not *a priori*, although s_0 is fixed = 0). Initially a matrix is completed which states the category C_i associated with the difference in attractiveness between the two alternatives being considered from a particular point of view. Consistency can be checked by ensuring that values increase or decrease consistently along the rows and columns. Another check (of so-called ‘semantic coherence’) is to ensure that if $aPb = C_i$, and $bPc = C_k$, then $aPc = C_k$ or C_{k+1} , so that small differences in preferences do not cause large jumps in categories. The value functions and the s_i were then determined by linear programming. Cook and Kress (1991) also convert ordinal preferences to interval scales, and Larichev and Moshkovich (1995), use ordinal information together with several other stages (e.g. conjoint analysis).

An approach which appears to combine the advantages of utility and fuzzy approaches in terms of expressing imprecision in preferences is that of Salo and Hämäläinen (1995). Originating in the pair-wise comparison of AHP, but combining this with a MACBETH like-approach, the authors use interval (i.e. a range rather than a point) statements of preference and the intervals are used as linear constraints on the weights. The feasible region derived from the preference intervals means that every alternative receives an interval of weights: alternative *a* shows absolute dominance over *b* if for all feasible local weights *a* has the higher weight. The dominance relations are determined through linear programming. The adaptation originated with Saaty and Vargas (1987) because of the acknowledged difficulties of eliciting exact ratio estimates. The use of interval statements may have advantages in group decision-aid as the interval can include the range of opinions of the group. These ideas have been incorporated in the software program INPRE which is AHP based and COMPAIRS which allows the analysis of value functions (as opposed to pairwise comparisons) in a similar way (Salo and Hämäläinen 1992).

Imprecision: Indices and indicators

Indices and indicators are comparative management tools that indicate performance on a pre-defined scale. Beinart (1995) classifies (after Morgan and Henrion 1990) these approaches as rights-based approaches (single- and multi-attribute environmental standards) and utility-based approaches (environmental indices). The use of the concept of utility is seldom if ever explicit. Indices used within environmental management contexts are often 'proxy' attributes in that they measure an aspect of the environment which is linked to the actual measure of interest. The functional relationship between the proxy and the attribute of interest is seldom explicit. It is also necessary to keep in mind the functional importance of different sub-indices when combining these. For example, the contribution of different gases to an index of global warming may be weighted by their relative radiation inhibition effects (Rogers 1999). In developing indices of biological importance or significance, some felt that aggregation to one index of 'ecological value' was desirable (or necessary), while others felt that this confused rather than clarified the issue (Smith and Theberge 1987). Supporters of the latter view (e.g. Turpie 1995) prefer separate indices or ranks for separate criteria as an alternative with exceptional value for one criterion may be hidden in the aggregation. Some felt that unless the ecological relationships between criteria were reflected by the mathematical specification of criteria and aggregation the final score would be meaningless (Smith and Theberge 1987). Either way such a composite index does require a strong theoretical basis as yet not established for sustainability or conservation evaluation. Increasing acknowledgement of the need to include social and economic aspects in conservation decisions means that MCDA approaches are beginning to be used in the development of indices (e.g. Lamberth and Joubert, in review, Turpie *et al.* 2002).

4.2 Applications in environmental management

Environmental management applications in which the range of MCDA methods have been applied include environmental impact assessments (EIAs), environmental conflict resolution (some examples were given in Section 4.1.3) and general environmental decision-making. MCDA (and CBA) in a sense originated with applications to water resources management in the USA and a variety of approaches have since been applied in many parts of the

world. As the applications have often been of an operational nature (e.g. release rules for dams), MODM methods are frequently used, while the use of AHP, SMART and outranking approaches are often used in planning. MCDA, often MAUT, has also been widely applied in the energy industry and in siting problem (e.g. nuclear waste sites).

The more general acceptance of MCDA and applications in environmental impact assessments, nature conservation and prioritisation has been more limited or *ad hoc* as the theoretical basis of MCDA techniques appears to have been slow to filter through to these fields. Bakus *et al.* (1982) suggest that this may be due in part to the jargon used and the tendency for articles to appear in specialist journals and to concentrate on issues of academic rather than practical interest. A multi-criteria approach is a necessity in many other applications such as in the prioritisation of types of habitats or ecosystems for conservation (e.g. wetlands, rivers, grassland areas) and environmental impact assessments. However, MCDA is often not explicitly used, more often scoring methods are built up from first principles, and in practice, sometimes methods are applied which are not theoretically justified. For example, a simple sum or weighted sum of ranks is often used. This implies that an increase in one rank based on criterion i is equivalent to an increase of one rank based on criterion j , and that an increase from rank k to rank $k+1$ is equivalent to an increase from rank l to $l+1$ (within and across criteria). However, this assumption is often not recognised.

This section briefly discusses the use of MCDA in EIAs and then describes some applications in the field of land- and water-resource management. The examples were included because of (a) the similarity of the area of application to that of the main case studies described in Part II, (b) the range of methods which were used and (c) they were reported in the ecological or environmental rather than operations research or multicriteria literature.

Environmental Impact Assessments

EIA emerged as a result of the pressures from a public desiring more accountability and participation in decision-making and a long history of criticism and misapplication of CBA. It originated with the passing of the National Environmental Policy Act (NEPA) in the USA in 1969. Smith (1992) discussed impact assessments in the context of sustainable resource management, and categorised the various types of EIA as checklists, interaction matrices, overlay mapping, networks, and simulation modelling. Checklists are the most frequently used and were listed in increasing order of sophistication by Smith (1992), as simple, descriptive, scaling, and multi-attribute utility theory checklists.

The Environmental Evaluation System (Dee *et al.* 1973) and the Sondheim method (Sondheim 1978) are two of the more well known scaled checklists (both are in effect MAUT/MAVT). The Sondheim method, which includes the opportunity for public involvement in the aggregation of scores, improved upon the former in that the weights attached to the criteria are project specific rather than pre-determined. Because of its complexity, MAUT is seldom applied in EIAs (e.g. Fuggle and Rabie 1992 p. 771, Smith 1992 p. 19) and “lacks the ease of understanding that is a pragmatic necessity” in impact assessment (Smith 1992 p. 20). The perception of complexity arises from exposure to or use of more ‘purist’ MAUT approaches. Methods such as SMART and software such as VISA (1995) are only recently becoming known within the EIA fraternity (e.g. Marttunen and Hämäläinen 1995, Joubert 1997), and may

go some way to softening perceptions regarding the use of MCDA in EIAs. As with MCDA itself, there has perhaps been “an excessive interest in methodologies and techniques” in EIAs (Clark and Herington 1988 p. 12) which may be counterproductive in terms of the rigorous application of any particular method.

EIAs in South Africa generally try to include various types of uncertainty through formally requiring specialists to state their ‘confidence’ and suggesting that the probability of the impacts should be mentioned (Dept. of Environmental Affairs 1992). Confidence in the estimation of the significance of the impact is often reported on a 3 or 5 point verbal scale and probability on a 4 or 5 point verbal scale (e.g. Gibb 1997). However, this information is not formally included in any decision-making process, or any final impact statement. In South Africa impacts are often reported in terms of their physical extent, duration, intensity, probability, and confidence in the estimates (incidentally this is remarkably consistent with Bentham’s dimensions of utility, which were: intensity, duration, certainty, propinquity/remoteness, together with fecundity, purity and extent: see Section 3.1). The impact reported combines these dimensions in an unspecified way to indicate the significance of the impact on a 4 or 5 point verbal scale (e.g. Gibb 1997).

Land- and water-use applications

Chang *et al.* (1994) used the MODM methods of multi-objective linear programming (MOLP) and compromise programming in a watershed-based land-use decision-making problem, to examine trade-offs among objectives for several planning scenarios. The objectives were grouped under the three broad headings of economic (regional and national) development, social welfare and environmental protection. The six objectives were to: minimise phosphorous and nitrogen pollution and discharge of biological oxygen demanding load, minimise the total sediment yield, maximise employment level and maximise income. Decision variables were the amounts of forest land, agricultural land, residential area, grassland, stock farming area and recreational area. The constraints were measured in terms of, for example, the minimum amount of forest land required for conservation by law. They used compromise programming (which determines the minimum distance from the ideal solution) to examine non-inferior solutions and trade-offs after rescaling the objectives to a [0,1] interval, and compared the results to those obtained by using a multi-objective simplex method. The solutions did not vary much with changes in the parameters of the compromise program nor with the use of the multi-objective simplex method.

Stansbury *et al.* (1991) evaluated ten alternative water transfer options with a three module decision support system consisting of 1) surface- and ground-water models, 2) an impact analysis which made use of GIS, and 3) an MCDA module. The MODM method used was composite programming, an extension of compromise programming. The criteria (“base-indicators”) were organised into a value tree and weights, “balancing factors” and worst and best values for each criteria were determined. The balancing factors “indicate the importance of the maximal deviations of the indicators and limit the ability of one indicator to substitute for another” (Stansbury *et al.* 1991 p.446). This means that “high balancing factors give more importance to large negative impacts on any indicator rather than allowing these impacts to be obscured by the trade-off process” (p. 446). The criteria were quantitatively defined from models and GIS, transformed to reflect non-linearities, normalised to the [0,1] interval and then aggregated

(composited) into first, second and final level indicators. The final level consisted of a composite indicator for ecology and one for socio-economics. Nine weight sets were defined for the different groups' perspectives (ecologists, water managers etc.) and the alternatives ranked on the basis of each of these. Apart from lack of clarity about how the weights and 'balancing factors' were determined, this approach appeared easy to use, problem focused and, importantly, easy for the reader to understand. In many respects it was similar to MAVT, with the exception of the use of 'balancing factors'.

CBA and weighted summation of scores were used to assess the effects of land-use conversions in an Indonesian watershed development project which was an attempt to increase forest and agricultural productivity while reducing sedimentation and stream-flow fluctuations (de Graaff and Kuyvenhoven 1996). The actual developments were compared to three hypothetical options (a conservation oriented option, an agricultural production oriented option, and an autonomous option where there is no government interference). The evaluation criteria were grouped under the headings efficiency (maximising net benefits of food, and of export produce or import substitutes, minimising investment and recurring costs), equity (share of income to people in watershed as compared to those downstream), and sustainability (minimise soil erosion and natural forest loss, maintain hydrological regime). Three stakeholder groups (central government, local agencies, upland farmers) gave rise to three weight sets on a scale of 1 to 8, from least to most important. These were translated into weights on a [0,1] interval by dividing the rank by the sum of the ranks. The CBA reflected the efficiency of the four options, but could not take into account biodiversity features of forests and equity attributes whereas using the weight sets of the different stakeholder groups to the original criteria allowed all aspects of the problem to be considered.

Prato (2000) used MCDA, specifically goal programming, for the selection of land-use and management plans at the site and landscape scales. He compares two cases, one in which the land in a watershed is mainly publicly owned and one in which the land is mainly privately owned, the criteria of interest differing between the two.

Faith and Walker (1996) used the concepts of utility and 'trade-off space' in selecting protected areas, but they limited themselves to two criteria: total biodiversity and total cost of a set of areas. A constraint was added in order to incorporate the concept of complementarity, so that the biodiversity contribution of each new individual area added should exceed its individual cost. The cost is initially expressed in terms of suitability for another land-use, and then translated into the 'equivalent cost' in terms of biodiversity.

Davies *et al.* (1995) in an application to dune management developed a checklist of various parameters and a scoring system applicable to Northern Europe. The scores (on a 0 - 4 scale) were either derived from non-linear relationships between various quantitative factors (e.g. pebble cover as a percentage of surface) or from partial ordinal rankings of qualitative data (e.g. high, medium, low sand supply input) which were related to a score between 0 and 4. These scores were combined to give a site vulnerability index, which was compared to an index of extant protection measures. The ratio of the two indices gave an indication of whether the dune area required more

protection, and, by assessing the detailed criteria that were used to derive the indices, the particular type of protection needed was determined.

Ridgley (1995) proposes an approach to setting instream flows that begins with designing flows using an MODM approach, after which a suitable subset can be evaluated using AHP or SMART. Ridgley and Rijsberman (1992 and 1994) used AHP to evaluate restoration policies for the Rhine estuary. The two examples in these three papers are similar in context to the cases described in Chapters 6 and 7, as they are all concerned in different ways with selecting an 'appropriate' river quality.

4.3 Summary

This chapter has provided a brief overview of MCDA history and the many approaches encompassed by this term. The main fields of MCDA (MODM, MADM) have developed and expanded since von Neumann and Morgenstern's development of expected utility and the early applications of linear programming, and illustrate the shift from an emphasis on positive economics and descriptive behavioural decision theory to prescriptive decision-aid in MCDA. The range of MCDA approaches also indicates the broad nature of the intellectual input into MCDA (e.g. psychology, operations research, economics and statistics).

Extensions to group decision-making, methods which include uncertainty, allow for imprecision and which used ordinal or semantic scales were described. Some applications in land- and water-use decision-making (the context of most of the case studies) were described which applied both different and similar approaches to the MCDA case studies.

This chapter provides the background for the description, in Chapter 5, of the particular method (SMART) which was applied in the case studies and the methodological context for the MCDA applications in Part II. Together with the outline of measurement theory in Chapter 2, this chapter also provides some of the background for the discussions in Part III.

5. Contingent valuation, conjoint analysis, travel cost and SMART

"The notion of consumers' surplus always seems on a superficial view to involve a bit of flimflam", WJ Baumol 1977b in Don Ross 1999, p. 260.

While previous chapters have addressed the general background and context of MCDA and EE, this chapter discusses in more detail the particular MCDA evaluation method and EE valuation methods that were applied in the case studies to be described in Part II. The overall aim of using any of these methods is to estimate the value of a project, policy or alternative and include that value in a decision-making process. Normally, the EE valuation methods would be used in estimating particular values within a CBA aggregating framework, while MCDA provides the evaluation tool as well as the process and aggregation framework. These decision-making processes are aimed at finding a 'best' alternative or at maximising social welfare in some sense. Therefore, a general description is given (Section 5.1) which links the MCDA and EE approaches in terms of their common aim of measuring multi-attribute value. The specifics of the methods used within the two paradigms then follow.

The EE methods applied (Section 5.2) were the contingent valuation method (CVM), contingent behaviour valuation (CBV), conjoint analysis (CA) and the travel cost method (TCM). Note that, in the context of the applications in this thesis, conjoint analysis is included as one of the EE valuation methods, although it is also an MCDA technique. These are all survey based methods, and the known survey biases are mentioned at the end of this section. The MCDA method applied in the cases was named SMARTx and combined SMART with the group-value sharing model and a process for the development of alternatives (Section 5.3). Section 5.3 ends with a discussion of various studies comparing MCDA methods, providing some justification for the use of SMARTx as the MCDA method applied in the case studies (where only one MCDA method could be applied).

The aim of this chapter is therefore to describe the approaches in detail, within the same overall descriptive framework in order to facilitate the descriptions of case studies and the subsequent discussion and comparison. The comparison in Chapter 9 considers a general social choice framework, and within this, assesses the two paradigms and their techniques in terms of four metacriteria (which are introduced more formally there). The first metacriterion concerns the validity of the paradigm or particular technique within the prevailing political milieu and particular decision context. This chapter (e.g. Section 5.1) informs that discussion together with the context of the case studies in Chapters 6 to 8). The second metacriterion considers other issues of validity that do not arise from the politico-decision context. These discussions are informed by the section on survey biases (Section 5.2.4), on the TCM (Section 5.2.5), the comparison of SMART with other MCDA methods (Section 5.3.4) and the case studies. The third metacriterion refers to the ability of the SMARTx and EE approaches to adequately include efficiency, equity and sustainability criteria and the fourth metacriterion refers to the practicality of the methods (e.g. simplicity,

transparency and flexibility). Besides the discussion in Section 5.3.4, there are details throughout this chapter that inform discussions on these two metacriteria.

5.1 Measuring welfare and multi-attribute value

SMART and EE valuation methods (if within a CBA framework) both aim to estimate the overall value (utility or social welfare) of an alternative. The EE paradigm is concerned with finding welfare measures for changes in Q , the environmental good, in our context, for inclusion in the cost-benefit equation:

$$V(a) = \sum_{j=1}^n B_j(a) - \sum_{i=1}^m C_i(a), \quad (5.1)$$

where B_j are the benefits and C_i are the costs of alternative a , which are usually discounted over time, and may be weighted. We will suppress the argument of $V(a)$ and related arguments in much of what follows. Within a CBA framework, a particular B_j may be provided by an environmental good or service at a particular ‘quality’. Depending on the type of good or service, a particular EE valuation method is applied in order to estimate the monetary value of B_j . The “subjective” EE valuation methods (*sensu* Dixon *et al.* 1986 see this thesis p. 36) are linked to the individual utility U_j^i of B_j (U_{Bj}^i) where the utility of a commodity or alternative is a function of the multiple attributes of the alternative. The EE methods may also explicitly consider the individual’s attributes (such as income, sex, age etc.) and estimate value to the individual i of the benefit B_j :

$$U_{Bj}^i = f(Q, P, K^i), \quad (5.2)$$

where U_{Bj}^i is utility in a general (economics) sense to the individual i of the environmental benefit (good or service) B_j , Q is a vector of the attributes of the environmental good, P is a vector of market prices or costs, and K is a vector of attributes of the person (or group). The EE methods usually assume that f is the same for all i and usually market prices P are assumed to remain constant with a change in Q , and are therefore omitted, thus equation (5.2) becomes:

$$U_{Bj}^i = f(Q, K^i, Y^i), \quad (5.3)$$

where Y , income, is now separated out from K .

The overall social utility or welfare measure of the environmental benefit B_j is then an aggregation of the U_{Bj}^i (see Section 3.2):

$$B_j = f(U_{Bj}^1, U_{Bj}^2, \dots, U_{Bj}^n),$$

thus providing one of the B_j in equation (5.1). Note that often an environmental good is valued alone, outside of the cost-benefit equation.

Usually, in applications of CVM, CBV, TCM and CA it is assumed that the value of an environmental good can be represented by the weighted addition of its attributes, although interactions with income are also sometimes included (and the value of an alternative by weighted addition as in equation (5.1)).

CVM, CBV, TCM, CA: $U^i(a) = \alpha^i + \beta_{Q1}^i Q_1(a) + \beta_{Q2}^i Q_2(a) + \dots + \beta_{YP}^i (Y^i - P(a)) + \dots + \beta_{K1}^i K_1(a) + \beta_{K2}^i K_2(a) + \dots$

where the β_{jk}^i are the weights (or coefficients).

The SMART equivalent to equation (5.1) is:

$$V(a) = \sum_{j=1}^n w_j v_j(x_j(a)), \quad (5.4)$$

where v_j is the ‘score’ or value of the alternative according to criterion j and w_j is the weight for that criterion and V is overall value in an MCDA sense (i.e. excluding risk or uncertainty), but generally comparable to V in equation (5.1). MCDA in general, and SMART in particular, does not explicitly include the attributes of the individual K , but rather, different individuals or groups may have different functions which directly or indirectly reflect their K ’s – in SMART therefore, they may have different v_j or w_j . The v_j might be the value of one environmental attribute, good or service, and the value v_j of a criterion to a group or individual is an implicit or explicit function of one or more attributes X :

$$v_j(a) = f(X(a)),$$

By the basic axioms of utility or measurement theory, individual or group i will choose a rather than b if and only if $U^i(a) \geq U^i(b)$ (from equation (5.3)) or $V(a) \geq V(b)$ (from equations (5.1) or (5.4)).

If CVM and CA are formulated according to the random utility model (RUM) (e.g. McFadden 1976), the probability of choosing a over b is:

$$\Pr(\text{choose } a) = \Pr[U^i(a) \geq U^i(b)] \text{ or } \Pr(a) = \Pr[V^{*i}(a) + \epsilon^i(a) \geq V^{*i}(b) + \epsilon^i(b)], \quad (5.5)$$

where V^* is the ‘observed’ component of utility and ϵ is the ‘stochastic’ component due to measurement, uncertainty or other errors. The form of the probability of choosing a then depends on assumptions around the distribution of ϵ .

Note that the development of the RUM model in economics relies on the compatibility of an individual repeated choice model with a population sample model. The individual choice model is a model of an individual subjected to repeated choices, who implicitly chooses one of a set of utility functions, and then chooses the alternative which maximises utility according to that utility function (the original setting of RUM, Thurstone 1927, McFadden 1976). The population sample model is a model in which the analyst draws a sample from a population who have ‘fixed’ different utilities and are offered a single choice (McFadden’s original setting being repeated revealed choice settings). The two models are regarded as equivalent, the latter allowing for modelling of economic choice (McFadden 1976). SMART does not consider the random component, uncertainty, nor use probabilistic analyses.

In CVM, CBV and EE applications of CA, the aim is to find the amount of money, willingness to pay (WTP), which will make the utility, to an individual, of two scenarios equivalent (compensating variation) when there is a change in quality Q . This is a measure of the consumer’s surplus (compensating surplus, or equivalent variation or surplus may also be relevant) (see Section 3.4). In TCM, Marshallian consumer’s surplus is found directly from the demand function.

All of the EE methods mentioned here and used in the case studies (CVM, CBV, CA, and TCM) are survey based techniques. CA might also be used in a workshop setting (as part of MCDA). SMART may be used in individual decision-making contexts (e.g. the manager of a firm) but in the context of this thesis it is a workshop-based technique using Krzysztofowicz’s (1979) group-value sharing model. EE valuation methods have been divided into revealed and stated preference approaches. As the names suggest, revealed preference approaches rely on choices

that individuals reveal in the market to develop a demand curve and impute value to alternatives. Stated preference approaches either rely on questionnaires which ask respondents about their preferences, or on focus / expert group workshops where participants state their preferences (see Table 3.1). TCM is a revealed preference technique, and CBV, CVM and CA are stated preference techniques. SMART is in effect a non-survey based stated preference approach (as are most other MCDA methods), although not addressed in economics literature as it does not necessarily use or aim to arrive at monetary values for goods.

5.2 Environmental economics valuation methods

The implied decision-making framework for EE is usually CBA (see Section 3.4.1). In a specifically environmental setting, this will mean:

1. identifying the costs and benefits of an alternative including the impacts on, or use and non-use values of, environmental goods and services of the ecosystem or amenity in question,
2. deciding on appropriate valuation methods for the environmental impacts,
3. for non-consumptive use and non-use, designing appropriate CBV, CVM, TCM, and CA survey instruments,
4. designing an appropriate sampling strategy and administering the survey(s),
5. using real and surrogate markets to estimate other values, and
6. discounting and summing all costs and benefits however derived (in certain cases, costs and benefits are differentially weighted, usually not) to give net present value of the potential change.

CBA can therefore give a statement of the desirability of a single project relative to the *status quo* while this is only possible in MCDA if the *status quo* is explicitly included as an alternative (and see Section 10.4). EE valuation is often undertaken outside of any decision-making context or framework and sometimes without a specific statement of what the new state Q might be, in other words to find the value of a particular amenity or environmental quality *per se*.

5.2.1 Stated preference: Contingent Valuation (CVM)

There are a variety of CVM methods both open-ended and closed-ended (or referendum), and many variations of each. Open-ended methods ask respondents how much they would be willing to pay (in cash, tax, rates or some other cost, such as distance travelled) in order to gain an improvement in the environmental good being valued. Referendum methods ask whether a respondent accepts or rejects an offer of an environmental good at the specified improved level of quality with an attached price (in the form of an increased tax or some other realistic payment vehicle). Referendum CVM methods are regarded as being easier to answer and less biased than open-ended (Section 5.2.4), but use information inefficiently, requiring far larger sample sizes, and also more complex statistical analyses, sampling strategy and survey design. Referendum approaches have been adapted recently to elicit and use more information (e.g. through ranking or rating of alternatives which include attributes and a bid WTP amount). These later adaptations have essentially created a continuum from traditional CVM to CA. The following is based primarily on Jakobsson and Dragun (1996) and Freeman (1993). Useful context and detail was also provided by

Mackenzie (1993), Loomis (1995), Boxall *et al.* (1996), Hanemann and Kanninen (1996), Bockstael *et al.* (2000), Farber and Griner (2000), Hanley and Spash (1993) and Stevens *et al.* (2000).

For either open-ended or referendum data one is trying to find the point at which the utility with an improvement in environmental quality and a payment (loss in income T) is the same as the utility before the environmental change with no payment (i.e. we are trying to find compensating variation). In other words, referring to equation (5.3), the point where:

$$U^i(q^1, K, Y^i - \text{WTP}) = U^i(q^0, K, Y) \text{ or where } U^{i1} = U^{i0}, \quad (5.6)$$

where the superscript i refers to the individual i and the superscripts 1 and 0 to the environmental quality after and before the change respectively.

The results from open-ended WTP questions ($T = \text{WTP}$), give mean and median WTP directly (Mackenzie 1993). For interest, and assuming a particular form, the WTP responses (rather than utility) can be regressed using ordinary least squares multiple regression against the attributes (Q, K) to find the relationships between them and to estimate marginal WTP for specific Q attributes from their regression coefficients (so utility is not estimated). The WTP response or other data may be transformed using logarithms or some other appropriate transformation. Thus, open-ended CVM directly answers equation (5.6). However, as the responses are likely to be strongly non-normally distributed (with a few extreme bids and many zeros), using the sample mean may be biased, therefore maximum likelihood methods based on particular proposed distributions are sometimes used (e.g. FAO 2000), extremes or 'protest zeros' may be removed, or the median might be used. Open-ended CVM was used in the cases described in Sections 8.3 and 8.4 (mainly because we felt that sample sizes would not be large enough to handle the statistical analysis of referendum CVM formats) and the sample mean and median WTP were used. Open-ended CVM was applied in the case studies described in Sections 8.3 and 8.4.

For referendum CVM the respondent i gives a binary response (yes/no) R^i to a proposed bid T : i.e. $R^i = g(q, K, A)$. The responses need to be analysed in terms of the probability of accepting or rejecting the proposed change in quality, as different respondents are proposed different 'bid' amounts for the change in quality (i.e. $\Pr[\text{yes}] = \Pr[\text{WTP}^i \geq T]$). Given the binary dependent variable, ordinary regression is unsuitable as predicted values for R^i of >1 and <0 will arise. Interpreting the yes/no response as binary (0,1), allows us to make use of binary response regression such as the logit and probit models which restrict the dependent variable to between 0 and 1. The aim is to model the probability that a respondent will accept or reject the bid, the probability that $U^{i1} \geq U^{i0}$. If T is the proposed cost of the improved environmental quality (e.g. an increase in tax), the individual responds yes (accepts the bid) if $U^{i1} > U^{i0}$, or if $U^i(q^1, K, Y - T) - U^i(q^0, K, Y) \geq 0$ and no otherwise. In a RUM framing (and remembering that the original RUM, an individual drawing from a family of utilities, is now a RUM of a sample of the population with fixed utilities), if $V^*(.)$ is the observable component of utility, the probability of a yes response is given by:

$$\Pr[\text{yes}] = \Pr[U^{i1} > U^{i0}] = \Pr[V^*(q^1, K, Y - T) + \varepsilon_1 \geq V^*(q^0, K, Y) + \varepsilon_0].$$

The logit model also fits with the assumption that the ε are independently and identically distributed with a type I extreme value distribution (Freeman 1993) or Weibull distributed (Jakobsson and Dragun 1996). Then the probability, ignoring K , is given by the logit (or logistic) probability density function. With, for example, a linear function $R^i = \alpha + \beta_j T$, this is:

$$\Pr[\text{yes}] = \frac{1}{(1 + e^{-(\alpha + \beta T)})}$$

If the random variables are assumed to have a normal distribution then the probit model is appropriate:

$$\Pr = \Phi(\alpha + \beta T), \text{ where } \Phi \text{ is the cumulative standard normal distribution.}$$

But what is actually of interest is if the bid T proposed to the respondent is greater than or less than her real WTP, so as to find the point where $V^{i*}(q^1, Y - \text{WTP}) + \varepsilon_1 = V^{i*}(q^0, Y) + \varepsilon_0$ (i.e. to find compensating variation). Viewing WTP as a random variable, then $\Pr[\text{yes}]$ can be viewed as a function of WTP (rather than as a function of T and Q as above) and then $\Pr[\text{yes}]$ is $\Pr[\text{WTP} \geq T]$. This allows one to view the problem as a cumulative density function, with WTP a random variable in V , where the expected value of WTP or $E[\text{WTP}]$ is the shaded area above the cumulative distribution function and below $\Pr[\text{no}] = 1$ (Figure 5.1). Attributes other than bid (T) are seldom included in the regression, and in the logit model above this means that $E[\text{WTP}] = -\alpha/\beta$ (both median and mean). Therefore, WTP (compensating variation) for q^1 is found when $V^{i*}(q^1, Y - \text{WTP}) + \varepsilon_1 - \varepsilon_0 = V^{i*}(q^1, Y)$, and this probability can be solved as above with the logit or probit models. The $E[\text{WTP}]$ can be multiplied by the relevant population to find total WTP.

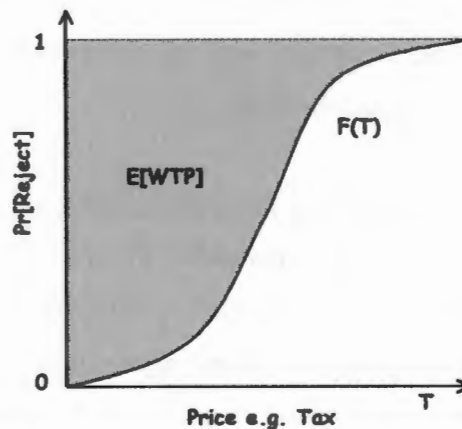


Figure 5.1. The cumulative density function for rejecting the offer of q^1 at a cost of T . The shaded area gives the expected value of WTP, $E[\text{WTP}]$ (from Freeman, 1993).

Cameron (1988) has developed a procedure for analysing referendum WTP data that does not rely on the random utility (or utility) model, but rather derives a WTP function directly. This allows estimation of the marginal values for any variable in the WTP function, while the utility function model (as above) does not (i.e. in order to obtain the $E[\text{WTP}]$ welfare measure, only the bid itself is included as an independent variable in the model).

5.2.2 Stated preference: Contingent behaviour valuation (CBV)

A variation of the classic WTP contingent question is to ask for changes in behaviour (or level of activity, e.g. frequency of visits) due to hypothetical changes in ecosystem quality (e.g. McConell 1986, Thayer 1981 in Freeman 1993 and Jayne *et al.* 1996 in FAO 2000). These changes in behaviour or activity can be translated into a price in some way, for example, by multiplying an increased frequency of trips by the cost of a single trip. Then changes in *status quo* travel cost obtained from the TCM, or changes in *status quo* WTP obtained from a CVM can be used to estimate changes in demand or WTP. Some feel that contingent behaviour questions are less susceptible to bias from implied value cues than contingent valuation (Freeman 1993). CBV was the approach used in Chapter 7 to find changed visitation behaviour with specific scenarios that could be associated with travel costs and utility estimates.

5.2.3 Stated preference: Conjoint analysis (CA)

As a point of clarification, we need to look at the different meanings and methods associated with the terms conjoint scaling, conjoint measurement, and conjoint analysis that are sometimes used interchangeably in the literature. Sarin explains that “The word conjoint is used to highlight that the objects and their components are measured simultaneously” (Sarin 1990 p. 279). In my terminology this quote would refer to alternatives and their attributes.

Conjoint *measurement* is an approach for deriving additive value functions. Keeney and Raiffa (1976) and Von Winterfeldt and Edwards (1986) use the term conjoint measurement to refer to a specific indifference technique for value measurement as opposed to numerical estimation techniques and value function construction. Using indifference questions, interval scale value functions are constructed (i.e. strength of preferences are not required). This requires joint independence which is that preferences in any subset of attributes must be independent of fixed levels in the remaining attributes. The procedure used has been called variously, the dual standard sequence, saw tooth, or lock-step procedure. This is because the value function is created by asking for the value of successive different x_2^* which will make the respondent indifferent between (x_1, x_2) and (x_1^*, x_2^*) where x_1^* is fixed at the same level throughout - this creates a saw-tooth shaped graph. Another approach is to directly compare two alternatives with two attributes each (x_1^*, x_2^*) , and (x_1^0, x_2^0) and to directly ask for the level $x_2^?$ which makes the two equivalent, i.e. towards which the respondent is indifferent. This is precisely the indifference curve approach to deriving utility found in economics (see Figure 3.2). This sort of procedure is practically possible only when the analyst and decision-maker directly interact (Sarin 1990).

The term conjoint *scaling* is used to refer to a process whereby alternatives are compared only two attributes at a time, which allows (with e.g. linear programming) the estimation of interval scaled single attribute value functions (e.g. Belton and Stewart 2002). A few attribute levels of two criteria x_1 and x_2 (call the levels x_{ij}) are defined and all combinations are laid out in a table. The respondent rank orders these two attribute alternatives. Then, if for example, alternative (x_{11}, x_{22}) ranks 1 and (x_{12}, x_{21}) ranks 2 then $v(x_{11}) + v(x_{22}) > v(x_{12}) + v(x_{21})$, etc. Set up as a linear programming (LP) problem where e.g. $v(x_{11}) + v(x_{22}) - v(x_{12}) - v(x_{21}) + \delta_1 \geq 0$, the sum of δ_i s can be minimised.

Conjoint *analysis* (CA) refers to the holistic rating or ranking of multi-attribute profiles, from which ‘part-worths’ (single attribute utilities or values) can be determined through some estimation procedure (regression, maximum likelihood, LP). This is usually the approach used in survey-based techniques such as market research (Green and Krieger 1993) and economic valuation. Sarin refers to this as a “statistical procedure for conjoint measurement” (Sarin 1990, p. 293), thus clarifying the distinction between the different conjoint methods.

To continue, rather than presenting respondents with one specific change in an environmental good, CA presents respondents with a number of alternatives which are combinations of attributes and attribute levels. The respondent typically gives a rating or ranking, and from this the part-worths of each the attributes can be determined. Where cost is included as one of the attributes, the economic value (shadow value, WTP or marginal benefit) of the other attributes can be estimated (Farber and Griner 2000) (as with SMART, see below). As with open-ended CVM, ordinary least squares could be used, but may be biased and therefore maximum likelihood estimation, logit or probit models may be used (Mackenzie 1993).

Various forms of CA have been applied in economics valuation exercises and tested for their usefulness, including ratings, ratings differences, rankings, and binary choice models with or without strength of preference indications (e.g. Farber and Griner 2000). The first presents a number of alternatives which are rated on a given scale (e.g. 0-10), the second uses the differences between the rating given to the *status quo* and to other alternatives in the analysis, the third only asks for the rank order of the options, and the fourth asks if a specified multi-attribute option would be chosen (more or less reducing to CVM), or which of a number of options would be chosen (with or without strength of preference for the choice). When using ratings (i.e. direct estimation of utility of a multi-attribute package), the issue of interpersonal comparisons arises, as people may use the scale differently (e.g. centre at different places, avoid the end-points, interpret end-points differently etc.).

Besides estimating the part-worths and trade-offs between attributes, the valuations can be used to find the compensating variation or WTP (amount of money which makes two compared alternatives equivalent in utility, Farber and Griner 2000) e.g.:

$$U(Q^1, K, Y - \text{WTP}) = U(Q^0, K, Y)$$

In a RUM framework, $U() = V() + \varepsilon$, and the probability of choosing one alternative over another is:

$$\Pr[V(Q^1, K, Y - \text{WTP}) + \varepsilon_1 > V(Q^0, K, Y) + \varepsilon_2] = \Pr[V^1 - V^0 > \varepsilon^0 - \varepsilon^1]$$

The functional form of $V(.)$ needs to be established and then the parameters in V can be estimated. Often in economics applications an additive form is used (requiring that attributes are preferentially independent or what von Winterfeldt and Edwards (1986) call joint independent) or income interacts with one or more of the other variables. Then, for the purposes of an economic valuation compensating variation (WTP) for the sample can be determined. If we assume a simple weighted addition model, then:

$$V = \alpha + \beta_1 Q + \beta_2 Y, \quad (5.7)$$

and once the parameters are determined using ordinary least squares or maximum likelihood, compensating variation for an individual or WTP^i is:

$$WTP^i = -\beta_1 (Q^1 - Q^0) / \beta_2 \text{ (Farber and Griner 2000)}$$

CA was used in the EE cases described in Chapter 7 and Sections 8.3 and 8.4. No monetary attribute was included, so ‘part-worths’ of the attributes were determined by relating them to monetary measures obtained by other means.

5.2.4 Stated preference approaches and survey biases

Choosing which model of these or the many other models proposed is appropriate (the form of the utility function, the distribution of the ϵ terms, etc.) given the economic theory constraints, is a subject of debate in the literature (e.g. Freeman 1993, Hanemann and Kanninen 1996, Jakobsson and Dragun 1996). The statistical and econometric models have become more and more sophisticated, as has survey design in terms of trying to avoid some of the known survey biases. The various recognised biases which may arise from the application of survey-based methods include (e.g. Dixon *et al.* 1986, Arrow *et al.* 1993, Carson *et al.* 1994, UNEP 1995):

- *Starting point bias*: Responses in bid CVMs may be biased by the reference point provided by the initial proposed bid.
- *Hypothetical bias*: Bias caused because the hypothetical amount will not actually be made, because there is no incentive to answer correctly or because the respondent might simply find the questionnaire unrealistic and boring.
- *Strategic bias*: The answer may depend on what the respondent thinks will be done with their answer, for example, if the respondent thinks she will actually have to pay the amount. Private goods (e.g. improved water supply) are generally less susceptible to this bias than public goods.
- *Information bias*: The information provided may not be sufficient to give a realistic valuation, or it might not be balanced.
- *Embedding and scale bias*: Respondents often give a positive WTP for a subset of the resource in question, but give the same amount for the entire resource (e.g. to prevent the death of 1% or 5% of birds by oiling, or to preserve different amounts of forests from logging). Associated to this is the fact that when asked about single issues (e.g. logging of old-growth forests), significant WTPs might be offered by respondents. However, given that they might be concerned about many other environmental problems (e.g. oiling of seabirds, threats of extinction to Bengal tigers), their aggregate WTP for all of these would be an unrealistic proportion of their budget.
- *Instrument bias or vehicle bias*: Respondents may be hostile to the means of payment (e.g. taxes) and have a zero WTP, whereas if asked the same question with another vehicle (e.g. a donation to a conservation organisation) they might have a positive WTP. Associated with this are protest bids when the good in question is usually free.
- *Interviewer bias*: The interviewer or enumerator may unconsciously provide cues to which the interviewee responds to ‘please’ the interviewer.
- *Warm glow bias*: Respondents may give unrealistically high WTPs as it gives them a “warm glow” to contribute to a worthy cause such as preserving open space areas for future generations.

- *Time to think bias*: Relative to those who are not given time, people may give significantly different WTPs if they are given time (a day or two) to think about the issue before giving their WTP (e.g. Whittington *et al.* 1992).

Referendum CVM is regarded as less biased than open-ended CVM as it avoids or reduces the effects of some of these biases and therefore was recommended by (Arrow *et al.* 1993) as the preferred approach. However, the large sample sizes required for this type of survey mitigate against its application in certain contexts. Note that any scoring approach (i.e. including MCDA approaches) may also be subject to a biases depending on whether changes are represented as gains or losses.

5.2.5 Revealed preference methods – the travel cost method (TCM)

Only one revealed preference method is discussed, TCM, which is commonly used to measure amenity, recreational or tourism value of an environmental amenity or recreational site. This is a survey-based technique relying on the concept of individual consumer's surplus and aggregate consumers' surplus (see Section 3.4.1). The zonal TCM is based on the idea that travellers from different zones (i.e. distances from the destination) to a recreational area, pay different amounts which increase with distance. Thus, those who stay closest to the destination enjoy a large consumer surplus, as, in theory, they would be willing to pay more to get to the destination because others from more distant zones, have shown willingness to pay more. In other words, inferences are made about an individual's true WTP (price paid + consumer surplus) based on others' revealed WTP, by developing a demand curve from the sample. A regression of actual quantity (e.g. number of visits per year or number of days stay) against actual price (travel cost in money and or time), gives parameters for estimating *predicted* quantity with an addition to the travel cost P in order to develop the Marshallian demand curve. The consumers' surplus can then be determined by finding the area under the demand curve above the actual price (Figure 5.2). Therefore, from the actual expenditure to get to the destination, and at the destination, TCM determines the 'total WTP' which is equal to the price paid plus consumer surplus. Some assumptions of TCM are that users from all zones have the same benefit which is equal to the travel cost of the most distant (marginal) user, the consumer surplus of the most distant (marginal) user is zero and that consumers in all zones would consume the same quantity Q at a particular price P (UNEP 1995).

In particular, in a zonal TCM, visitor numbers per population of the zone of origin are regressed against travel cost (TC) for each zone, often using the semi-log form: $\text{Ln}(\text{visits per inhabitant}) = \alpha + \beta \text{TC}$ (e.g. Strong 1983, UNEP 1995). Then, using the parameters derived for this equation, the number of visits per zone are estimated with a series of hypothetical *additional* costs P , allowing a demand curve to be set up e.g. $\text{Ln}(\text{visits per inhabitant}) = \hat{\alpha} + \hat{\beta}(\text{TC}+P)$, where $\hat{\alpha}$ and $\hat{\beta}$ are the estimates from this new regression. Consumer surplus for each zone (using the original TC and P equal to zero) can then be determined using $\hat{\alpha}$ and $\hat{\beta}$ by finding the 'area under the demand curve'. In general terms if quantity is a function of price, $Q = f(P)$, then the area under the demand curve from p_K to p (the maximum price) is the consumer surplus for the consumer paying p_K . This is given by the integral $\int_{p_0}^p f(P) dp$

(Haeussler and Paul 1996). In our case, where $p_K = TC_K$, consumer surplus or $CS_K = [e^{(\hat{\alpha} + \hat{\beta}TC)} - e^{(\hat{\alpha} + \hat{\beta}TC_K)}] / \hat{\beta}$.

Since at the maximum price $e^{(\hat{\alpha} + \hat{\beta}TC)}$ will be zero this reduces to:

$$CS_K = -e^{(\hat{\alpha} + \hat{\beta}TC_K)} / \hat{\beta} \text{ for each } TC_K \text{ or zone } K.$$

Graphically, Zone 1 in Figure 5.2, for which travel cost is P_1 , has a consumer surplus of P_1OP^* , while Zone 2, has the consumer surplus P_2MP^* . The consumer surpluses are added to give total consumer surplus, for example $CS = CS_1 + CS_2 = P_1OP^* + P_2MP^*$ in the two zone example in Figure 5.2.

Zonal TCM was applied in the case studies described in Chapter 7 and Section 8.3.

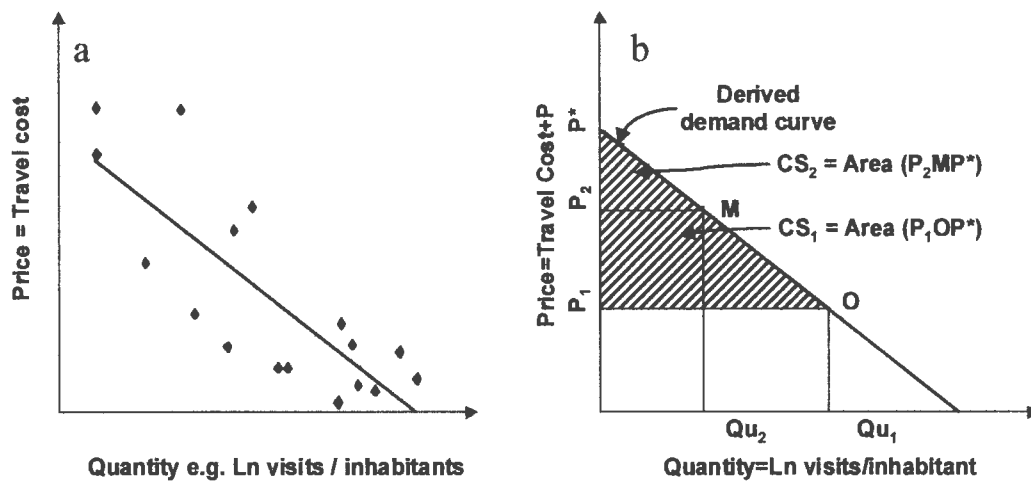


Figure 5.2. Determination of consumer surplus from travel cost. (a) - hypothetical zonal travel cost and visitor number data and the regression line. (b) - demand curve derived from the regression in (a), consumer surplus CS_1 and CS_2 for two zones.

5.3 The simple multi-attribute rating technique (SMART)

SMART (introduced in Section 4.1.1) is a simplified MAVT-MCDA technique developed by Edwards (1971, 1977), and adapted in various ways since then (see e.g. von Winterfeldt and Edwards 1986). At its heart is a weighted summation of the scores for different attributes of an alternative (equation (5.4)). Viewed from the EE perspective, SMART is in effect a stated preference approach. We applied a form of SMART described by Belton and Stewart (2002 Chapter 5, although not specifically referred to as SMART), Goodwin and Wright (1998 Chapter 2) and von Winterfeldt and Edwards (1986 Chapter 8). This implies the use of scoring on a 0-100 interval scale (either directly or using linear and non-linear value functions) for the single attribute values and the elicitation of swing weights. We also imply a workshop (or decision conference) format, where participants interact with an analyst, and a fairly extensive analysis phase which includes sensitivity analyses, and other explorations of the information elicited in the workshop(s). The weights and scores are supplied by specialists, representatives or stakeholders in a workshop in interaction with the analyst. In this setting, we are therefore applying Krzysztofowicz's (1979) group value-sharing model (see Section 4.1.3 and 5.3.3). We refer to this version as SMARTx and our application of it reflects the growing acceptance of the idea that preferences and beliefs are constructed rather than revealed (e.g. Bana e Costa

and Pirlot 1992). SMARTx together with a process for alternative development (AD, see next section) was applied two of the case studies (Chapter 6 and Section 8.1), and SMARTx on its own in one (Section 8.2).

Various stages are followed in a SMARTx decision-making process (Figure 5.3):

1. *A problem structuring stage*: Define alternatives and criteria. Develop a value tree (the terminology of Belton and Stewart 2002, Goodwin and Wright 1998, von Winterfeldt and Edwards 1986) or objectives hierarchy (Keeney and Raiffa 1976) - a hierarchical organisation of criteria and goals (e.g. Figure 5.4). This stage is iterative and non-linear within itself and with the following stages,
2. *An evaluation phase*: Evaluate alternatives and give them scores for each criterion j separately either directly $v_j(a)$ or via value functions $v_j(x_j(a))$,
3. *An aggregation phase*: Give swing weights w_j to the criteria (see Section 2.2.6). Aggregate the scores to get overall value $V(a)$ applying equation (5.4) or (4.2), and
4. *An analysis phase*: Do sensitivity and robustness analyses, further analysis of trade-offs etc..

5.3.1 Problem structuring

Various 'soft operations research' methods have been used in the problem structuring stage of SMARTx (e.g. cognitive mapping (Rosenhead 1989)), together with 'post-it' sessions and brainstorming. A typical process might start with an initial session, during which the participants respond to questions from the analyst/ facilitator on post-its. These questions are fairly broadly phrased (e.g. "what are the issues of concern in the catchment?"). Once a number of points have been written down, the participants (or the analyst/facilitator) stick the post-its on an open area of wall or board. As this happens, similar ideas start to be grouped together (by the participants or analyst/facilitator). From the ideas and notes on the post-its, criteria, objectives, and elements of alternatives begin to emerge. Criteria and objectives need to be identified and discussed by the group, 'pruned', and means-ends relationships identified etc., leading to the creation of a value tree. The elements relating to alternatives may include a wide range of issues ranging from definite proposals to vague notions, and these are taken to the next stage of alternative structuring. The criteria and alternative focussed stages that follow this initial stage, interact with and inform each other, and are not performed in a linear fashion.

Criteria and value trees

Guidance given in selecting criteria is that they should be (Stewart *et al.* 2001):

- *Complete* to ensure that all substantial interests are incorporated,
- *Operational* so that they are meaningful and understandable to all role-players,
- *Difference independent* where interval scale preference can be given for alternatives according to one criterion without reference to the performance for other criteria,
- *Non-redundant* so as to avoid double-counting, and
- *Of minimum number*.

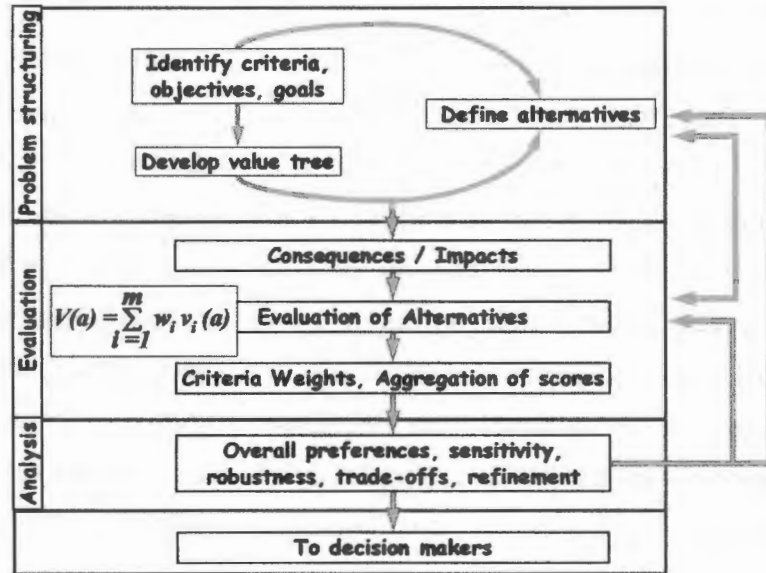


Figure 5.3. The SMARTx framework.

There are three approaches to value tree construction, the top-down (analytical), the bottom-up (synthetic) and mixed approach (von Winterfeldt and Edwards 1986). The first approach corresponds with the concept of 'value focused thinking' which Keeney (e.g. 1996), regards as preferable, as it focuses on fundamental values and may lead to more creative thinking. The second approach corresponds with 'alternative focused thinking', as the approach concentrates on finding measures to differentiate between alternatives, subsequently grouping them. In practice, value tree construction is likely to be the combined approach. In terms of 'pruning' the resultant tree, the criteria given above are relevant, however, there are further issues with respect to structure for which there is little guidance. In constructing a value tree in the top-down approach, the analyst should check that the new level adds meaning, that the lower level criterion or goal has a functional (e.g. means-end) relationship, and that lower level criteria are unique to the group in which they are situated (von Winterfeldt and Edwards 1986). How more complex functional relationships are represented in a value tree and where a criterion is positioned, if it could be included within two higher level goals (e.g. employment might be an economic or social criterion) remain 'judgement calls' made by the analyst and participants.

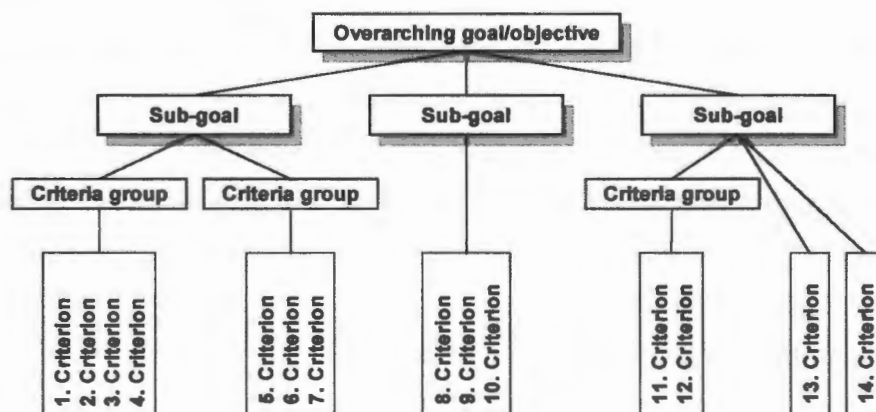


Figure 5.4. A value tree.

Alternatives

In many situations the alternatives to be evaluated are not pre-defined. Various ‘soft’ and ‘hard’ approaches are available for this stage. For example strategic choice can help to design options (Rosenhead 1989), while goal programming can be used to screen a large number of alternatives in order to select a subset for more detailed evaluation (e.g. Stewart *et al.* 2001). Stewart and Scott (1995) developed the ‘scenario based policy planning’ approach which involved the definition of ‘policy elements’ (e.g. levels of expansion or shrinkage of afforestation and irrigation, dam locations), discrete levels of which are combined to form scenarios. These are evaluated on the basis of attributes with natural scales and our found from running hydrological or economics models, for example. Applying a range of weights, the scenarios which are most frequently preferred are then taken forward for more detailed evaluation, including on the basis of more intangible criteria which could not be included in the technical screening stage. A less comprehensive, but less data demanding approach, is to determine ‘worst’, ‘best’ and intermediate alternatives from a number of points of view (e.g. ecology, economics), once the elements of alternatives (‘policy elements’) have been defined. These can be combined to form a wide range of scenarios. The latter process of alternative development (which we refer to as AD) was applied together with SMARTx (then referred to as AD-SMARTx) in the case studies described in Chapter 6 and Section 8.1 (SMARTx alone was used in Section 8.2).

5.3.2 Evaluation and analysis

The overall value or multi-attribute value V of an alternative a is therefore represented by a weighted addition of single-attribute values (equation (5.4)) on interval scales (usually 0-100) given directly $v_j(a)$ or indirectly via a value function $v_j(x_j(a))$, given that independence and other assumptions are satisfied (Chapter 2.2.4). Different approaches to the elicitation of scores were briefly mentioned in Section 2.2.6. SMARTx makes use of the swing weighting procedure described in Section 2.2.6 for weight elicitation.

In the SMARTx cases described in Chapters 6 and 8, direct scoring, value scale construction, linear, piecewise linear, and mathematical non-linear (e.g. logarithmic) value functions were used. Use was made of so-called ‘thermometer scales’ for direct scoring (Figure 5.5). VISA (1995) and Excel (Microsoft 1995) were used during workshops for calculation, and analysis. Environmental values (or impacts) may be specified by defining specific levels of performance (or impact) and associating scores with these (i.e. a ‘scoring system’ or value scale is developed), by direct judgement scoring or if a natural measurement scale is available, via a value function. Which approach is used therefore depends partly on the level of information available, but also on the particular context. For example, where a change in land-use will affect the number of ‘natural’ land-types conserved, the value of this changed may be represented by a linear or non-linear value function relating the number of land-types conserved to value (Figure 5.5) or, if species lists were available for each land-type, the number of species contributed by each land-type conserved could be calculated, and this translated to value.

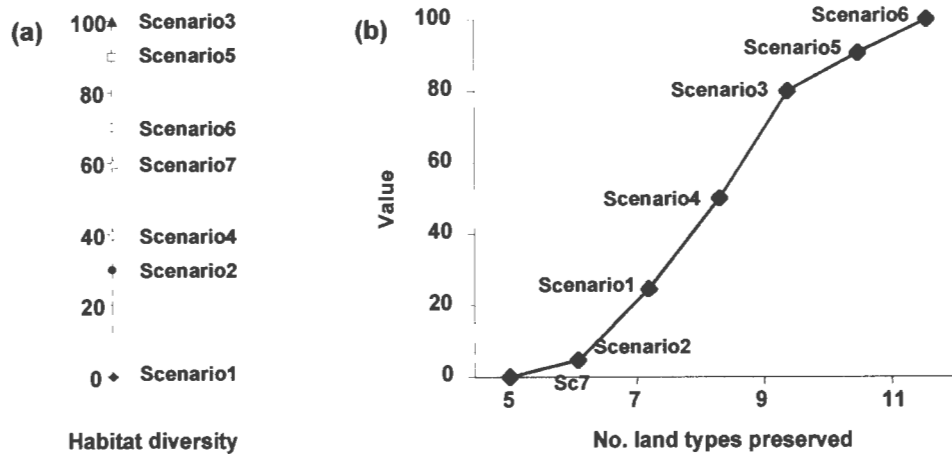


Figure 5.5. (a) A thermometer scale for habitat diversity, (b) a non-linear value function for number of land-types preserved.

Where a criterion v_2 is derived from a quantitative attribute x_2 (e.g. Rands, Mm^3), the average value in units of x_2 , C_{v_2} , of a change in v_2 can be found (i.e. x_2 per unit v_2 , e.g. Rands per value point):

$$C_{v_2} = (x_2'' - x_2') / (v_2(x_2'') - v_2(x_2')) \quad (5.8)$$

The trade-off between a pair of criteria, v_1 and v_2 , is given by their weight ratios, where w_1 and w_2 are the respective weights. By definition a one value point change on v_1 is 'worth' a w_1/w_2 value point change on v_2 , or a one value point change in v_2 is worth a w_2/w_1 change in v_1 . Therefore, if $w_1=0.6$ and $w_2=0.4$, then a decrease of 1 on v_1 is exactly compensated for by an increase of 1.5 on v_2 .

Thus, we determine how much of one criterion we are prepared to give up in order to gain in another criterion. This trade-off or marginal rate of substitution is given in value points. Where one criterion derives from a quantitative attribute we can find this trade-off in terms of the units of that attribute. Combining equation (5.8) with the weight ratio above we find the value in units of x_2 (i.e. the units of the attribute underlying criterion v_2) of a change in v_1 , or in other words, the change in x_2 that compensates for a unit change in v_1 . This is referred to as $\hat{C}_{v_1}$ as it is an indirect value, whereas C_{v_2} comes directly from the v_2 value function:

$$\hat{C}_{v_1} = (w_1/w_2) \times C_{v_2} = (w_1/w_2) \times (x_2'' - x_2') / (v_2(x_2'') - v_2(x_2')) \quad (5.9)$$

Therefore, where a monetary attribute were used, the monetary trade-off values can also be estimated (see Section 6.3), and therefore WTP in some sense can be determined. The concept of WTP may also be used directly in MCDA although it seldom appears to be applied in practice. For example, Keeney and Raiffa (1976) describe a process to directly elicit WTP by comparing two multi-attribute profiles with one monetary attribute. The process asks for a price that will make the decision-maker indifferent between the two alternatives, or that would keep her overall utility constant. Essentially this would then be a compensating or equivalent surplus (see Section 3.4). The simplest form of this procedure, where each attribute can be 'priced out' directly and separately, relies on two assumptions:

1. "The money attribute taken together with any other single attribute is preferentially independent of the complementary set of attributes", and
2. "The marginal rate of substitution between money and any other attribute does not functionally depend on the monetary level" (Keeney and Raiffa 1976 p. 126).

Point (2) means in effect that the monetary criterion should be linear (that value should have a linear relationship to value). However, Keeney and Raiffa (1976) feel that it may be easier to develop the preference structure without using the WTP transformation.

Another approach to determining WTP, suggested by Goodwin and Wright (1998), is to comparing aggregate benefits (from equation (5.4)) against a single cost measure. For example, where the decision-maker is reluctant to make direct trade-offs between benefits and monetary costs, financial costs could be kept separate. For example, in a dam site selection problem, alternative dam sites may be compared in terms of a number of 'benefit' attributes. Once aggregated, the benefits can be compared with the price of building in each location. The 'WTP' question can then be related to the alternatives on the efficient frontier, as the problem is now two-dimensional. The ratio of the change in benefits moving from site 1 to 2 to the change in costs gives a price per value point change in benefits. This can be compared with the ratio of moving from site 2 to 3 on the frontier, etc. (Figure 5.6). The decision-maker then has to assess how much she is willing to pay for a value point change in benefits, and chooses the option accordingly.

However, in various situations the approaches of Goodwin and Wright or Keeney and Raiffa may not be appropriate. For example, it is not always possible to keep attributes neatly in 'costs' and 'benefits' categories, and there may be more than one 'cost' criterion. Similarly, although mutual additive independence may apply, monetary criteria might not be linear, and so Keeney and Raiffa's indifference procedure becomes unwieldy. Therefore, in comparing alternatives with environmental impacts, use of the 'implied WTP' arising from the trade-offs determined in the SMARTx process (calculated with equation (5.9)) may be more broadly applicable, and was applied in the SMARTx case studies. However, as with Keeney and Raiffa's procedure it is also complicated by non-linear value functions, but in interpretation not elicitation. In this case marginal rates of substitution can be determined for fixed intervals of v along the non-linear value function or, if the function is differentiable, instantaneous rates of change can be determined. This means, therefore, that the interpretation of these values requires some care (see Section 6.3). For example, the marginal rate of substitution depends on the 'starting point' – the monetary level itself or the scenario from which one is moving – and so different values of 'WTP' are available (e.g. Figure 5.7). Therefore, Wenstøp *et al.* (1997) in calculating similar WTP values limited their application to linear value functions.

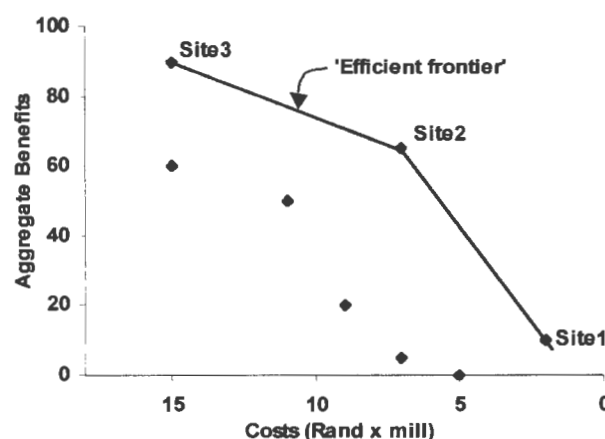


Figure 5.6. Trade-off between aggregate benefits and costs of different dam sites (adapted from Goodwin and Wright 1998).

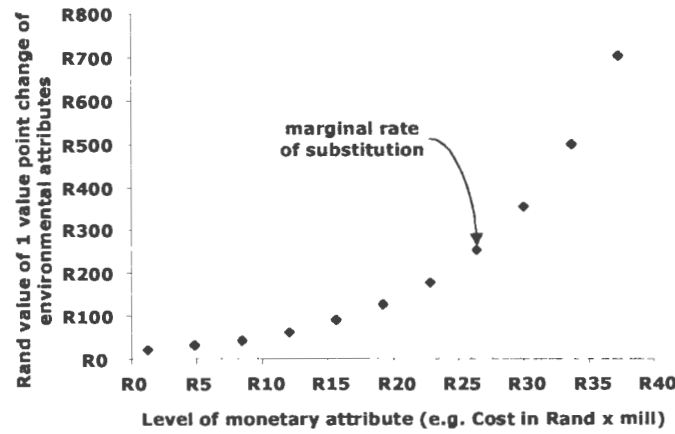


Figure 5.7. Marginal rate of substitution (MRS) of an environmental criterion and a cost attribute, where the cost attribute has a logarithmic value function. The MRS of a 1 value point change in the environmental attribute changes continuously.

5.3.3 Group decision-making model for SMARTx

Decision contexts often arise in which there is not really a single decision-maker. Rather, a group of people is given responsibility to undertake a particular decision process and make a recommendation, and a higher level decision-maker may only ratify the decision (Roy and Vanderpooten 1996). This is the context in which we have applied SMARTx. The examples of Chapter 6 and Section 8.2 were both ‘Phase 1’ processes, which were to be followed by broader stakeholder participation processes. In the Section 8.1 case, the work was also preliminary – a prelude to more in depth investigation of certain aspects. For these reason, Krzysztofowicz’s (1979) group value-sharing model is an appropriate group model. This model is explained in Section 4.1.3 and Figure 5.8. In summary, this model assumes that the group establishes *group* value functions g_j and group weights k_j directly rather than through aggregating individual views (Figure 5.8). We differ from the original in that we do not assume that each subgroup G_M^k corresponds one to one with an attribute. The case studies tended to have three main criteria groups (social, economic and environmental issues), and the specialists or G_M^k s corresponded roughly to these areas. However, there were situations when criteria cut across these disciplinary boundaries in some way. For example, employment, conventionally an economic criterion, was also a social issue, while estimating availability of secondary and natural resources for harvesting relied on the input of ecologists (these examples are from Chapter 6).

1. Define (G_M, d) = Group Decision Maker, where G_M is the set of the members of the group, and d is the decision rule to make individual to group translations.
 2. Define v_j as a group value function on criterion j .
 3. We have a function H on the v_j such that $W = H(v_1, v_2, \dots, v_n)$. We assume that H is weighted summation: $W = \sum_{j=1}^n w_j v_j$.
 4. G_M determines subgroups G_M^k whose membership corresponds to their expertise related to criteria groups.
 5. Then if individuals within G_M^k , G_M^k and G_M apply formal measurement theory techniques to find each v_j , then v_j will be the subgroup G_M^k ’s marginal value function, which means that v_j is also the whole group G_M ’s marginal value function. Translation from individual values to g_j occurs through application of rule d , which may be implicit and unknown.
 6. Using the same approach (rule d) each subgroup G_M^k finds criteria weights w_j within criteria groups.
 7. And using the same approach (rule d) the whole group G_M finds between criteria groups weights
- IN SUMMARY:** group value functions v_j are obtained directly (from subgroups or the whole group). Weights w_j within groups are found by subgroups, weights w_j between groups are found by the group. The v_j and w_j form part of W directly. Trade-offs are between v_j (criteria) rather than V^i (individuals). (Although ultimately different individuals have different interests in each v_j .)

Figure 5.8. Summary of group-value sharing model applied in SMARTx (see also 4.1.3 for the original version).

In Section 4.1.3, we applied the notation and formula $GV = \sum_{j=1}^n k_j g_j$, where g_j was the group or subgroup's partial value function, and the k_j were the group or subgroups weights. We have retained the conventional notation and formula $V = \sum_{j=1}^n w_j v_j$ in most of this section and in the case studies, where now w_j and v_j apply to group values and weights.

5.3.4 Comparison of MCDA methods - the use of SMARTx in environmental decision-making

SMARTx was exclusively applied in the MCDA cases described (together with AD in Chapter 6 and Section 8.1). While it is clear that different MCDA methods may have different roles to play in different stages of the decision-making process, it is not an obvious decision which method to choose in a particular context. There is probably no way of prescribing a particular technique for particular classes of problems, and those interested in using decision analysis approaches may be faced with a bewildering array of approaches and software packages. Rather damningly, Arrow and Reynaud (1986 p. 2) say that "The accent has been on building a large collection of multicriterion decision-making recipes without being able to decide which of them were best". Tecle (1992) claims that there are more than 70 MCDA techniques and at least 49 criteria for evaluating them. While theoretical papers propose techniques which may be interesting, useful or offer greater discriminating power, as has already been seen, these are often complex and may be of little use in real world situations where expertise and backgrounds may be widely divergent, and techniques need to be easily understood and decisions accessibly recorded. In addition, an analyst's preference for a technique may be based on experience with that particular technique rather than on its inherent superiority, and this experience will naturally influence the people with whom an analyst may work. This is clearly seen in the studies which compared the effectiveness and ease of use of different approaches from particular points of view (e.g. Tecle 1992 and Hobbs *et al.* 1992): those techniques which are favoured in one comparative study are ranked low in the next. Hobbs (1985) and Hobbs *et al.* (1992) found that choice of method could significantly affect results, whereas Goicoechea *et al.* (1992) found high consistency of results in applying different methods. Given this background, the section below describes various studies comparing MCDA methods. These studies provide some justification for the choice of SMARTx as the MCDA method in the case studies described here. However, another author, reading the same reference as those cited here and below, may have come to a different choice.

Comparative studies

Hobbs *et al.* (1992) compared the use of four different multi-criteria methods (goal programming, ELECTRE I, additive value functions, and multiplicative utility functions) and three techniques for choosing weights (direct rating, indifference trade-offs and AHP) using a real water resource planning problem in the USA. Ten alternative plans were evaluated using ten criteria (both qualitative and quantitative data). Eight hypotheses were tested relating to appropriateness and ease of use, validity, and differences in results of the multi-criteria and weighting approaches. They concluded that the simplest methods were preferred (particularly additive value functions but also goal programming), but that no method was endorsed by a majority of the 21 participants. They advised caution with the use of a direct rating approach to determining weights, as the relationship between the weights and the range of values of the criterion was seldom kept in mind, and so the weights may fail to represent the trade-offs that users are

willing to make. Methods which were poorly understood (AHP, ELECTRE, utility functions which included risk attitudes, gambling questioning) were not favoured.

Fifteen different multi-criteria methods which were applied to a watershed resources management problem were compared by Teclé (1992). Twenty-four criteria were used to assess the performance of the different techniques (a particular MCDA technique, composite programming, being used for this assessment). The criteria were divided into four groups which relate to: (a) the characteristics of the decision-maker and/or analyst involved, (b) the characteristics of the algorithm for solution, (c) the characteristics of the problem, and (d) the nature of the obtainable satisficing solution (the preferred solution which may not be 'optimal'). The evaluations stem from the authors own evaluations, those of graduate students and results of other comparative studies. Compromise programming was ranked highest overall and composite (a variation of compromise programming) second. MAUT, goal programming and the surrogate worth trade-off method were ranked 10th, 14th and 15th, respectively. The study highlighted that those approaches which are easiest to understand and flexible were preferable.

Stansbury *et al.* (1991, see Section 4.2) noted that preference functions for MAUT may be difficult to assess for some problems, while goal values for goal programming may be difficult to find for others (e.g. for ecological or social criteria). They used composite programming for the following reasons: ease of use, the double weighting scheme provided flexibility, the aggregation of criteria through several levels improves weight analysis and analysis of the system and the graphical output improves understanding.

In the context of "sustainable development", De Montis *et al.* (undated) examine seven different methods (MAUT, CIE, Evamix, ELECTRE III, Regime, goal programming and multi-objective programming using 22 criteria). Unsurprisingly, they found that each had their strengths and weaknesses, and that some might be more appropriate for different stages of the decision-making process. They found that overall Regime (an outranking method using ordinal information and the idea of 'probability of dominance' of one alternative over another) and MAUT were preferable. However, they felt that in the context of analyses of sustainability, and the need to encourage learning, communication and consensus, the particular method may be less important than that function, which all could fulfil.

AHP and SMART type approaches were used in the assessment of environmental impacts of two water development projects by Marttunen and Hämäläinen (1995). Twenty-four individual computer aided interviews were held to clarify the values of the nine different stakeholder groups. The attributes were assumed to be linear, and the importance of the criteria were determined using the swing weighting approach. The interviewees either responded numerically, graphically or verbally. In the latter case AHP-like ratio scales were used. One of the conclusions of this study was that the formulation of the value hierarchy was the most important step in ensuring a good framework for the process, particularly as the weights of the attributes were influenced by the number of attributes within a branch. The use of computers meant that the results were immediately visually available to the interviewees and sensitivity analyses could be done simultaneously. Of the two approaches AHP was found to be cumbersome and time-consuming and the motivation of the participants was affected, possibly influencing the results. The SMART

approach was found to be easier to understand and apply, required less time, and helped the participants to clarify issues and view them from a wider perspective.

Bakus *et al.* (1982) compared the use of two group decision-making approaches (the Delphi method and the nominal group technique) and three decision analysis methods (the goals achievement matrix, ELECTRE and SMART). They recommended a combination of Delphi and the nominal group technique for the formulation of values and SMART for the evaluation phase. They emphasised that the problem structuring stage was crucial to the process, and that the group interactive stages helped both to clarify individual and group views and to reduce conflict between groups. SMART was preferred to the other approaches as it was able to handle different types of input, and was felt to be simple to understand and defensible.

Barzilai and Lootsma (1997) felt that the pairwise comparisons of multiplicative AHP meant that information was collected in a fragmented way, while SMART allowed decision-makers to maintain a holistic view on the set of alternatives. In addition, they found that working with SMART was usually faster.

The comparisons above related to studies where 'formal' methods were applied. The inevitable proliferation of derivatives of methods seen in the fields of CBA, EIA and MCDA, is in a sense counterproductive to those interested in their practical application. Naturally, methodological improvements occur along the way, but the sheer number of methods may result in less rigour in application rather than more as the underlying assumptions, theory, concepts and philosophies become buried in the discussion of details, and practical guidelines may be hidden in the grey literature. Each problem encountered in natural resource management situations is often tackled afresh and a 'new' system developed. This leads to both the 'reinvention of the wheel' in various forms and the incorrect use of some techniques. Smith and Theberge (1987) looked at 20 different studies which had evaluated natural areas in terms of biological importance or significance. The type of scales used in each study (nominal, ordinal, interval, or ratio), the type of analysis (simple additive weighting, utility theory etc.) and method and degree of aggregation of scores were assessed. A wide range of approaches were used, often *ad hoc*, and some incorrect (e.g. adding ordinal ranks). Most of the studies used simple additive weighting methods (7/20), and the so-called expected value method (6/20), or a disjunctive model (either an overall ranking is based on the highest rank each area has for any criterion, or if an alternative meets a minimum standard of at least one criterion, regardless of its level for other criteria). The simple additive weighting approaches were often misapplied due to, for example, incorrect approaches to weighting or standardisation, while 6/20 of the studies incorrectly summed ordinal ranks to obtain overall scores. The approach called 'expected value method' in this study is basically an additive weighting approach, but where the criteria are ordinally ranked in terms of importance and in two of the studies the alternatives are also ordinally ranked for each criteria. The final score is a sum of the ranks for each criteria multiplied by the criteria score or rank (the latter two studies thus also performed non-permissible numerical operations on ordinal numbers). The disjunctive approach avoided the problem of 'hiding' areas of particular importance for a specific criterion (where a site may be selected because it has high diversity, irrespective of how low it may score for other criteria).

Continuing on the theme of ‘rigour’ of application, the effects of over-linearisation of value functions were assessed by Stewart (1996) and Rowe and Pierce (1982) in simulation experiments, both finding that this could significantly affect results. In testing a large number of methods with a small group of climate change experts, Bell *et al.* (2001) found that there were generally low correlations between formal models and holistic valuation, low correlations between results by the same decision-maker applying different methods, and low correlations between the results from different models. The highest correlation between decision-makers was in applying non-linear value functions with weights assessed by explicitly determining trade-offs. Disturbingly, most of the participants preferred using holistic judgement, although they felt that the formal methods were helpful in exploring the decision problem. Edwards (1977) (in agreement with Bell *et al.* (2001)) found that the application of a formal model (SMART) improved the agreement of participants compared to holistic judgements. Arrow and Reynaud (1986) and others therein, find that ordinal models produce more consistent results. Overall, Hwang and Yoon (1981) found that simple weighted models provide results that are close approximations of more complex models.

Ozernoy (1997) concludes that “different users will always prefer different methods” (p. 108), and proposes that an expert system might help with the choice of method, while taking into account individual expertise and preference, and allowing the expert system user to increase her knowledge of other methods. Given the wide range of MCDA methods available and the potential impact of choice of method on results, Bell *et al.* (2001) recommend that a number of methods should be applied.

The use of SMARTx

There is nothing conclusive about the studies mentioned above, except for a general consensus that simpler and less time-consuming approaches are preferred by decision-makers, and may be more reliably applied in everyday situations. However, oversimplification was shown to also potentially cause problems of validity. Of four well known approaches which have been applied in contexts similar to ours, AHP, ELECTRE, MAUT and SMART, we chose SMART. This was mainly for reasons of practicality given the action research setting. However, it also seemed a fairly representative MCDA approach and had similar (utility) roots to the EE approaches with which we were to compare it. Our rationale for the choice in the action research setting is expanded below.

Clearly, in environmental decision-making contexts, uncertainty and risk are important features, and therefore one might feel that MAUT would be an appropriate MCDA approach. This approach was rejected *a priori* for the case studies because of its relative complexity, a characteristic that was undesirable in a setting where people were unfamiliar with and somewhat suspicious of MCDA. Its high cognitive demands were also felt to be undesirable for the workshop settings of the case studies and where there were strong time constraints. Also, the fact that MAUT concentrates on eliciting probabilities and utilities for hypothetical lotteries rather than on the actual problem situation is also considered an undesirable characteristic (Tochner 1977). Therefore, insofar as they were considered in the SMARTx cases, uncertainty and risk were included in two ways. Firstly, through the definition of criteria as ‘the risk of...’ (e.g. the risk of exotic plant and animal invasion (Chapter 6) and ‘the risk of project failure due to lack

of consumer buy-in' (Section 8.2)). Secondly, where differences of opinion arose as to scores or weights, or where people were uncomfortable with giving a precise score, ranges were recorded and included in analyses.

Although approaches that use linguistic judgements or required only qualitative or ordinal inputs were often favoured in the discussions above, the exhaustive pairwise comparisons required by AHP were unfavourably compared to SMART (e.g. Marttunen and Hämäläinen 1995). The thresholds defined in ELECTRE are regarded by some as relatively arbitrary (Arrow and Reynaud 1986 p. 3) as they are often adjusted until a satisfactory partial ranking is obtained. Arrow and Reynaud (1986 p. 112) wonder why, if one has the information available for linear value functions one would need to apply such a sophisticated method as ELECTRE.

However, SMART has been subject to various criticisms, summarised by suggesting that it is 'dumb'. Specific criticisms are that it (a) might be easily manipulated, (b) might be subject to framing effects in the finding of scores and weights, and (c) that the inclusion of an alternative which performs particularly poorly on one criteria, but which no-one really is interested in anyway, can influence the results (i.e. too high a weight might be given to this criterion) (J-C Pomerol pers. comm.). Belton and Vickers (1990) feel that SMART-like approaches are demanding on the decision-maker and that the onus is still on the analyst (and decision-maker) to consider 'incomparability' issues when alternatives have similar scores. They feel on the other hand that SMART is transparent and flexible and can incorporate more sophisticated methods within the framework as and when required.

In contrast to Belton and Vickers (1990), the studies above and others (e.g. Goodwin and Wright 1998) feel that, compared to other methods, the judgement required by SMART (or simple weighted addition methods not necessarily labelled SMART) are simple, and the analysis easy, transparent and speedy, and results reasonably robust. Faced with the action research setting of the case studies, we had to choose one method to apply, and preferred SMART because we felt it to be the simplest to apply in workshop settings, especially where participants would include a wide range of disciplines, expertise, technical abilities, education and sophistication. It also appeared *flexible* in a number of ways. Firstly, it is flexible in terms of the specific way in which scores and weights are recorded (e.g. using dedicated computer software such as VISA (1995), general spreadsheet software such as Excel, paper, or stones distributed in a matrix drawn on the sand). Secondly, it is flexible in terms of the level of theoretical sophistication of application (e.g. to what extent non-linearities are examined). Thirdly, it is flexible in its ability to do both 'back of envelope' and highly detailed evaluations. Fourthly, and perhaps most importantly, it can include quantitative, qualitative and intangible issues. Our experiences in the workshops largely supported these views (Section 9.7), however, we cannot say that another approach (e.g. AHP or ELECTRE) might not have worked as well or better.

5.4 Summary and conclusions

This chapter has described five methods of supplying values as part of the appraisal of projects or policies with environmental consequences. Four of these were EE methods that can be used to find different types of value, and the fifth was SMARTx, an MCDA approach. The EE approaches can also value the particular environmental good

per se without the context of a particular project, although normally one should specify the value of a specified *change* in the quality of that good. On the other hand, SMARTx provides both the evaluation tools and a framework within which values can be compared and aggregated. The decision-framework for EE valuation, if one is used, is usually CBA.

Although all of the methods are intended to help in choosing between projects by supplying values for the impacts of those projects, their *modus operandi* are significantly different. The EE valuation methods rely on the hypothesis that individual WTP from surveys (while other values in a CBA might be based on non-individual monetary values), will reflect societal values for environmental goods and services. In contrast, SMARTx is a workshop-based technique. Therefore, as applied here, to projects with environmental impacts, SMARTx does not address individual values, rather its use is based on the implicit or explicit view that the values supplied in the workshops represent societal values. The case studies described in the following three chapters had potentially wide environmental and social implications, which, depending on the paradigm applied were therefore either measured on the basis of individual or ‘group’ values.

EE valuation methods are intended to broaden the applicability of, for example a CBA decision-making framework, by allowing environmental impacts to be valued in monetary terms (individual WTP). Therefore the concepts of use and non-use values are defined, and the various methods are applicable to different types of value. The TCM is applicable to use value (specifically tourism or recreation), while CVM, CBV and CA are often used to supply non-use, specifically ‘existence’, value. SMARTx does not pre-define the types of value to be included, and, as monetary values are not required, any type of impact may be included, even where only qualitative information is available.

The descriptions of SMARTx evaluation and EE valuation methods (within a CBA framework), show that, although similar ‘steps’ are followed, in application these steps take on quite different flavours within the two paradigms. Therefore it is most clear in application that they are different: the next three chapters describe the implementation of both approaches. Issues arising from the case studies, together with more fundamental philosophical differences, will be discussed in Chapters 9 and 10 in light of the metacriteria mentioned at the beginning of this chapter.

PART II - CASE STUDIES

6. Development and analysis of catchment land- and water-use alternatives using SMARTx

This chapter describes an approach, consistent with the philosophy of integrated catchment management and integrated water resource management, to the evaluation of different ‘futures’ for a catchment. The philosophy of integrated water resource management is embodied in the National Water Act of South Africa (NWA, Act 36, RSA 1998), which the study largely predated. The use of alternative development (AD) and SMARTx as decision support in the formation and evaluation of potential catchment land-use alternatives is described. The results are presented, and the implications of these in terms of trade-offs between criteria and alternatives, and the costs, benefits and sources of value of different land-uses are discussed. The approach used in this case study (together with the SMARTx studies described in Sections 8.1 and 8.2) is compared in Chapter 9 with the EE applications in Chapter 7 and Sections 8.3 and 8.4. The study is therefore an important source of many of the conclusions of this thesis. Of particular relevance are the criteria that were selected (Section 6.1.3 and Figure 6.2), the value functions (Figure 6.3), the analysis of implied trade-offs and derivation of indirect compensatory values (Section 6.3.2), and some of the practical issues that arose (Section 6.3.4).

The Sand river subcatchment was chosen for the broader project (see Pollard *et al.* 1998), of which this study formed a part (see Joubert and Pollard 2001), as it was recognised that the natural resources of the catchment were degraded. This was due to inappropriate land- and water-use, which precipitated socio-economic problems which, in turn, exacerbated the environmental problems. The perceived water resource and land-use management problems in the catchment were, amongst others, water-use by exotic plantations in the upper catchment, water-use by the irrigation schemes, lack of payment for water services, lack of bulk supply to some areas, shortages of water in the lower catchment, inappropriate land-use (e.g. irrigation schemes in a water-poor environment) and poor land-use practices (e.g. plantations in riparian zones and on steep slopes), which contribute towards erosion and high silt loads in the Sand River.

The Sand river subcatchment (1910km²) is a subcatchment of the Sabie catchment and contributes about 20% of the Sabie’s mean annual runoff. There is high inter-annual and spatial variation of rainfall: the escarpment in the west has an average annual rainfall of about 2000 mm, while the eastern side of the catchment has about 550 mm. Three hydrological studies for the catchment under present afforestation levels, arrived at widely different estimates of mean annual runoff ranging from 96 to 215 Mm³ (Pollard *et al.* 1998). In terms of land-use patterns, the catchment can be broadly viewed as three zones (Figure 6.1). The upper catchment has some 5000 ha of forestry plantations, the lower catchment is commercial and state nature conservation, while the middle catchment is where most human activity occurs, including some government irrigation schemes, grazing, dryland crop farming, small garden plots,

and small urban areas. The 1998 population was approximately 337 000, amongst whom there was 40-80% unemployment (Pollard *et al.* 1998).

6.1 Methods

As part of the development of an integrated catchment management plan, use was made of SMARTx (see Section 5.3) to evaluate the alternatives and a simplified form of “scenario-based policy planning” (Stewart and Scott 1995) for the development of alternatives (referred to as AD) (see Section 5.3). The combined AD-SMARTx process provided the framework and tools to develop and evaluate hypothetical land- and implied water-use alternatives for the Sand river subcatchment. The objective was to recommend an overall ‘preferred’ land-use combination scenario for implementation, and to be able to assess what the gains and losses would be if this alternative or another were chosen. The land- and water-use scenarios chosen would retrospectively be translatable into a management class as required by the NWA, an exercise that is presently underway.

The group process is described below (Section 6.1.1) followed by a description of the development of the catchment alternatives (Section 6.1.2). The selection of criteria, formation of the ‘value tree’, and evaluation of alternatives follows (Sections 6.1.3 and 6.1.4). The derivation of the weights used in the summation of scores is then described (Sections 6.1.5 and 6.1.6).

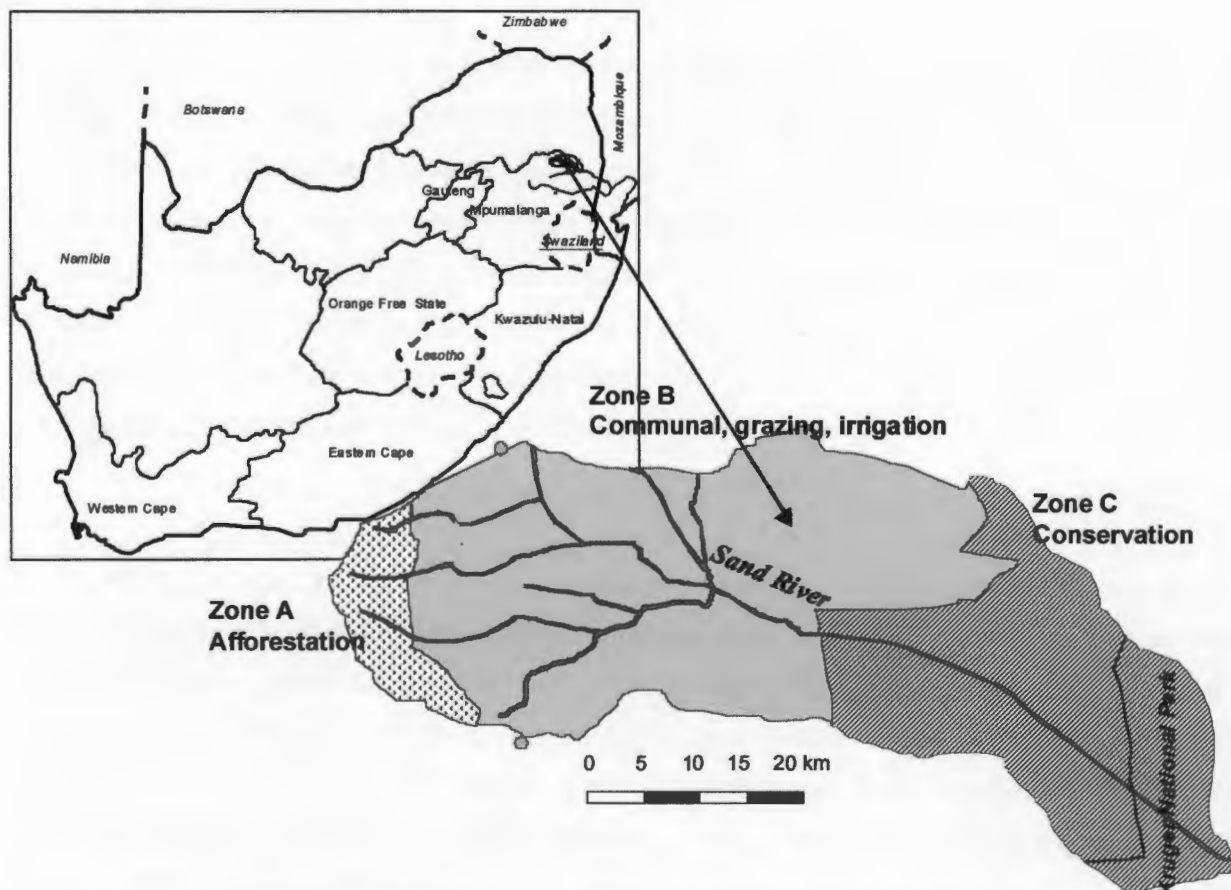


Figure 6.1. Zones used for the development of alternative alternatives within the Sand River catchment.

6.1.1 Group membership and group process

A team of specialists (mainly from AWARD⁸ - Table 6.1), were tasked (by the client, the Dept. of Water Affairs and Forestry) with the development of an integrated catchment management plan with associated land-care and water conservation plans for the Sand river subcatchment, Mpumalanga, under the co-ordination of Dr Pollard. The time frame of the project was approximately three months, during which time the project team had to collate all available information for the catchment, as well as do further research where necessary and possible (e.g. hydrology, economics and water-use of irrigation schemes). In parallel, the team participated in the AD-SMARTx workshops, which formed part of the broader study. The various stages of the AD-SMARTx process occurred in four day or day and a half long workshops with the project team who represented various points of view (viz. ecological, social, economic - Table 6.1) while intermediate analyses were completed between workshops. The process included problem structuring, formulation of alternatives, identification of criteria for the evaluation of alternatives, aggregation of scores, and sensitivity analyses. The group process (rule *d*) included initial post-it sessions, group discussions, breakaway sessions for evaluations within specialist areas (subgroups G_M^k), and feedback sessions.

Table 6.1. Participants in the SMARTx workshops.

Participant	Organisation	Areas of expertise
Sharon Pollard	AWARD	Aquatic ecology, medicinal use of plants
Charlie Shackleton	AWARD	Terrestrial ecology, natural resource use & valuation
Tony Poulter	Working for Water	Exotic plant clearing, forestry
Juan-Carlos Perez de Mendiguren	AWARD	Economics, social use and value of water
Willie Lubbe	Economist	Economics, especially agricultural
Phillip Walker	AWARD	Social issues
Peter Sekgobela	AWARD	Social issues
Temelo Mashego	AWARD	Social issues

The workshops were run and scores and weights were derived during them without the use of a data projector and decision support software. This was partly a conscious test of the application of SMARTx with these constraints (we were concerned that computers might be off-putting in some contexts), and in part due to logistical problems (specifically the availability of a data projector and its transport). Various software was used for intermediate analyses, including inbetween sessions, during which participants would group around the computer. This software included Excel (Microsoft 1995), Visual Interactive Sensitivity Analysis (VISA 1995), and Decision Explorer (1996). Results were reported back to the project team using an overhead projector and flipcharts. Only Excel was essential to the process (and was also used inbetween sessions where the group or subgroups might cluster around a computer).

6.1.2 Problem structuring: Development of catchment land-use alternatives

Note that the development of alternatives took place in parallel to the selection of criteria during the first and second workshops, and the alternatives were further refined during the third workshop. Considerable time and effort went into defining which parts of the broader project we could realistically tackle, and thus to defining appropriate criteria and alternatives.

⁸ Association for Water and Rural Development, Private Bag x483, Acornhoek, 1360.

In scenario based policy planning (Stewart and Scott 1995), alternatives are formed by combining discrete levels of elements of the alternatives (Section 5.3.1). With a number of elements and levels defined, the number of possible alternatives quickly becomes very large. This larger set (or a subset) forms the background set, which in scenario based policy planning is pre-analysed in various ways to select a 'foreground set' small enough for the team to evaluate. Time and information constraints meant that this process could not be undertaken (hydrological and economic data would have been a minimum prerequisite and neither were available until the final workshop), and a manageable set of alternatives had to be identified in other ways.

The catchment was divided into three management zones based on areas of similar present land-use patterns, climatology, topography, and demographic patterns (Figure 6.1 and Table 6.2). The team was asked to identify 'alternative elements' (e.g. land-uses, dam sizes and location, water allocation plans), which, by combining in different ways, could form catchment plans. After initial discussions, it was agreed that the primary alternative element (i.e. the primary element of any catchment management plan in the context of the Sand river subcatchment) was land-use, and that considering the time constraints, trying to analyse more complex alternatives would be counterproductive. Land-use was the driving force behind all other economic activity in the catchment (there being no heavy or service industry apart from that associated with tourism), and was also a direct cause of most of the environmental problems in the catchment. Therefore different levels (i.e. areas) of the land-uses were combined to form alternatives, which we called scenarios, within each of the three zones. In order to keep the number of scenarios to a manageable size (for the timely completion of hydrological and economic models), the team was prompted to imagine their worst and best case scenarios from their respective points of view (e.g. extremes such as 'all plantations removed'). These were modified to be more realistic (e.g. clearing 50% of plantations in the next 20 years) when considering the real land-use-potential and political constraints and intermediate scenarios also identified. The final set of alternatives was then a combination of worst, best and intermediate scenarios from each of the three main points of view. Thus, land-use scenarios were developed and evaluated separately in each of the three zones (Figure 6.1): eight in Zone A, three in Zone B, and four in Zone C (Table 6.2).

The number of hectares of each land-use in the scenarios was based on the realistic potential for certain land-uses in the different zones. For example, there was some potential for more irrigation in Zones A (about 1890 ha) and B (about 3250 ha), based on slopes and soils, and there was some potential for increased afforestation in Zone B (about 7300 ha), based on slope, soils and rainfall. The eight land-uses in the catchment were: conservation, woodlands (with grazing), exotic commercial afforestation, dense residential, residential with garden plots, permanent irrigation, annual irrigation and dryland agriculture. Note that where reference is made to either conservation or woodlands, the implied land-cover is indigenous grass, bush and woodland in both cases, with some coppiced bush and overgrazed grassland occurring in woodlands.

Zone A

Zone A was approximately 11 600 ha, 43% of which was under commercial plantations of exotic species (mainly

pine). The zone was delimited by the current extent of afforestation (apart from a small section in Zone B) on the eastern boundary, and the catchment limits on the western boundary. The rest of this zone was a combination of bushland, indigenous forest, woodland and grassland, about 20% of which was used for grazing, the remainder being inaccessible. This land-use pattern was modelled as **Scenario 1** - the status quo - enabling the group to assess whether or not keeping the present level of afforestation had benefits which outweighed its environmental impacts.

The development of other scenarios in this zone was predicated on the fact that an estimated 25% of the afforested area violated present forestry practice code as it was on steep slopes, riparian and wetland areas and therefore would have to be cleared. This was therefore a minimum requirement and modelled as **Scenario 8**.

The remaining scenarios removed 50% of forestry (2497 ha), and replaced half of this (1248.5 ha) with another land-use. The remaining 1248.5 ha would be cleared to allow the return of indigenous vegetation and remain under forestry management (except Scenario 7) as most other land-uses would also be unsuitable. The expense of rehabilitating previously afforested soils was not addressed.

In **Scenario 2**, 50 % of the presently afforested area would be cleared, and half of this (1248.5 ha) would be replaced with irrigated permanent crops (trees). **Scenario 3** would replace the same area with dryland cultivation, **Scenario 4** with woodlands, **Scenario 5** with irrigated annual crops, and **Scenario 6** with residential and garden plots. For **Scenario 7**, the entire 2497 ha was assumed to be used for conservation: in this case 'community conservation' (Table 6.2).

Zone B

The larger portion of the Sand river subcatchment, Zone B (109 370 ha), comprised land under communal tenure, lying between the forestry area in the west and the conservation area in the east. This was the zone in which the majority of people live and work. Land-uses included government irrigation schemes, dryland agriculture, woodlands (for grazing and natural resource use), residential areas with garden plots used for small-scale vegetable growing (i.e. 'sparse'), and more dense residential areas (i.e. 'dense'). Potential for expansion of irrigation, afforestation and conservation was used as a basis for the scenarios. Three scenarios were formulated, all of which took into account the likely increase in population to the year 2010 and therefore the increased extent of residential areas. These expansions all occurred at the 'expense' of woodlands.

Scenario 1, the status quo for the 1998 population rather than that of 2010, was not evaluated further, but included as a reference point. In **Scenario 2** (max) irrigation and afforestation were expanded to their maximum potential levels, while also expanding conservation into the zone based on a community conservation model (Table 6.2). **Scenario 3** (mid) was similar but considered to be more realistic, as no increase in afforestation was proposed, and smaller increases in irrigation and conservation were proposed than for Scenario 2. **Scenario 4** was simply a projection of the status quo to 2010, taking into account the increase in population and the concomitant increase in residential areas.

Note that Zone B was the only area in which population growth was incorporated into the scenarios through increases in the area required for housing and in water demand, as the most likely future would be that all increases would be accommodated in Zone B. In order to calculate the expansion of residential areas, the current population figure of 336 638 was projected, at a growth rate of 2.4%, to a population of 447 469 in the year 2010. The current population divided by the current residential area (17 859 ha) gave a density of 18.9 people per hectare. At the same density the projected population for 2010 would require 23 739 ha. Split in the same proportions as at present between dense and sparse residential (0.15:0.85) gave areas of 3 656 ha and 20 083 ha respectively.

Zone C

Zone C (70 000 ha) consisted of private and state game reserves (i.e. 'commercial conservation'), and was defined by the present western borders of the game reserves and the catchment limits in the east. **Scenario 1** was the status quo: commercial conservation on 70 000 ha. For **Scenario 2**, 20 % of this land was converted to woodlands for grazing and natural resource use. **Scenario 3** was the same as Scenario 1, except that one of the game reserves (Manyeleti, 3 622 ha) came under community management with no natural-resource harvesting, while **Scenario 4** was the same as Scenario 3, but included harvesting on 20 % (13 173 ha) of the current commercial conservation area (Table 6.2).

Table 6.2. Land-use scenarios for the Sand River catchment. Measurements in hectares. SQ = Status quo, For = Forestry, PI = permanent irrigation, AI = annual irrigation, DA = dryland agriculture, WL = woodlands, R&GP = residential and garden plots, CmCon = commercial conservation model, CCon = community conservation model, NRH = natural resource harvesting, Man = Manyeleti.

ZONE A								
Total area (excl water) – 11581.7 ha	Forestry	Unused*	Permanent irrigation	Annual irrigation	Dryland agric	Woodlands	Community cons ^a , NRH	Residential + garden
Sc1-SQ	4994	5270	0	0	0	1318	0	0
Sc2-For↓50%,25%→PI	2497	6519	1249	0	0	1318	0	0
Sc3-For↓50%,25%→DA	2497	6519	0	0	1249	1318	0	0
Sc4-For↓50%,25%→WL	2497	6519	0	0	0	2566	0	0
Sc5-For↓50%,25%→AI	2497	6519	0	1249	0	1318	0	0
Sc6-For↓50%,25%→R&GP	2497	6519	0	0	0	1318	0	1249
Sc7-For↓50%,50%→CCon	2497	5270	0	0	0	1318	2497	0
Sc8-For↓25%	3746	6519	0	0	0	1318	0	0
ZONE B								
Total area – 108725 ha	Forestry	Permanent irrigation	Annual irrigation	Dryland agric	Woodlands	Residential and garden	Dense residential	Community cons ^a , NRH
Sc2-max	7307	989	4843	10275	53382	20083	3656	8190
Sc3-mid	0	724	3545	10275	66347	20083	3656	4095
Sc4-SQ2010	0	438	2145	7743	74659	20083	3656	0
Sc1-SQ1998	346	438	2145	7743	80194	15109	2750	0
ZONE C								
Total area – 69770 ha	Woodlands	Commercial conservation	Dense residential	Community conservation	NRH			
Sc1-SQ	0	69487	283	0	0			
Sc2-↓CmCon,↑WL	13897	55589	283	0	0			
Sc3-Man→Ccon	0	65865	283	3622	0			
Sc4-Man→CCon, NRH in 20% CmCon	0	65865	283	3622	13173			

* Newly cleared riparian areas, wetlands and steep slopes (1249 ha) would be forestry managed.

6.1.3 Problem structuring: Criteria and value tree formation

The terms of reference of the project team specified as the overall objective the rehabilitation of the Sand river subcatchment through the principles of 'integrated catchment management' and 'landcare'. In turn, promoting

rehabilitation and sustainability would be achieved through economic growth, equitable access to water, and sustainable and appropriate land- and water-use. Some of the criteria for evaluating scenarios developed naturally through the further refinement of these objectives while others were obtained in 'brainstorming sessions' during the workshops. The objectives and criteria were organised into a value tree (iteratively), and the scenarios evaluated on the basis of 19 criteria (Figure 6.2).

Criteria that contributed to the goal of rehabilitation and sustainability (RS) were grouped into two categories: those relating to terrestrial ecology and those relating to aquatic ecology. Effects on terrestrial ecology were assessed in terms of terrestrial species richness and habitat diversity, soil erosion, spread of exotic invasive species, all of which are directly affected by land-use. Aquatic habitat diversity, water quality and catchment water yield (i.e. effects on runoff) are directly and indirectly affected by land-use. Clearly 'sustainability' is the overall objective of an integrated catchment management plan. The inclusion of the term in this group indicates that the sustainability of resource *use* is mainly measured in terms of impacts on ecology.

The criteria contributing to the goal of economic growth (EG) were: the operating margin or profit resulting from the different scenarios (the total profit to the catchment zone accruing from all land-uses in the scenario), and total income earned in formal occupations and informally through harvesting of secondary and natural resources. A financial analysis of earnings from harvesting secondary and natural resources was adjusted to include the non-cash costs of harvesting and transportation. The resulting incomes from harvesting were considered conservative as the costs included factors such as the cost of transportation to urban centres (whereas indications are that most produce is sold locally, Pollard *et al.* 1998). Only economic implications from primary land-uses were assessed and no multiplier effects were included. It is likely that multiplier effects for any of the land-uses proposed here would be similar, all being agriculturally based, and therefore involving mainly transportation and packaging. Multipliers from tourism could be higher if more services were based outside conservation areas. Other suggested criteria were: contribution to gross geographic product, ability to attract investment capital etc., but these were rejected as it was felt that available data would not support their determination, and also that other criteria already partially measured these.

Although employment is conventionally considered an economic criterion, the two employment criteria (total number of informal and formal jobs) were included in the group of criteria contributing to the goal of social upliftment and equity (SE) as employment was the primary means of achieving this higher level goal. The criterion 'water equity' was intended to be a measure of how many people could be supplied with different levels of water supply, as access and distribution are patchy and skewed. Land equity was a rather 'fuzzy' criterion. It related to land tenure systems and access to and use of resources associated with different land-uses. For example, most of the land under natural vegetation in Zone B is communal. Access to these means that pasture and natural resources for harvesting are available. The criterion 'greenspace and aesthetics' related to how much uncultivated land remained in the catchment and 'river access' to whether or not, under new land-use arrangements, residents would have access to the river for drinking, washing, social, ritual and gardening use. The 'social value of harvesting secondary and natural

resources' related to both the cultural aspects associated with harvesting and the fact that being able to harvest meant that those with no other income or resources gained a sense of worth through relative self-sufficiency from this source. The social criteria relating to health, crime and infrastructure were regarded as fairly standard criteria to use, but for the most part were not directly affected by the scenarios being assessed. Indirect effects would mainly be due to changes in employment, remuneration and equity, which were already addressed elsewhere. However, the social specialists felt that these criteria should be retained, rather than give the impression that these issues were not considered. It was argued that double-counting could be counteracted by giving them low weights.

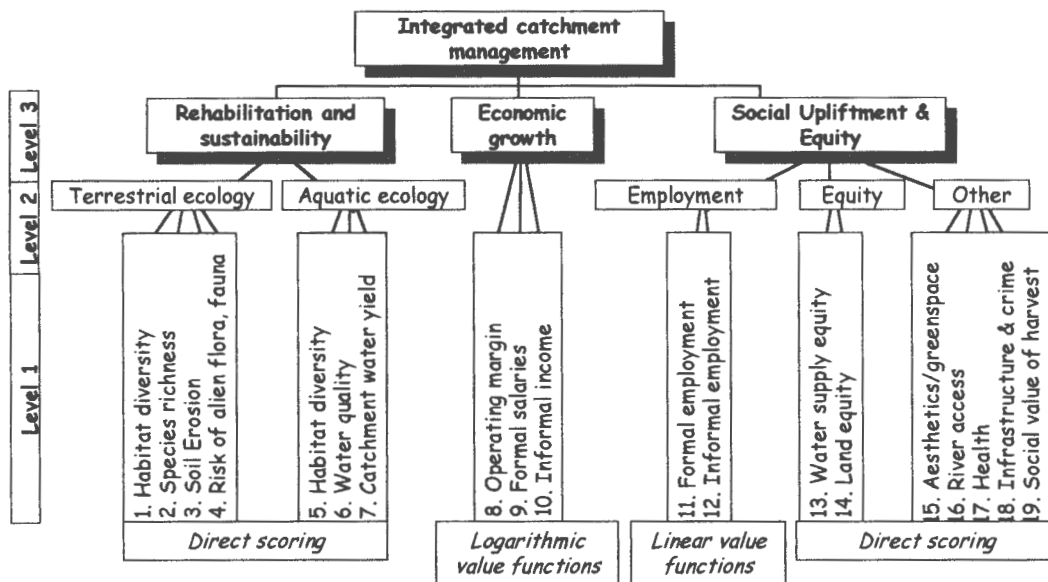


Figure 6.2. Value tree structure, criteria used and their associated scales. Criteria 7 and 13 could be quantitative once appropriate hydrological information were available.

6.1.4 Consequences and evaluation of scenarios

The consequences of the scenarios were therefore examined in terms of criteria 1 to 19 in the value tree (Figure 6.2). The quantitative or qualitative evaluations were based either on data arising from this project and previous studies, or on the opinion of the relevant specialist on the project team based on their previous experience and work in the area. Therefore, both direct judgemental scoring and value functions were used (see below), in all cases the worst alternative having a score of 0 and the best a score of 100.

Direct scoring by the relevant specialist(s) on a 0-100 scale, using thermometer scales on printed paper, was carried out for 13 criteria j (Figure 6.2 and Figure 6.3a). The score, $v_j(a)$, related indirectly to one or more unmeasured attributes, x_k , or consequences of the scenario a . In other words, $v_j(a) = f(x_1, x_2, \dots, x_k)$, where the x_k included the hectares of different land-use, but may have included other issues. Although for criteria such as species richness, some comparative, quantitative assessments were available, the use of a value function relationship was not felt to be necessary. In the absence of final hydrological models for the scenarios, specialist judgement was used for the criteria relating to aquatic habitats, catchment water yield and water equity. Value functions might have been used if hydrological model results were available.

Linear value functions (v_j) translated data ($x_j(a)$) into 'value' on a 0-100 scale for two criteria and non-linear value functions translated data to 'value' on a 0-100 scale for three criteria (Figure 6.3b). In other words, these five criteria were directly related to an attribute measured on a natural scale. Ideally, one would then assess the changing marginal levels at different levels of the attribute. Using the bisection or mid-value splitting technique one can develop a piecewise linear relationship between value and the attribute (e.g. Belton and Stewart 2002). In the case of the three non-linear value functions (operating margin, formal and informal income), the team agreed that they were of the 'diminishing marginal value' type (although this terminology was not used). Therefore, for these criteria, an increase of Δx_j from $x_{\min}$ was 'worth more' in some sense, than the same increase near $x_{\max}$. Due to time constraints, we felt that, rather using the bisection or similar method, a logarithmic shape would well represent this relationship. This was incidentally the relationship between income and value which Bernoulli had proposed would solve the St Petersburg paradox in 1738, and is a relationship often found in psychophysical measurement (Lootsma 1997). For the linear attributes, the diminishing marginal utility relationship was felt not to apply. For example, considering the high unemployment levels in the catchment, the team felt that the 'flattening off' of a logarithmic function at higher levels would be inappropriate, because unemployment would still be so severe. Therefore a linear value function seemed appropriate for employment (Figure 6.3b).

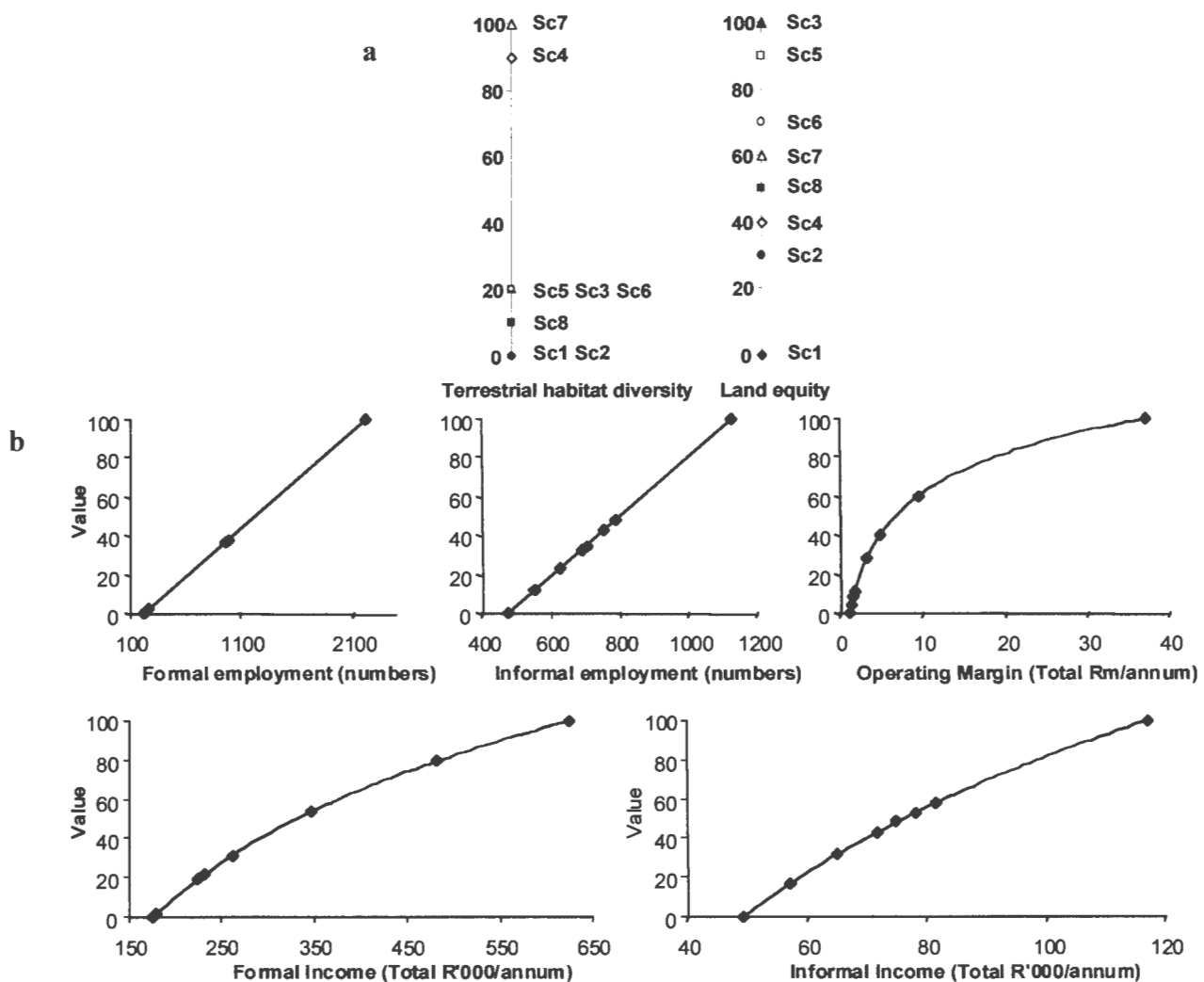


Figure 6.3 Example of (a) 'thermometer scale' direct-scoring, (b) linear and non-linear value functions.

6.1.5 Evaluation: Aggregation of scores for each zone

In each zone, the scenario's scores v_j on each criterion j were aggregated using a weighted summation of scores (i.e. SMARTx) to give an overall value $V(a)$ for each scenario a (see equation (5.4)). A range of values for the weights were used to assess sensitivity of the model (see below).

The scores could be aggregated to various levels up the value tree to guide future decision-makers. Thus, the scores for the four criteria in the group 'terrestrial ecology' were summed to give a 'terrestrial ecology' score, the terrestrial and aquatic ecology scores were summed to give a 'rehabilitation and sustainability' score, and the RS, EG and SE scores were added to give overall scores for the scenarios relative to each other. Preferred scenarios or 'directions of preference' could then be identified overall, or from different points of view, for different zones in the catchment. Overall performance could thus be compared with performance on any of the 19 criteria or with performance on the three grouping criteria. For example, a scenario may have performed well overall but poorly from the point of view of 'formal employment' or from the aggregate SE point of view. This can either help to highlight potential new scenarios, or indicate that a scenario which performed slightly less well overall was preferable because it did not score very badly for any one criterion, or was 'more balanced' in some sense.

6.1.6 Evaluation: Criteria weights

The use of weighted addition presupposes that an improvement in one criterion can compensate for a decrease in another criterion. The scales and weights used determine this trade-off and the use of a swing-weight approach will, at least roughly, provide the correct trade-off.

To find weights or scaling constants for the criteria *within* the three criteria groups RS, EG and SE, the team was divided into three groups, with the expertise of each group corresponding to each of these issues (Table 6.3). The weights were given by the relevant members of the team using the swing-weighting approach. The groups were asked the following question: "If a scenario existed which had all the criteria at the same (worst or middle level supplied), which one criterion would you want to improve to its best level if you could improve just one?". This criterion was ranked 1 and the question repeated until all criteria were ranked. By convention, the rank 1 criterion is given a weight of 100. The group then gave the importance or value of a swing from worst (or middle) to best on the next ranked criterion relative to this one (as a percentage). The procedure continued until all criteria were weighted and the weights were standardised to sum to one. The three groups evaluating within-criterion group weights, independently developed different strategies to help them in assessing weights:

- The *social group's* strategy was to develop a trade-off between the number of hectares of cattle grazing land, used as a proxy for land-equity, and the number of people formally employed.
- The *ecological group's* strategy was to decide which 'rehabilitation activity' they would choose if they could spend a million dollars on just one activity.
- In comparing formal and informal *employment* the whole group decided that they should be treated as equal, but that sensitivity analyses should assess the impact of weighting one formal job as worth two informal jobs or vice

versa, as arguments were given to support both of these ideas (this was done, but, as it made little difference, the results are not reported here).

The weights *between* the three criteria groups, RS, EG and SE were determined by the group as a whole. In practice, while appropriate elicitation of weights at the lowest (criterion) level is possible using methods such as that described above, at higher levels, determining the swing weights between criterion groups is less reliable. To determine appropriate weights for the criterion groups, two approaches may be adopted in workshops, both of which use the swing weighting questioning procedure. In the first approach, the criterion groups can be directly compared (“which group of criteria would you swing from worst to best?”), in which case, it is likely that the ‘intrinsic importance’ of the group or a criterion within the group will determine the weights. In the second approach, lower level criteria can be directly compared across all groups (“which criterion would you most like to swing from worst to best?”) and a criterion group weight inferred. The latter approach is more likely to reflect actual trade-offs between criteria and criteria groups. Due to time constraints in the final workshop, the former approach was used in this study. As the whole group was involved in discussions of the within criteria group weights it seems likely that a common frame of reference was achieved so that there was some internal reference to the implied trade-off between criteria groups, rather than just intrinsic importance. Also the later examination of trade-offs and weight sensitivity would help to reveal inconsistencies in these weights. In general, the rank order of the between criteria group weights was not disputed by the group although the relative weights differed. The weights for which there was most consensus are in future referred to as the consensus weights. The full range of weights suggested (Table 6.4) was tested in a sensitivity analysis to determine if preferences would be affected.

Table 6.3. Weights for each sub-criterion, global weights (levels refer to the levels in Figure 6.2) and ratio of each criterion weight w_i to w_{OM} (i.e. weight of operating margin) used in Section 6.3.2.

	Criterion	Contributing weights			Global weight	$\frac{w_i}{w_{OM}}$
		level 1	level 2	level 3		
RS	Soil erosion (SEr)	0.18			0.0455	0.554
	Terrestrial biodiversity	0.27			0.0682	0.831
	Risk of exotic flora, flora	0.24	0.5		0.0606	0.738
	Terrestrial habitat diversity	0.3		0.5	0.0758	0.923
	Aquatic habitat diversity	0.32			0.0804	0.979
	Water quality	0.32	0.5		0.0804	0.979
	Water yield	0.36			0.0893	1.088
EG	Operating margin	0.41			0.0821	1
	Formal income	0.51		0.2	0.1025	1.249
	Informal income	0.08			0.0154	0.188
SE	Formal employment	0.76	0.43		0.0986	1.202
	Informal employment	0.24			0.0318	0.387
	Land equity	1	0.39	0.3	0.1174	1.430
	River access	0.36			0.0189	0.231
	Aesthetics	0.27	0.17		0.0142	0.173
	Social value of harvest	0.36			0.0189	0.231

Table 6.4. Sets of weights suggested for the main criteria groups for the different zones. These were some of the weights used in sensitivity analyses. The 'consensus' weights are in bold.

Criterion group	Zone A			Zone B			Zone C		
Rehabilitation and sustainability	0.50	0.67	0.59	0.30	0.24	0.14	0.25	0.22	0.19
Economic growth	0.20	0.13	0.12	0.33	0.36	0.41	0.25	0.22	0.19
Social upliftment and equity	0.30	0.20	0.29	0.37	0.40	0.45	0.50	0.56	0.63

6.2 Results

In brief, the preferred scenarios for the three zones from the point of view of aggregated scores, were the removal of some plantations for community conservation in Zone A (Scenario 7), some expansion of irrigation and community conservation (but neither to the maximum possible) in Zone B (Scenario 3), and the inclusion of harvesting of natural products in some of the commercial conservation areas in Zone C (Scenario 4). This means that the preferred RS scenario was chosen in Zone A, the preferred SE scenario in Zone B, and the preferred EG and SE scenario in Zone C (Figure 6.4). The full set of results is given in Appendix 6.

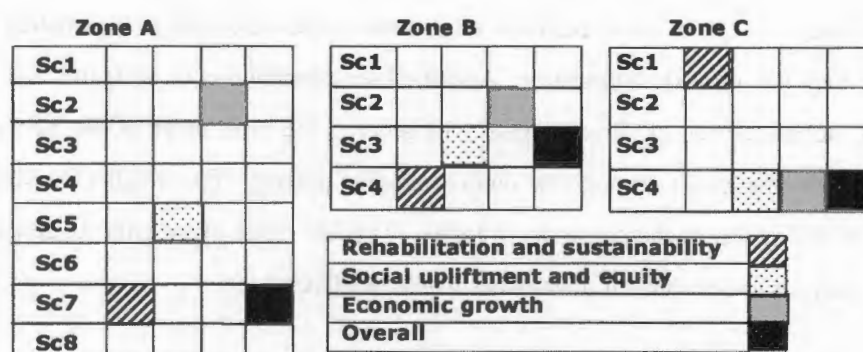


Figure 6.4. Preferred scenarios from different points of view for each zone.

6.2.1 Preferred scenarios for Zone A

Using the consensus weights the preferred option overall in Zone A was Scenario 7 - community conservation on 2497 ha of previously afforested area. Although Scenario 7 performed poorly in terms of the number of people formally employed, it had the highest level of informal employment (Figure 6.5), because harvesting of secondary and natural resources was allowed. The aggregated score for Scenario 7 was 42 % higher than the next preferred overall, Scenario 4, which had approximately the same aggregate score as Scenario 5 (preferred from the SE point of view). The scenarios naturally divided into three groups, Scenario 7 standing alone as preferred, Scenarios 4 and 5 being equivalent and possible compromise solutions, and the remaining scenarios probably being unacceptable. Note that the profile of Scenario 4 is very different to, and less 'balanced' than, that of Scenario 5. While it performs substantially better than Scenario 5 in terms of RS, it performs particularly badly on all of the EG criteria and most of the SE criteria relative to Scenario 5 (and most of the other scenarios). Also note that logarithmic relationships were used for all of the non-linear value functions (i.e. operating margin, formal and informal income), but the more important conclusions did not change if these were all or singly made linear.

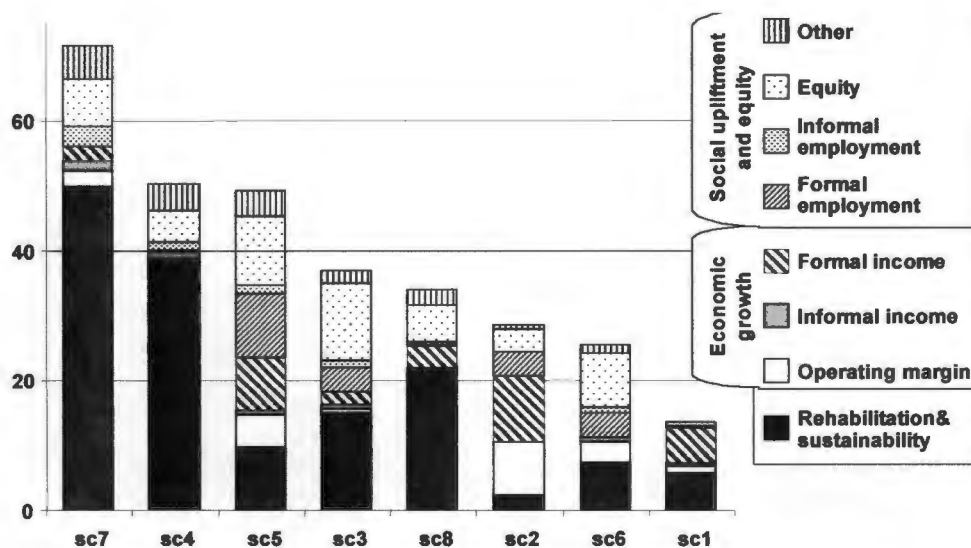


Figure 6.5. Relative contributions of criteria to overall scores of scenarios for Zone A.

6.2.2 Preferred scenario for Zone B

Only Scenarios 2, 3 and 4 were compared for Zone B, the middle catchment (Scenario 1 was not included, being a statement of the status quo without population growth). Using the consensus weights, Scenario 3 was the preferred option overall and from the SE point of view (Figure 6.6). This implied some increase in permanent and annual irrigation, dryland farming and the expansion of conservation into this zone (all at the expense of grazing). The conservation model proposed was 'community conservation', which allowed harvesting and, by assumption, employed 20% more people than that of the commercial conservation current in Zone C. Scenario 4, the projected status quo and preferred from the RS point of view ranked second overall, while Scenario 2, preferred from the EG point of view ranked third overall. There was little difference in overall score between Scenario 2 and 4.

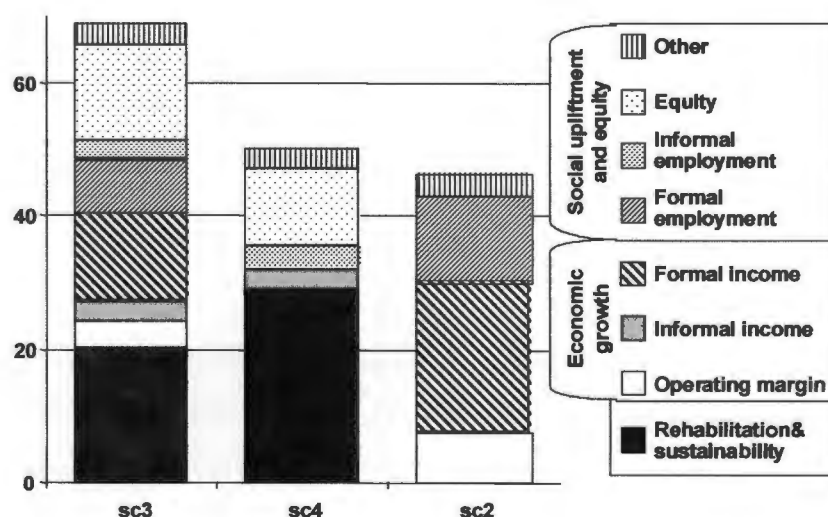


Figure 6.6. Relative contributions of criteria to overall scores of scenarios for Zone B.

6.2.3 Preferred scenario for Zone C

Using the consensus weights, the preferred option overall was Scenario 4 with community management of some game reserves, and some harvesting allowed on 20% of other conservation areas. Scenario 2 (converting 20% of commercial conservation land to rangelands) was second most preferred overall (due to informal employment and

earnings from resource harvesting), while Scenario 3 ranked third and Scenario 1 ranked fourth (Figure 6.7). There is a large gap in the overall score of Scenario 4 and those of the other three scenarios, and Scenario 1 and 3 are basically equivalent.

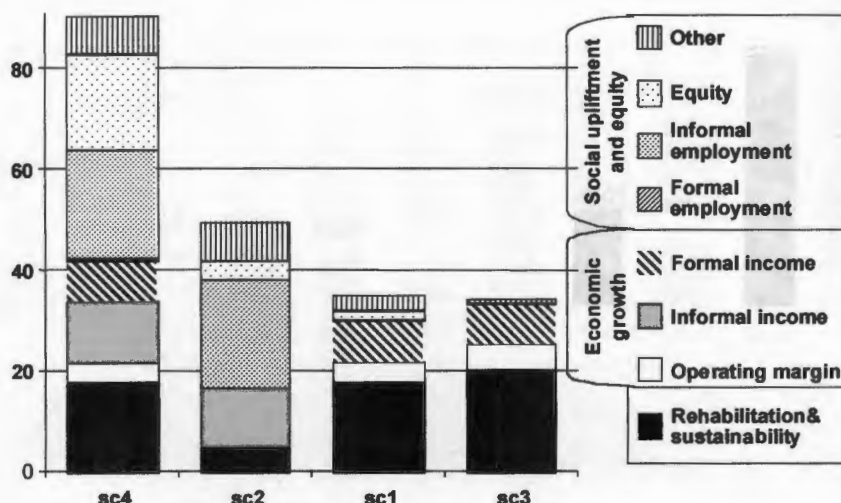


Figure 6.7. Relative contributions of criteria to overall scores of scenarios for Zone C.

6.2.4 Sensitivity to weight changes

Changing the weights of the three criteria groups within the range suggested in Table 6.4 made little difference to the preference order of the scenarios in all the zones and the preferred scenario remained the same. In Zone A, only when the ratio of weights for RS:EG:SE changed to (100:40:180), (100:151:60) or (100:108:108) did Scenario 5 become preferred to Scenario 7. Scenario 4, although second in overall score, never became preferred with changing weights at this level. Only if the ratio were (100:173:60) would Scenario 2 be the preferred option. In Zone B, only even more extreme weight changes would change the preferred option: for example, Scenario 4 would be preferred with a ratio of (255:90:100) and Scenario 2 would be preferred with a ratio of (80:293:100). In Zone C, even more extreme weight changes at this level are required to change the overall preferred option. Clearly, sensitivity to lower level weights also need to be tested, but are not illustrated here.

6.3 Discussion

6.3.1 Costs and benefits of preferred scenarios

A useful and intuitive formulation of a decision problem is that of specifying costs and benefits, which also helps to highlight areas where preferred scenarios could be improved before implementation. This may be done graphically by comparing 'value profiles' or graphs (e.g. Figure 6.5, Figure 6.6 and Figure 6.7) showing the relative contribution of each criterion to the overall score. This shows, for example, that for Zone A, Scenarios 4 and 5 have very similar overall scores, but for very different reasons.

Cost and benefits can also be explored by explicitly comparing two scenarios. For example, in Zone A, Scenario 7 - converting some presently afforested land to conservation under natural land-cover, was preferred to Scenario 2 -

converting some forestry land to irrigated tree crops, the latter being far more financially profitable than conservation. The benefits of preferring Scenario 7 to Scenario 2 stemmed from improvements in terrestrial and aquatic ecology and resources due to the gain of 2497 ha of formal conservation land, the gain of R67 500 informal income, 649 informal employment opportunities, improved equity and other social issues. The costs of choosing Scenario 7 stemmed from a loss of R34 mill operating margin (OM), R394 000 formal income, and 741 formal employment opportunities (Table 6.5). The MCDA process translated these attribute differences into value differences, the overall value difference being 41 'points', i.e. Scenario 7 is 41 points better than Scenario 2. Applying the appropriate weights the positive contributions to this difference came from RS (47.7), informal income (1.5); and from aggregated informal employment, land equity and 'other' social issues (9.3). The negative contributions or costs of choosing Scenario 7, came from OM (5.9), formal income (8.0) and formal employment (3.6) (Table 6.5).

Table 6.5. Benefits and costs of choosing Scenario 7 (preferred from rehabilitation and sustainability (RS) point of view), rather than Scenario 2 (preferred from economic growth (EG) point of view of). (SE=Social upliftment and equity).

Criterion Group	BENEFITS in choosing 7					COSTS in choosing 7		
	RS Aggregate	EG Informal income	SE Informal employment	Equity	Other	EG Operating margin	Formal income	SE Formal employment
Criterion								
Actual difference	2497 ha	R67 521	649 people			R33 993 516	R393 527	741 people
Difference in score	95.4	100.0	100.0	30.0	50	71.8	78.2	36.8
Weighted difference			0.24					0.76
Weighted difference			24.4					27.8
Weighted difference		0.08	0.4	0.39	0.17	0.41	0.51	0.43
Weighted difference		7.7	10.6	11.7	8.7	29.5	40.1	12.1
Weighted difference	0.5	0.2	0.3	0.3	0.3	0.2	0.2	0.3
Weighted difference	47.7	1.5	3.2	3.5	2.6	5.9	8.0	3.6
Contribution to score difference	47.7	1.5	9.3			13.9		3.6
Aggregate score difference	58.6					17.5		
Overall net benefit					41.0			

6.3.2 Implied trade-offs and indirect compensatory values

Viewing the global weights graphically (Figure 6.8) rather than numerically helps when identifying the swing weights and to make trade-offs more apparent to the group(s). However, this view does not explicitly inform the group(s) (nor the decision-maker) what their weights imply with regards to what they (and by implication, society) were 'willing to pay' for improvements on certain criteria. Further examination helps to reveal the exact nature of those trade-offs for the team to verify.

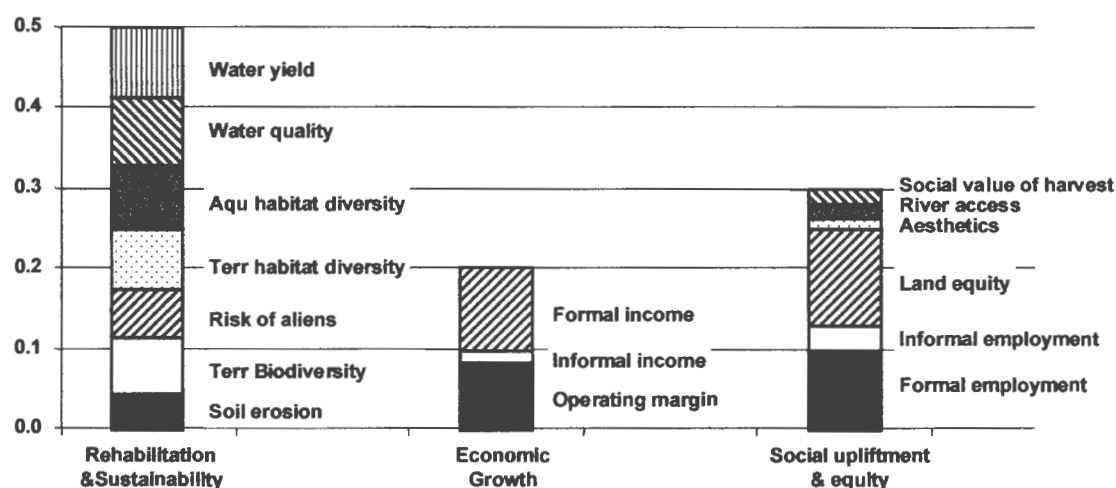


Figure 6.8. Relative contribution of criteria weight to the three criteria group weights in Zone A.

Besides determining the implied trade-offs between any pair of criteria, and given the current interest in the valuation techniques of resource and environmental economics, the differences in values given to scenarios and criteria weights can be further examined in at least two ways. Firstly, trade-offs can be used to determine the implied monetary values of issues not explicitly valued in the study. These explicit trade-offs should be of interest to the group(s) and can help to ensure the internal consistency of the problem. For example, the weights given to RS and OM can be used to estimate a monetary 'value' for the non-monetary criterion group RS. Secondly, differences between specific scenarios can be examined to look at the implied 'monetary benefit' of choosing one scenario over another.

As discussed in Section 5.3, a one value point change in v_1 is worth a w_1/w_2 value point change in v_2 . Where one criterion v_2 is linearly related to a quantitative attribute x_2 , the amount of the quantitative attribute that compensates for the change in v_1 can be determined with equation (5.9). For example, the change in OM (v_2) from x_2' to x_2'' that compensates for a change in soil erosion (SEr, v_1) can be determined. This is the implied trade-off value of SEr (C_{SEr}) in monetary terms (the number of Rands which compensates for, or is compensated by, a change in SEr):

$$C_{SEr} = (w_1/w_2) \times ((x_2' - x_2'') / (v_2(x_2') - v_2(x_2''))),$$

and in Zone A $w_1/w_2 = w_{SEr}/w_{OM} = 0.5538$. However, where x_2 is non-linearly related to value v_2 , the trade-off value will be different at each level of OM. Therefore, average values could be determined at appropriate intervals (e.g. at 20 point or 10 point intervals) (Table 6.6). In this case Rands OM was related to its value (v_{OM}) with a logarithmic value function:

$$v_{OM} = 100 \times (\ln x_2 - (\ln x_2)^{\min}) / ((\ln x_2)^{\max} - (\ln x_2)^{\min}),$$

and because the maximum and minimum Rands OM are constants we have:

$$v_{OM} = 100 \times (\ln x_2 - C) / (K - C)$$

Therefore an instantaneous rate of change between Rands and value could be determined (Table 6.6), where the relationship between Rands OM and value of OM is:

$$R_{OM} = e^{\left[\left(\frac{K-C}{100} \right) v_{OM} + C \right]}$$

The amount of OM compensating for a change in SEr from v'_{SEr} to v''_{SEr} , or $\hat{R}_{SEr}$ when the initial OM value is v'_{OM} is:

$$\hat{R}_{SEr} = e^C \left[e^{\left[\left(\frac{K-C}{100} \right) v'_{OM} \right]} - e^{\left[\left(\frac{K-C}{100} \right) v''_{OM} \right]} \right], \quad (6.1)$$

where
$$v''_{OM} = v'_{OM} - \left[(v'_{SEr} - v''_{SEr}) \times \frac{w_{SEr}}{w_{OM}} \right] \quad (6.2)$$

Thus, in the region of Scenario 5 (OM = 60.1), a 1 value point improvement in SEr, all other criteria held constant, compensates for a decrease in OM of from R199 000 to R180 000, depending on whether linear, 20 or 10 point intervals or instantaneous rates of change of Rands are used to calculate compensatory amounts (Table 6.6). In the region of Scenario 7 (OM = 28.25), a 1 value point improvement in SEr compensates for a decrease in OM of from R61 000 to R66 000 based on the piecewise linear and instantaneous calculations (Table 6.6) and R199 000 for the

linear version. Thus the monetary 'value' of changes in SEr depends on perspective: are we improving soil erosion in a situation where OM is high, or in a situation where OM is quite low? In the former case, we are willing to give up more OM for improvements in SEr than in the latter case. But all this actually means is that we are more ready to give up OM when we have more of it (the differences in compensating amount have no implied inference for the value of SEr itself).

Table 6.6. Value of 1 point change in OM. Inferred SEr value calculated linearly, at intervals and instantaneously.

Linear		Piecewise linear intervals						Instantaneous		
R'000 1 unit OM	R'000 1 unit SEr	20 value pt Interval	R'000 1 unit OM	R'000 1 unit SEr	10 value pt Interval	R'000 1 unit OM	R'000 1 unit SEr	Value Point	R'000 1 unit OM	R'000 1 unit SEr
R359.8	R 199.3	80-100	R 920.3	R 509.7	90-100	R 1 075.8	R 595.8	100	R 1 269.8	R 703.2
		60-80	R 465.1	R 257.6	80-90	R 764.8	R 423.5	90	R 902.7	R 499.9
		40-60	R 235.0	R 130.2	70-80	R 543.7	R 301.1	80	R 641.7	R 355.4
		20-40	R 118.8	R 65.8	60-70	R 386.5	R 214.0	70	R 456.2	R 252.6
		0-20	R 60.0	R 33.2	50-60	R 274.7	R 152.2	60	R 324.3	R 179.6
					40-50	R 195.3	R 108.2	50	R 230.5	R 127.7
					30-40	R 138.8	R 76.9	40	R 163.9	R 90.8
					20-30	R 98.7	R 54.7	30	R 116.5	R 64.5
					10-20	R 70.2	R 38.9	20	R 82.8	R 45.9
					0-10	R 49.9	R 27.6	10	R 58.9	R 32.6
								0	R 41.9	R 23.2

Table 6.7. Trade-offs between all criteria and OM and the benefit of choosing Scenario 7 over Scenario 5 (Zone A) translated into monetary terms. All weights are given in Table 6.3.

monetary terms. All weights are given in Table 6.5.						
	Criterion	w_i	Δv_j	Benefit of Sc7 over Sc5	Benefit of Sc7 over Sc5	Benefit of Sc7 over Sc5
		w _{OM}	Sc7-Sc5	R'000 Linear	R'000 20 pt intervals	R'000 Instantaneous
RS	BENEFITS					
	Soil erosion (SEr)	0.554	100	R 19 928	R 10 982	R 8 087
	Terrestrial biodiversity	0.831	80	R 23 913	R 13 178	R 8 540
	Risk of exotic flora, fauna	0.738	80	R 21 256	R 11 714	R 8 257
	Terrestrial habitat diversity	0.923	80	R 26 570	R 14 642	R 8 760
	Aquatic habitat diversity	0.979	80	R 28 183	R 15 531	R 8 868
	Water quality	0.979	50	R 17 615	R 9 707	R 7 734
	Catchment water yield	1.088	100	R 39 143	R 21 571	R 9 294
EG	Informal income	0.188	46.96	R 3 171	R 1 747	R 2 474
	Informal employment	0.387	57.69	R 8 042	R 4 432	-R 5 083
SE	Aesthetics	0.173	50	R 3 119	R 1 719	-R 2 439
	Social value of harvest	0.231	60	R 4 991	R 2 750	-R 3 592
EG	COSTS					
	Operating margin	1	-31.8	R 6 310	R 6 310	R 6 310
	Formal income	1.249	-57.9	R 25 158	R 14 349	R 103 022
	Formal employment	1.202	-100	R 43 244	R 23 831	R 565 902
	Land equity	1.430	-30	R 15 439	R 8 508	R 31 667
	River access	0.231	-60	R 4 991	R 2 750	R 5 766

Therefore, although the compensatory relationships between criteria hold for any scenario, the actual trade-off value between two criteria would depend on which scenario were being considered, because of the non-linear relationship of OM value to Rands. It may be more useful to assess the effective trade-offs made between pairs of scenarios. I have therefore called these values *indirect compensatory values*. The amount of Rands OM which will compensate for the particular change in value of each criterion when moving from Scenario 7 to Scenario 5, calculated linearly,

with intervals or instantaneously are shown in Table 6.7. The compensatory values calculated when moving from various other scenarios to Scenario 7 are shown in Table 6.8 (instantaneous) and Table 6.9 (linear).

Table 6.8. Compensating OM value for gains or losses in each criterion when changing from Scenario X to Scenario 7, using instantaneous rates of change. Δv_j refers to the value point gain/loss for each criterion. ICV Δv_j to OM refers to the equivalent worth or indirect compensatory value of Δv_j (i.e. the entire change) in terms of OM.

Criterion	$\frac{w_i}{w_{OM}}$	Scenario 2 to Scenario 7		Scenario 5 to Scenario 7		Scenario 1 (SQ) to Scenario 7	
		Δv_j	ICV Δv_j to OM R'mill	Δv_j	ICV Δv_j to OM R'mill	Δv_j	ICV Δv_j to OM R'mill
Soil erosion	0.5538	+100	R31.59	+100	R8.09	+100	R1.539
Biodiversity	0.8307	+100	R35.02	+80	R8.54	+100	R1.706
Risk of exotic invasion	0.7384	+90	R33.36	+80	R8.26	+100	R1.667
Terrestrial habitat diversity	0.9230	+100	R35.61	+80	R8.76	+100	R1.735
Aquatic Habitat Diversity	0.9790	+90	R35.37	+80	R8.87	+100	R1.749
Water quality	0.9790	+100	R35.89	+50	R7.73	+50	R1.472
Catchment water yield	1.0878	+90	R35.89	+100	R9.29	+80	R1.720
Informal Income	0.1876	+100	R17.60	+47.0	R2.47	+68.0	R0.639
Informal Employment	0.3874	+100	R27.29	+57.7	R5.08	+76.9	R1.157
Land Equity	1.4303	+30	R28.60	-30	R31.67	+60	R1.716
Aesthetics, greenspace	0.1734	+50	R9.53	+50	R2.44	+60	R0.542
River access	0.2312	+40	R10.07	-60	R5.77	+10	R0.137
Social value of harvesting	0.2312	+60	R14.03	+60	R3.59	+100	R0.989
Operating Margin	1.0000	-71.75	R33.99	-31.82	R6.31	+16.8	R1.403
Formal Income	1.2491	-78.22	R1 006.47	-57.93	R103.02	-31.8	R5.229
Formal Employment	1.2018	-36.77	R130.89	-100	R565.90	-2.62	R0.206

Note that an improvement of 100 points in SEr may be 'worth' between R31.6 mill (changing from Scenario 2 to Scenario 7) and R1.5 mill (changing from Scenario 1 to Scenario 7), all other criteria being held constant (Table 6.8). This is because Scenario 2 has a very high level of OM (the maximum with a value of 100), and Scenario 1 a low level (with a value of 11.5). The calculations for moving from Scenario 2 to Scenario 7 are as follows (from equations (6.1) and (6.2), weights in Table 6.8 and values in Appendix 6):

$$v_{OM}'' = 100 - [(100 - 0) \times 0.5538]$$

$$\hat{R}_{SEr} = 1\,226\,402.45 \left[e^{\left[\left(\frac{17.43 - 14.02}{100} \right) 100 \right]} - e^{\left[\left(\frac{17.43 - 14.02}{100} \right) 44.62 \right]} \right] = 31.5 \times 10^6$$

Also note that if we were to 'give up' 100 points of SEr from a particular scenario (e.g. from Scenario 7) to any other scenario (e.g. to Scenario 1 or 2 or 5), the value of OM which would keep utility constant, all other criteria being kept constant, would be R18 mill in all cases (Table 6.9). If we had chosen to have a linear relationship between Rands OM and value of OM, a change of 100 value points of SEr would have the same 'value' (i.e. R19 mill) regardless of which scenario we were changing from. Therefore a linear relationship may have allowed a slightly easier interpretation of results. However, the non-linear relationship of OM to value was chosen by the participants as a better reflection of reality than a linear relationship (and conforms to the notion of decreasing marginal utility of money). The actual relationship (i.e. logarithms) was chosen for convenience of calculation, as we did not formally assess mid-points. As shown by Stewart (1995) assuming a linear value function when the 'real' relationship is non-linear, can lead to fairly severe errors in overall ratings.

Table 6.9. Value of changes from Scenario X to Scenario 7, with a linear value function for OM. Δv_j refers to the value point gain/loss for each criterion. ICV Δv_j to OM refers to the equivalent worth or indirect compensatory value of Δv_j (i.e. the entire change) in terms of OM.

Criterion	$\frac{W_j}{WOM}$	Scenario 2 to Scenario 7		Scenario 5 to Scenario 7		Scenario 1 (SQ) to Scenario 7	
		Δv_j	ICV Δv_j to OM R'000	Δv_j	ICV Δv_j to OM R'000	Δv_j	ICV Δv_j to OM R'000
Soil erosion	0.5538	+100	R 19 928	+100	R 19 928	+100	R 19 928
Biodiversity	0.8307	+100	R 29 891	+80	R 23 913	+100	R 29 891
Risk of exotic invasion	0.7384	+90	R 23 913	+80	R 21 256	+100	R 26 570
Terrestrial habitat diversity	0.9230	+100	R 33 213	+80	R 26 570	+100	R 33 213
Aquatic Habitat Diversity	0.9790	+90	R 31 706	+80	R 28 183	+100	R 35 229
Water quality	0.9790	+100	R 35 229	+50	R 17 615	+50	R 17 615
Catchment water yield	1.0878	+90	R 35 229	+100	R 39 143	+80	R 31 315
Informal Income	0.1876	+100	R 6 752	+46.96	R 3 171	+67.9	R 4 584
Informal Employment	0.3874	+100	R 13 940	+57.69	R 8 042	+76.9	R 10 723
Land Equity	1.4303	+30	R 15 440	-30	R 15 440	+60	R 30 879
Aesthetics, greenspace	0.1734	+50	R 3 119	+50	R 3 119	+60	R 3 743
River access	0.2312	+40	R 3 327	-60	R 4 991	+10	R 832
Social value of harvesting	0.2312	+60	R 4 991	+60	R 4 991	+100	R 8 318
Operating Margin	1.0000	-71.75	R 33 994	-31.82	R 6 310	+16.8	R 1 403
Formal Income	1.2491	-78.22	R 35 155	-57.93	R 26 038	-31.8	R 14 309 790
Formal Employment	1.2018	-36.77	R 15 901	-100	R 43 244	-2.62	R 1 135

Whatever approach used, the wide range of 'values' of SEr given by the indirect compensatory values are informative and interesting. However, they might cause confusion to workshop participants and even more so to other stakeholders not directly participating. It would therefore be useful to develop dynamic graphical approaches which can be used within workshops to help to interpret these trade-offs, calculate the indirect compensatory values and illustrate the changes when moving from one scenario to another. For example, Figure 6.9 shows that if one were to change from scenario A to A' in terms of SEr (an increase of 60 points, when weighted = 12 points), utility would be kept constant with a loss of 12 points OM (weighted), which is the same as a loss of R3.7 mill. However, moving from B to B' (the same change in value of SEr), is compensated by a loss of R7.9 mill (also 12 weighted points OM).

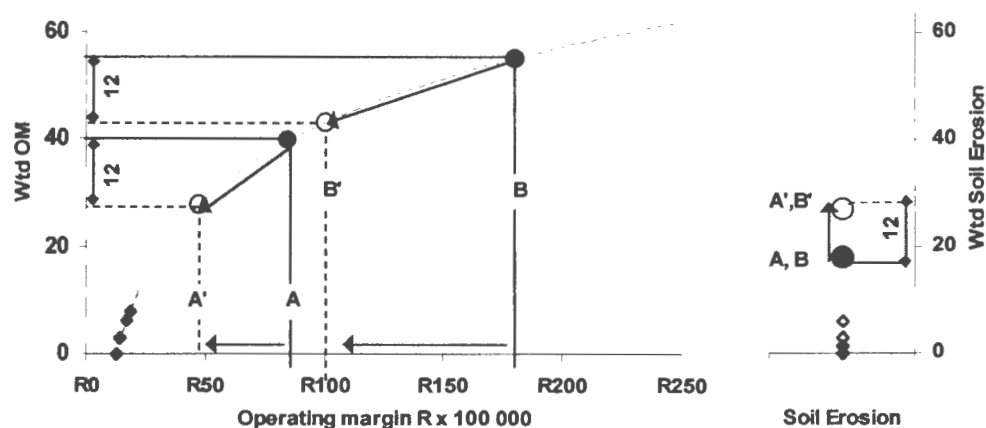


Figure 6.9. The trade-off between SEr and OM moving from A to A' and from B to B'. The same change in value of SEr is compensated for by different Rand OM. The value change on both scales for both cases is 12.

6.3.3 Sources of value

In an environmental and resource economics formulation, the direct-, indirect-, and non-use monetary value of the land- and water-use scenarios would need to be determined. This division of sources of value into direct, indirect, and non-use is a useful typology and may help to ensure that all types of value are considered.

Some *direct-use values* were included explicitly in this study, in particular to include contributions to people's livelihoods from informal sources. For example, income and employment from harvesting secondary and natural resources, and from small-scale irrigation on garden plots were included. *Indirect-use value* was included in the RS criteria, in particular, soil erosion, water yield, and water quality which indirectly indicate the extent of 'ecosystem services' supplied by the different scenarios. Most of the SE criteria include some aspects of indirect-use. Resource and environmental economics tools to determine indirect-use values would require intensive data collection about ecosystem services (e.g. costs of supplying services such as flow regulation, water quality treatment, replacement of topsoil), and the implications of degraded environments on profits from various land-uses. A production function approach might be used where the effect of different levels of an environmental input are modelled in terms of various economic outputs (e.g. crop yields).

Existence value stems from the value people gain from knowing that a particular ecosystem, habitat, or species exists and is usually determined using contingent valuation surveys, asking questions about WTP to conserve a particular environment, or accept in compensation for its loss. WTP for conservation of the upper catchment among the general public in the Sand river subcatchment would be likely to be low if determined in this way. However, other types of value in terms of WTP could derive from two sources. Firstly, the upper catchment forms part of the escarpment, many areas of which have high tourism value. The area could develop into a tourist destination to realise use value, and the increased awareness would increase existence value (Scenario 7 captured some of these ideas). Secondly, owners of the conservation areas in Zone C might have a high WTP for changed land-use in the upper catchment, because of the effects of upstream use on their area. This has already been demonstrated in their partial funding of the project of which this study formed a part, and their willingness to litigate regarding forestry and irrigation practice in Zones A and B. Existence value plays a role in some of the RS (e.g. species richness) and SE criteria (e.g. land equity; scoring on this criterion was based more on the perception of land being accessible than on the redistribution of land). The flow of sources of value and associated criteria in the Sand river subcatchment is illustrated in Figure 6.10.

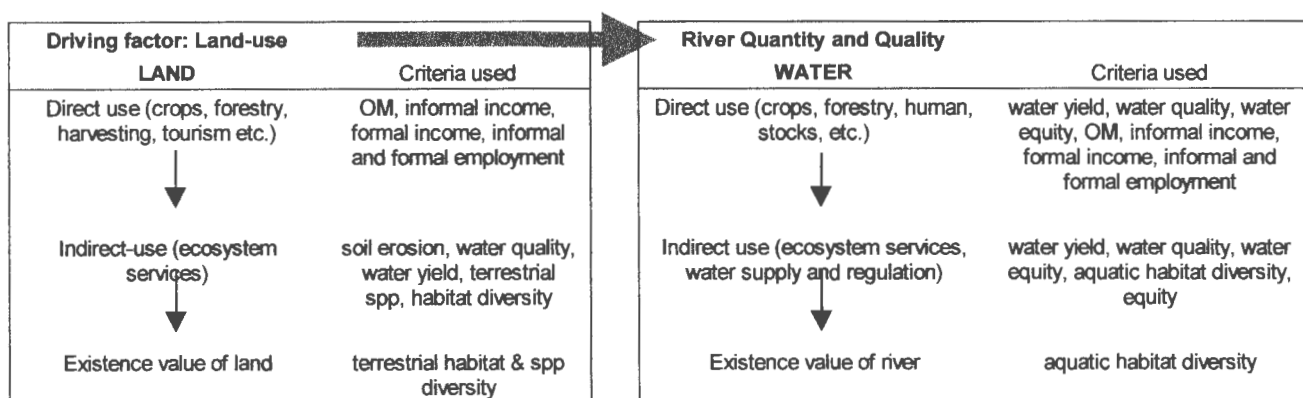


Figure 6.10. Sources of value for the catchment land-use scenarios (spp =species).

In summary, we elicited scores, as representations of the relative *preference* for levels of criteria or of performance measured according to criteria, without requiring a reference to ability or WTP. Compensatory amounts of OM for

various changes in levels of criteria were calculated, but because of the non-linear OM value function, these were dependent on scenario. Subjecting the implied compensation OM amounts to 'reality checks' might be worthwhile in certain situations as it might improve the accuracy of trade-off representations (however, the explicit introduction of OM compensatory amounts may well introduce conflict into a group where there were none before). Note that the actual costs of implementing the scenarios (e.g. the costs of clearing plantations and rehabilitating the cleared land) were not included in the analyses.

However, whether translated into money or whether left as value scales, the relative values are important rather than the absolute values. The relative values clearly underline the policy directions, and would be highly unlikely to change with realistic weight changes as seen in the sensitivity analysis.

6.3.4 Practical issues

Due to the time and funding constraints most of the results presented here could not be reported back to the project team for feedback and refinement after the workshop where final scores and weights were obtained. This allowed no interactive sensitivity analysis, or review of values. This shortcoming might have been avoided if it had been possible to use software such as VISA (1995) during the workshops, rather than only in-between.

In Zones B and C the number of scenarios could have affected results and their interpretation. In Zone B, only three scenarios were considered (as decided by the project team), and it was almost inevitable that a 'middle ground' scenario would be rated as 'best' (i.e. it was by its very nature a compromise). The division of the catchment into three zones was essential from a practical and management point of view. Each of the preferred zonal scenarios was preferred from a different point of view (RS for Zone A, EG for Zone B, and SE for Zone C, (Figure 6.4). The project team felt that, taken together, the scenarios were feasible and would satisfy the objectives of integrated catchment management. However, combinations of scenarios across zones could have been examined in more depth if time had permitted.

Most of the project team found the use of thermometer scales fairly intuitive and were comfortable indicating the relative value of scenarios on the scale. In some cases, verbal cues or definitions associated with scores would have been useful (e.g. poor, very good etc.). The value measuring techniques could have been adapted for less numerate participants by using beans or stones to indicate scores (using some of the techniques developed in participatory rural appraisal). However, this adaptation was not necessary for this stage of the project, where general public participation was not required.

The preferred scenarios, or indicated preferred directions for change, are currently being implemented in the Sand River catchment and the scenarios and criteria will be integrated into the initial stages of the process to implement the new NWA (Act 36, RSA 1998) (i.e. designating a management class etc.). A shortcoming of the analysis was the lack of information on the actual financial costs of implementation. For example, the harvesting of the plantation forests in the upper catchment in order to implement Scenario 7 is a cost borne by the Dept. of Water Affairs and

Forestry and was not included in the analysis. Another issue was that although this study was supposed to be a Phase 1 analysis with subsequent more detailed examination of the 'directions of preference' indicated, the subsequent analysis was not undertaken. In the particular case of Scenario 7 in Zone A where half of plantation forests were to be removed, this preference has led to a recommendation that *all* plantation forests should be removed in this catchment. There is no *a priori* reason to believe from our analysis that removing *all* forests would necessarily be preferred to removing half of the forests.

6.4 Summary and conclusions

Various land- and water-use combination scenarios for the Sand river catchment were examined using AD-SMARTx. These were evaluated on the basis of ecological, economic and social criteria within the three zones of the catchment. The preferred scenarios for each zone corresponded to the preferred rehabilitation and sustainability scenario, the preferred social upliftment and equity scenario, and the preferred economic growth scenario in Zones A, B and C respectively. Results were not sensitive to the weights chosen. As an illustration, indirect compensatory values were calculated between OM and SEr. This showed that the non-linear relationship between Rands of OM and the value of OM influenced the implied trade-offs. Thus, if changing from a scenario with a high level of OM one would be willing to give up more in order to improve SEr than if one were at a low level of OM. Though complicating the interpretation, this was a more realistic interpretation than that implied by a linear OM value function. The criteria used could be related to use and non-use values as used in EE analyses.

The decision-support was perceived to be particularly useful for problem structuring and for the integration of different types of information. The SMARTx and scenario developing techniques were fairly simple and intuitive, and closely resembled frequently-used common-sense approaches (e.g. ranking, creating indices). The participants appreciated that there was an advantage in applying a more formal approach to improve the theoretical basis of their recommendations and help to avoid some pitfalls of less rigorous approaches (e.g. interpreting ranks as scores, and designating weights which do not relate to the range of consequences being considered).

The following chapter describes an application of EE to the valuation of river health scenarios, a context comparable to the case discussed in this chapter. Further SMARTx and EE cases are discussed in Chapter 8, which took place within quite different contexts to these river health cases. The application of SMARTx and EE to the six cases and the practical and philosophical implications are discussed in Chapter 9, with emphasis placed on the Sand river case and the case in the following chapter.

Appendix 6A. Consequences of scenarios, scores and weights⁹.

Zone A

Zone A

	Criterion Group: Rehabilitation and Sustainability									Aggreg
	Terrestrial ecology				Aggreg	Aquatic ecology			Aggreg	
	Soil erosion	Bio-diversity	Risk of exotic invasion	Terrestrial habitat diversity		Aquatic habitat diversity	Water quality	Catchment water yield *		
Sc1 SQ	0	0	0	0	0	0	50	20	23.2	11.6
Sc8-25%	40	10	10	10	15.5	90	90	35	70.4	42.9
Sc2-PI	0	0	10	0	2.4	10	0	10	6.8	4.6
Sc3-DA	0	20	20	20	16.4	50	40	40	43.2	29.8
Sc4-WL	90	90	30	90	75.5	80	90	70	79.6	77.6
Sc5-AI	0	20	20	20	16.4	20	50	0	22.5	19.4
Sc6-R+GP	0	20	0	20	11.5	20	0	30	17.1	14.3
Sc7-CCon	100	100	100	100	100	100	100	100	100	100
Weights	0.18	0.27	0.24	0.31		0.32	0.32	0.36		
Weights					0.5				0.5	
Group weight										0.5

* Estimated - hydrology unavailable

	Criterion Group: Economic Growth						Aggreg
	Operating margin		Informal Earnings		Formal Income		
	Rand	Score 0-100	Rand	Score 0-100	Rand	Score 0-100	
Sc1 SQ	1 812 728	11.45	65 303	32.11	346 727	53.6	23.7
Sc8-25%	1 414 817	4.19	57 427	17.16	261 096	31.3	19.1
Sc2-PI	37 209 420	100	49 551	0	624 925	100	92.3
Sc3-DA	1 664 694	8.95	75 102	48.37	224 198	19.3	17.3
Sc4-WL	1 226 402	0	81 580	57.99	179 447	1.8	5.4
Sc5-AI	9 526 301	60.07	78 185	53.04	482 982	79.7	69.6
Sc6-R+GP	4 824 824	40.14	71 787	43.12	175 465	0	19.8
Sc7-CCon	3 215 904	28.25	117 072	100	231 398	21.8	30.5
Weights		0.4104		0.077		0.513	
Group weight							0.2

Criterion Group: Social upliftment and Equity											Aggreg
	Formally employed	Employment		Aggreg		Equity	Aesthetics greenspace	Other River access	Social value of harvesting	Aggreg	
		Score 0-100	Informally employed	Score 0-100	Total	Land Equity					
Sc1 SQ	271.9	2.6	627.7	23.1	7.6	0	40	30	0	21.8	7.1
Sc8-25%	271.9	2.6	552.8	11.6	4.8	50	80	40	30	47.3	29.9
Sc2-PI	959.6	36.8	477.9	0	27.8	30	50	0	40	28.2	28.7
Sc3-DA	984.6	38.0	702.6	34.6	37.2	100	10	70	40	42.7	62.7
Sc4-WL	235.5	0.8	788.8	47.9	12.3	40	50	60	90	68.2	32.9
Sc5-AI	2233	100	752.6	42.3	85.9	90	50	100	40	64.6	83.8
Sc6-R+GP	984.6	38.0	690.1	32.7	36.7	70	0	50	50	36.4	49.7
Sc7-Ccon	219	0	1 127.1	100	24.4	60	100	40	100	78.2	47.7
Weights		0.756		0.244		1	0.27	0.36	0.36		
Weights					0.435	0.39				0.17	
Group weight											0.3

	Rehabilitation and Sustainability	Economic Growth	Social upliftment and Equity	OVERALL
Sc1 SQ	11.6	23.7	7.1	14.9
Sc8-25%	42.9	19.1	29.9	34.2
Sc2-PI	4.6	92.3	28.7	29.4
Sc3-DA	29.8	17.3	62.7	37.2
Sc4-WL	77.6	5.4	32.9	49.7
Sc5-AI	19.4	69.6	83.8	48.8
Sc6-R+GP	14.3	19.8	49.7	26.0
Sc7-Ccon	100	30.5	47.7	70.4
Weights	0.5	0.2	0.3	

⁹ See Text and Table 1 for descriptions of the scenarios. Where value functions were used to translate data to 0-100 scales, both the estimated data and the resulting scores are included. Weights were rescaled to sum to one.

Zone B

Criterion Group: Rehabilitation and Sustainability									Aggreg	
Terrestrial ecology					Aggreg	Aquatic ecology				Aggreg
Soil erosion	Bio-diversity	Risk of exotic invasion	Terr habitat diversity			Aquatic habitat diversity	Water quality	Catchment water yield *		
Sc1-SQ1998										
Sc2-Max	0	0	0	0	0	0	0	0	0	
Sc3-Mid	85	100	25	100	84	90	50	70	54	
Sc4-SQ2010	100	95	100	95	78	100	100	100	100	
Weights	0.42	0.33	0.13	0.13		0.32	0.32	0.36		
Weights					0.5				0.5	
Group weight										0.3

Criterion Group: Economic Growth							Aggr
Operating margin Rand	Operating Margin 0-100	Informal Earnings Rand	Informal Earnings 0-100	Formal Income Rand	Formal Income 0-100		
Sc1-SQ1998							
Sc2-Max	146 764 712	100	2 315 999	0	2 520 835	100	91
Sc3-Mid	126 474 335	55.4	2 462 001	89.04	1 909 203	58.16	60
Sc4-SQ2010	105 118 599	0	2 480 598	100	1 297 398	0	9
Weights		0.23		0.0912		0.678	
Group weight							0.333

Criterion Group: Social upliftment and Equity													Aggr
Formally employed Total	Employment		Informally employed Total	Informally employed 0-100	Aggr	Equity Land Equity	Aesthetics greenspace	River access	Other Infra- structure	Crime	Social value of harvesting	Aggr	
	Formally employed 0-100												
Sc1-SQ1998													
Sc2-Max	28 920	100	22 189	0	79	0	30	0	100	100	0	52	43
Sc3-Mid	26 559	62.64	23 629	85.64	68	100	70	70	60	60	20	50	77
Sc4-SQ2010	22 601	0	23 870	100	21	80	80	100	0	0	100	44	48
Weights		0.79		0.21		1	0.05	0.13	0.32	0.19	0.32		
Weights					0.43	0.39						0.17	
Group weight													0.37

* Estimated as hydrology unavailable

Rehabilitation and Sustainability		Economic Growth	Social upliftment and Equity	OVERALL
Sc1-SQ1998				
Sc2-Max	0	91	43	46.4
Sc3-Mid	69	60	77	69.2
Sc4-SQ2010	99	9	48	50.2
Weights	0.3	0.333	0.37	

Zone C

Criterion Group: Rehabilitation and Sustainability										Agg
	Terrestrial ecology				Aggr	Aquatic ecology			Aggr	
	Soil erosion	Terr species diversity	Exotic invasion	Terr habitat diversity		Aqu habitat diversity	Water quality	Catch water yield *		
Sc1-SQu	100	0	100	0	66	100	100	100	100	83
Sc2-WL	0	100	0	100	35	0	0	00	0	17
Sc3-CCon	70	0	70	10	48	90	95	90	92	70
Sc4-CCon+NRH	70	0	70	10	48	90	95	90	92	70
Weights	0.31	0.17	0.34	0.17		0.32	0.32	0.36		
Weights				0.5				0.5		
Group weight										0.25

	Criterion Group: Economic Growth						Agg
	Operating margin	Operating Margin	Informal Earnings	Informal Earnings	Formal Income	Salary	
	Rand	0-100	Rand	0-100	Rand	0-100	
Sc1-SQu	76 492 296	100	0	0	1 334 143	95.5	49
Sc2-WL	63 525 794	0	356 519	100	1 111 643	0	50
Sc3-CCon	75 694 862	94.4	0	0	1 345 733	100	50
Sc4-CCon+NRH	75 694 862	94.4	356 202	99.9	1 345 733	100	99
Weights		0.18		0.49		0.33	
Group weight							0.25

Criterion Group: Social upliftment and Equity											Agg
	Employment		Agg	Equity Land Equity	Aesthetics green-space	Other River access	Social value harvest	Agg			
	Formally employed Total	Formally employed 0-100							Informally employed Total	Informally employed 0-100	
Sc1-SQu	1 668	79.3	0	0	2	0	100	0	0	9	2
Sc2-WL	1 612	0	3460	100	98	20	0	100	90	86	65
Sc3-CCon	1 682	100	0	0	2	10	100	30	20	32	10
Sc4-CCon+NRH	1 682	100	3425	98.9	99	100	100	60	100	82	96
Weights	0.02		0.98			1	0.1	0.5	0.45		
Weights			0.44			0.39			0.17		
Group weight											0.5

* Estimated as hydrology unavailable

	Rehabilitation and Sustainability	Economic Growth	Social upliftment and Equity	OVERALL
Sc1-SQu	83	49	2	34.1
Sc2-WL	17	50	65	49.4
Sc3-CCon	70	50	10	34.9
Sc4-CCon+NRH	70	99	96	90.3
Weights	0.25	0.25	0.5	

7. Estimating changes in the tourism value of a river with EE

The case reported in this chapter was part of a broader study (Turpie *et al.* 2000, Mander *et al.* 2001) to investigate whether the ecological, social and economic effects of different river qualities could be measured in economic terms for inclusion in a decision-making process to choose a desired river quality. As with the case in Chapter 6, this work took place in the context of the National Water Act (NWA, Act 36, RSA 1998). According to the NWA, all rivers need to be placed into a management class indicating levels of protection or development, and thus river quality. A “Reserve” then needs to be determined which is the amount of water required for that river quality, together with basic human needs. The river management class and Reserve need to be chosen so as to balance ecological, social and economic considerations. The particular focus of this study was that of the economic value of environmental goods and services, as these are often ignored, resulting in environmentally detrimental decisions, which also have negative socio-economic impacts. Of particular relevance to the comparisons in Chapters 9 and 10 are Figure 7.2 which summarises the study design, the consumer surplus and consumer behaviour valuation results in Section 7.2.1, the comparison of conjoint models in Section 7.2.2, and the discussion in Section 7.3.

Conventionally, different values would be determined as part of a CBA. In this context, the total economic value of different river quality scenarios would need to be estimated. This would include analysis of agricultural and industrial outputs (direct consumptive use values), non-consumptive use (e.g. tourism), indirect use (e.g. erosion control) and non-use values (existence value), and importantly, in a developing world setting, subsistence use (e.g. fishing and harvesting of reeds) (see Table 3.1), and changes in these values with different scenarios. Specifically, the potential effects of changes in river quality on the tourism value of the Kruger National Park (KNP) were investigated in this case study. This case study, therefore, has a very similar context to that of the case discussed in the previous chapter, and the two cases are comparable in terms of their overall objectives (choosing a river quality scenario) as is further discussed in Chapter 9. The travel cost method (TCM – a revealed preference approach), contingent behaviour valuation (CBV – a stated preference approach based on contingent valuation, CVM) and conjoint analysis (CA-a stated preference approach) were used.

The Sand river, the setting for the previous case study, is a tributary of the Sabie river some of which flows through the KNP, and which is in turn a tributary of the Crocodile river which forms the southern border of the KNP. Many of the issues arising in the Sand river subcatchment discussed in the previous chapter also arise in the Crocodile catchment as a whole, and are not reiterated here.

7.1 Methods

The study focused on the sections of the Crocodile river and its tributaries which pass through the KNP. The main river forms the southern border of the KNP, and several tributaries flow through the southern part of the park,

including the Sand (the setting for the previous case study), a tributary of the Sabie, in turn a tributary of the Crocodile (Figure 7.1). At the time, a detailed Reserve assessment was to take place for the Crocodile catchment and it was envisaged that this study would contribute to that. The detailed Reserve determination did not take place, but various economic values other than those from tourism were investigated in Mander *et al.* (2001) and some are reported here, while part of this chapter was published in Turpie and Joubert (2001).

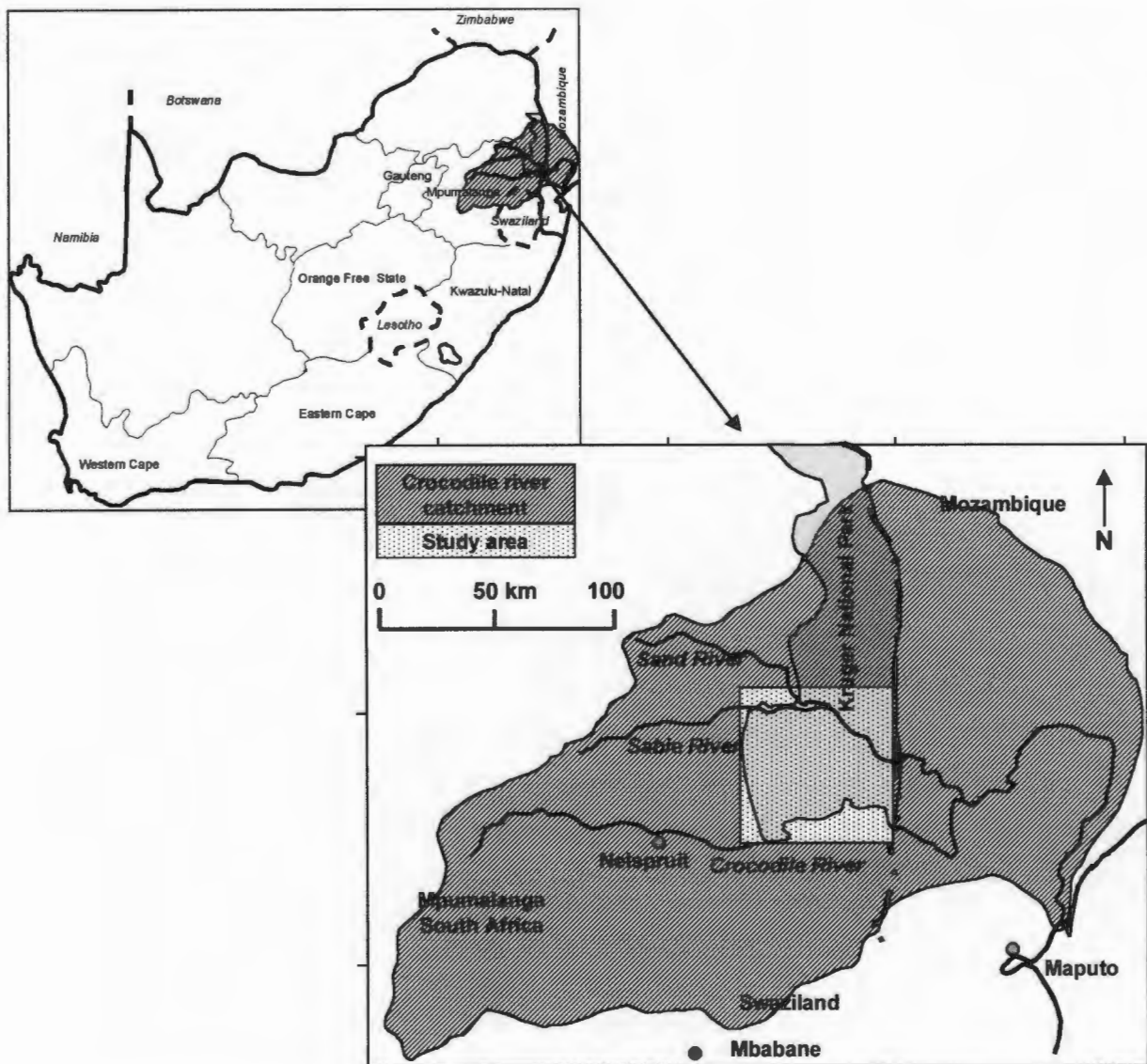


Figure 7.1. Map of the study area.

A combination of methods was used to deal with the complex issues involved. These methods were all combined in a single questionnaire which was administered during May to June 2000 to 183 tourists in KNP. Firstly, the tourists' actual travel costs were found, then potential changes in the length of their stay with two changes in river quality, then preferences for different river scenarios made up of several attributes (Figure 7.2). The TCM was used to determine actual costs and consumer surplus (Section 5.2.5). CBV was used to determine changes in lengths of stays given two changes in river quality (Section 5.2.1), and CA was used to determine preference for combinations of river attribute levels which reflected river quality (Section 5.2.3). The contingent and conjoint valuations were linked through the use of a common *status quo*, ideal and worst case scenarios in the two questions, thus allowing changes in river attributes to be linked to changes in length of visit. Ultimately, by linking these results to the travel

cost results (which also reflected the *status quo*), changes in tourism value could be predicted based on changes in river attributes (Figure 7.2). Cameron (1992) explores the linking of TCM and CBV results, but uses far more sophisticated econometric models than those reported here. The results of the initial study and model are examined and sensitivity to model specification is tested. A preliminary examination of different CA model specifications was performed by Cohen (2001). I have repeated this work with some modifications.

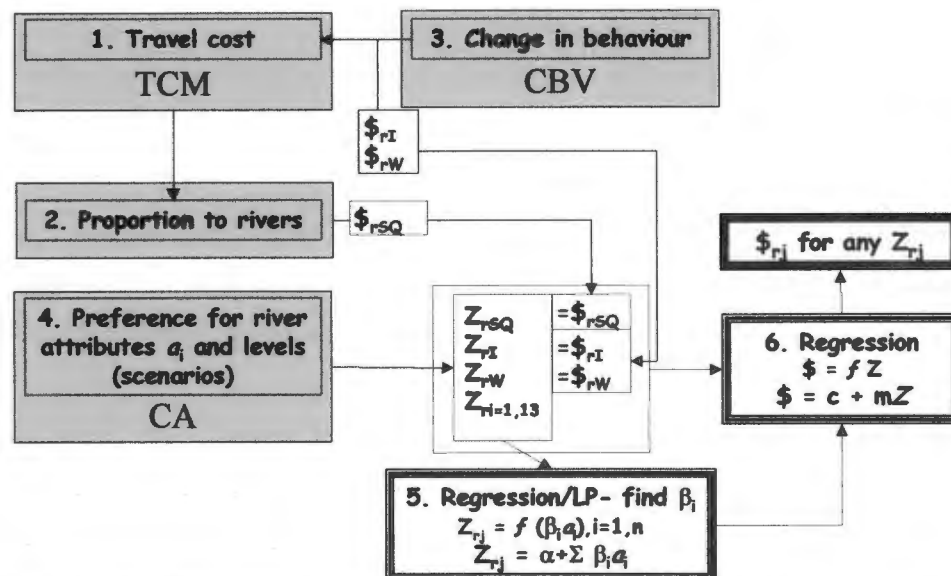


Figure 7.2. Links between the different sections of the questionnaire and the analyses to find the effects of river quality on tourism value. $\$_{rSQ}$ refers to the tourism value of the river (r) with *status quo* quality, as reflected by TC and $\$_{rI}$ and $\$_{rW}$ to travel cost for the ideal and worst case scenarios as determined by the CVM question. The Z_{rN} refer to the scenario scores from the CA question. The shaded blocks indicate sections of the questionnaire, the blocks with double borders indicate subsequent analysis, and the block with the bold border indicates the final output, i.e. a monetary value for tourism for any scenario's Z_{rj} score derived from any combination of the four river attributes.

7.1.1 The travel cost method (TCM)

The TCM is conventionally used to investigate the monetary value of an environmental amenity. Besides eliciting the actual expenditure to get to the destination, and at the destination, TCM determines the 'total WTP' which is equal to the price paid plus consumer surplus (Section 5.2.5, Figure 5.2, Figure 7.3). Visitor numbers per population of the zone of origin were regressed against travel cost (TC) using the log form:

$$\ln(\text{visits per inhabitant}) = \alpha + \beta TC \quad (7.1)$$

Then, with α and β , the number of visits were estimated with a series of hypothetical additional costs P , allowing a demand curve to be set up:

$$\ln(\text{visits per inhabitant}) = \hat{\alpha} + \hat{\beta} \times (TC + P),$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the estimates the new regression. Then assuming the additional cost P is zero, the surplus currently enjoyed (the area under the demand curve) is calculated to determine the consumer surplus for each zone (see Section 5.2.5 for the derivation):

$$\text{Consumer surplus Zone A} = - \frac{e^{\hat{\alpha} + \hat{\beta} TC_A}}{\hat{\beta}}$$

The consumer surpluses from each zone are added to give total consumer surplus (see Figure 5.2, Figure 7.3).

Tourists were asked their place of origin for the trip, the amount they had spent to get to the destination, and the amount they had spent at the destination. We did not calculate the opportunity costs of time. They were also asked to estimate what proportion of the reason for going on the trip was attributable to KNP to account for multiple destination trips. To be able to apportion to rivers the appropriate amount of the tourists' travel cost, they were asked:

- (a) how much of their total time at KNP they spent alongside rivers
- (b) to trace the routes they had travelled on a map (which allowed us to calculate how much of their travel distance was spent along rivers), and
- (c) how much of their total enjoyment of KNP came from rivers (compared to two other habitats, namely bushveld and waterholes).

In addition, the proportion of roads that were laid next to rivers was calculated, as a measure of park management's estimate of the relative value contributed by river habitats, and for comparison with tourists' time spent at rivers.

Note that TCM can also be used to assess the *change in utility of visitors*, in which case the presence of substitutes needs to be considered (if they can go elsewhere and enjoy themselves as much for similar costs, then they have not lost utility). Here we were interested in the *contribution of a river to the value of the KNP* (and to the catchment). If visitors' behaviour changed due to river quality (e.g. they went elsewhere, to another park with rivers of better quality), that value would be lost to the KNP or catchment. Therefore in this case, we are not concerned with substitutes.

7.1.2 Contingent behaviour valuation (CBV)

The contingent behaviour valuation question (also called contingent activity) asked the tourist if they thought they would increase or decrease the time spent at KNP if the river quality changed from the *status quo* to a better quality (hypothetically the 'ideal' river quality scenario) and to a worse quality scenario (hypothetically the 'worst' river quality scenario). Similar questioning was used in McConell (1986) and Thayer (1981) (in Freeman 1993). The changes in river quality for the 'worst' and 'best' were described in terms of the attributes used in the CA question which followed, and responses were given in days or percentages and all converted to percentage changes.

7.1.3 Conjoint analysis (CA)

Four river attributes were identified which it was felt would be relevant to tourists, and four levels of each of these were defined. The river attributes were the number of crocodiles and hippopotamuses seen (CH), the number of species of water birds seen (B), the general riverscape (R), and the numbers of tall trees (T). The number of possible combinations of four attributes with four levels each is $4^4 = 256$, although some of these are infeasible (e.g. high numbers of crocodiles and hippos are not found in dry riverbeds). The three scenarios previously mentioned, i.e. the *status quo*, the ideal and the worst case were included. In order to limit the number of scenarios which the respondents had to rate, scenarios intermediate to these were designed through randomly selected levels of each attribute, and infeasible combinations rejected. A subset of 16 scenarios was thus developed which captured both

extremes and a number of intermediate values. Five versions of the questionnaire each contained four scenarios, the *status quo* in each case plus three others (Table 7.1).

Respondents were asked to rate the scenarios relative to each other on a scale of 0-10 where 10 indicated the most favoured and 0 the least. Interpersonal comparisons of utility are implied by including all responses in a single model, therefore some standardisation of the utility scale was needed. The *status quo* was included in each version to provide a benchmark both for the tourist (although this scenario was not identified to them as the *status quo*) and for standardisation across questionnaires of the given scores to a common *status quo* score. To standardise, the given scores were examined for means and variance. The mean score for the *status quo* in the questionnaire version containing the ideal, z_{SQ}^* , was taken as the benchmark (that which included the worst case could equally have been used). The difference between z_{SQ}^* and the mean score for the *status quo* in each other questionnaire version z_{SQ}^i was added to each respondent score for each scenario (see results, Table 7.2). This seemed justifiable given the small variation between mean scores for the *status quo* across questionnaires and the low variance in the scores for the *status quo* within questionnaires, and the generally similar variances for all scenarios' scores (see results Table 7.2). Similar adjustments to utility scales were mentioned in Mackenzie (1993) and Roe *et al.* (1996).

Table 7.1. The five survey versions for the conjoint analysis showing attributes and attribute levels.

Scenario	Crocodiles & hippopotamus CH		Water-birds B		Riverscape (level) R		Trees (density) T	
A 1 SQ	10	CH3	15+	B4	dominated	R2	some (~15 large)	T1
A 2	1	CH2	very few (2-3)	B1	dry riverbed	R0	none	T0
A 3 Worst	0	CH1	0	B1	dry riverbed	R0	none	T0
A 4	1	CH2	common (5)	B3	uniform	R1	few (~ 5 large)	T3
B 1 SQ	10	CH3	15+	B4	dominated	R2	some (~15 large)	T1
B 2	20	CH4	very few (2-3)	B1	uniform	R1	none	T0
B 3	10	CH3	common (5)	B3	uniform	R1	many (~30 large)	T4
B 4	10	CH3	very few (2-3)	B1	dominated	R2	few (~ 5 large)	T3
C 1 SQ	10	CH3	15+	B4	dominated	R2	some (~15 large)	T1
C 2	10	CH3	15+	B4	uniform	R1	few (~ 5 large)	T3
C 3	1	CH2	0	B1	uniform	R1	few (~ 5 large)	T3
C 4	1	CH2	0	B1	uniform	R1	some (~15 large)	T1
D 1 SQ	10	CH3	15+	B4	dominated	R2	some (~15 large)	T1
D 2 Ideal	20	CH4	15+	B4	diverse	R4	many (~30 large)	T4
D 3	1	CH2	very few (2-3)	B1	diverse	R4	few (~ 5 large)	T3
D 4	0	CH1	very few (2-3)	B1	dry riverbed	R0	some (~15 large)	T1
E 1 SQ	10	CH3	15+	B4	dominated	R2	some (~15 large)	T1
E 2	1	CH2	15+	B4	uniform	R1	some (~15 large)	T1
E 3	0	CH0	common (5)	B3	dry riverbed	R0	few (~ 5 large)	T3
E 4	20	CH4	very few (2-3)	B1	diverse	R4	some (~15 large)	T1

Normally, a CA question to estimate WTP would directly include an attribute which measured cost in some sense (travel expenses, travel time, permit, access fee etc.). The only possible attributes to include in this case would have been time taken to travel to a river within the park, if the river were a specific destination or an increase in park entry fee. We felt that the latter would not be useful as realistic changes would be too low relative to total travel costs. However, in the context of a trip within a game park, travel time is unlikely to be felt as a cost as game and habitats are enjoyed along the way. On the other hand, it was also felt to be unrealistic to ask tourists CVM questions on several scenarios – which would have been necessary to determine WTP for different river attributes. Therefore the CBV and CA questions were designed to complement each other and to allow a link between river quality (attribute

levels), and length of stay in KNP and ultimately total WTP through the TCM which provided WTP for the *status quo*. The links between the various stages of this analysis are illustrated in Figure 7.2.

There are many ways in which to analyse the data from the conjoint question and, therefore, different model specifications were examined so as to assess their potential use in decision-making processes. These followed and expanded on the study of Cohen (2001). This required assessing (a) the influence of assumed linearity of responses in the parameters CH, B, and T, in the original model (b) the stability and robustness of the models in predicting scores, and (c) whether cardinal responses were necessary (in respondent's scores) or whether ordinal information would be sufficient. It was important to assess the use of ordinal information only, as respondents may find it easier to give rank orders and ordinal information would not be affected by respondents using different parts of an interval type scale as might be the case with other models. The stability of models and model predictions was tested by looking at results of the same regressions omitting one questionnaire version. The relationship between the scores and the attribute levels was therefore investigated using four different model specifications, all of which were additive combination of attributes. Briefly the models were as follows:

1. In the first model (Model1) three of the attributes were assumed to be continuous and linear, and one categorical, and the coefficients were found using ordinary least squares (OLS or L_2 -norm) in Excel.
2. In the second model (Model2), all variables were assumed to be categorical, and the coefficients were found using OLS regression in Excel.
3. The third model (Model3) used a linear programming (LP) formulation of the problem, with all variables categorical, and therefore minimised absolute deviations (L_1 -norm) as opposed to squared deviations using LINDO (LINDO 2000).
4. The fourth model (Model4) made use of only the ordinal information from the responses, and also used LP to minimise absolute deviations using LINDO (LINDO 2000).

Model1 used OLS regression in Excel, to estimate regression coefficients β_i for CH, B and T and the u_{ij} which showed the utility of moving from the lowest level (u_{i1}) of that attribute to level j for R (riverscape). The indicators I_{ijr} take on values 0 and 1 (0 if level j does not occur, 1 if level j does occur in a particular scenario r). The utility score Z for respondent k , scenario r is:

$$Z_r^k = \alpha + \beta_1 CH + \beta_2 B + [u_{32} I_{32r} + u_{33} I_{33r} + u_{34} I_{34r}] + \beta_4 T,$$

where CH, B, and T were continuous (and linear) variables and $u_{ij}(I_{ijr})$ refers to attribute R which was categorical and u_{31} (for R level 1) is subsumed in α . The term 'utility' is used in its economics sense, where no risk or uncertainty is implied. In this case u_1 is given by $\beta_1 \times CH$, u_2 by $\beta_2 \times B$, and u_4 by $\beta_4 \times T$.

Model2, Model3 and Model4 were all based on the idea of assessing the value of moving from the lowest level of an attribute (e.g. CH₁ - no crocodiles and hippos) to another level (CH₃ - 10 crocodiles and hippos). Reformulating the attributes as u_{ij} to u_{4j} , this example would be the value of moving from level u_{12} to u_{13} (Stewart 1997).

Formulating this as an ordinary least squares regression (OLS) (so-called L_2 -norm) model for Model2 gives (Stewart:

$$Z_r^k = K + \sum_{i=1}^4 \sum_{j=2}^4 u_{ij} I_{ijr},$$

where 12 u_{ij} parameters are estimated instead of 3 β_i and 3 u_{ij} , the u_{i1} are subsumed in K , and I_{ijr} takes on the values 0, 1 depending on whether that particular attribute level is represented in a scenario.

Similarly, we can formulate this as a linear programming (LP) problem (an L_1 norm) for Model 3, solving:

$$\begin{aligned} &\text{minimise} && \sum_{i=1}^4 \sum_{j=2}^4 \delta_r^k + \varepsilon_r^k \\ &\text{subject to} && Z_r^k - K - \sum_{i=1}^4 \sum_{j=2}^4 u_{ij} I_{ijr} + \delta_r^k - \varepsilon_r^k = 0, \text{ for every respondent } k \text{ and scenario } r. \end{aligned}$$

The LP finds values for K and u_{ij} which minimise the sum of absolute deviations. One can add further constraints, for example, such that that $u_{i,j+1} \geq u_{ij}$ for $j = 2, 3, 4$. This embodies the principle of ‘ordinal dominance’ or simply ‘more of a good thing is better’. However, in this particular case it might not be true as it may be that more of a particular variable is generally better, but beyond a certain point, more might represent a disutility. For example, seeing many crocodiles beyond a certain number might detract from a wildlife viewing experience, as they might be perceived to be ‘common’. However, within the range of the variables considered here, this was deemed a reasonable constraint:

$$\begin{aligned} u_{i,j+1} - u_{ij} &\geq 0.01 \\ u_{i2} &\geq 0.01, \text{ where } 0.01 \text{ was felt to be a small enough difference.} \end{aligned}$$

The problem was also formulated for Model4 so as to only make use of the ordinal information available. This may be desirable because the internal scales used by respondents may be different, and the ordinal information while maintaining preference information, ignores the scales. Respondent k 's scores indicate that she orders the four scenarios s_p^k such that $s_1^k \succ s_2^k \succ s_3^k \succ s_4^k$, etc. If the indicator $I_{ij}(s_p^k)$ is for attribute i level j applied to scenario s_p^k then u_{ij} should satisfy:

$$\sum_{i=1}^4 \sum_{j=2}^4 [I_{ijr}(s_p^k) - I_{ijr}(s_{p+1}^k)] u_{ij} > 0, \text{ for } p=1,2,3$$

and all respondents k , and setting $u_{i1}=0$ for the utility of a scenario with worst levels for each attribute. Adding deviation variables, and weighting these to give greater importance to those deviations which are associated with preference statements with which more people agree, this can be set up as an LP model:

$$\begin{aligned} &\text{minimise:} && \sum_k \sum_{r=2}^4 w_p \delta_{kp}, \text{ where } w_p \text{ is the total number agreeing with preference statement, } p. \\ &\text{subject to:} && \sum_{i=1}^4 \sum_{j=2}^4 [I_{ijr}(s_p^k) - I_{ijr}(s_{p+1}^k)] u_{ij} + \delta_p^k > 0 \\ &\text{and} && u_{i,j+1} - u_{ij} \geq 0.01 \\ &&& u_{i2} \geq 0.01, \text{ as in the previous model.} \\ &\text{and:} && \sum_{i=1}^4 u_{i4} = 10 \end{aligned}$$

7.2 Results

7.2.1 Travel cost, consumer surplus and contingent behaviour

Park revenues and total numbers of visitors were obtained from KNP statistics. Annual revenue for 1999/2000 was R135 793 000 and the total number of visitors was 739 582.

Proportion of travel cost due to KNP and to rivers in KNP

In terms of the importance of KNP to their motivation for their trip, KNP was on average 85% of the reason for South Africans, and on average, 51% of the reason for the trip for international tourists. The average budget for their entire trip could thus be apportioned so as to indicate the total travel costs both in the park and outside, which were attributable to KNP. For international tourists, on average R5 000 per person was attributable to the park (based on a median length of stay of 3 days), and for South Africans, on average R1 200 was attributable to KNP (based on a median length of stay of 5 days). Based on these data the total tourist expenditure attributable to KNP was estimated at R267 mill. Of this, 22% (R58 mill) was due to the Crocodile catchment portion of the KNP, as this accounted for 22% of the total bed nights of KNP (R30 mill of this was on-site expenditure).

Approximately 28% of roads in this part of the park have been laid next to rivers, but tourists drove about 32% of their total within park drives along rivers. The percentage of time spent by tourists along rivers was 25%, the percentage enjoyment (out of three habitats) was 30%. Therefore, between 25% and 32% of the R58 mill of on- and off-site expenditure was, in turn, attributable to the Crocodile river and its tributaries (about R18 mill).

Consumer surplus

Visitors from different origins could be grouped in different ways to represent travel zones, and groups could be included or excluded from the analysis if the sample size was felt to be too small. The way in which the countries of origin were grouped, included or excluded made a vast difference to consumer surplus, giving rise to values ranging from R1.3 bill to R24.2 bill. A figure of R6.4 bill is used here because the grouping of data explained the highest variance ($r^2=0.65$) in the initial regression (Figure 7.3). Of this, R1.4 bill is apportioned to the Crocodile catchment area within the KNP, and R419.6 mill to the river itself (and its associated attributes). The designation of travel zones, and the aggregation of consumer surpluses from different zones implies that incomes within and between zones are reasonably homogenous which would not be the case. Including gross national product per capita as a variable in the original visits/TC regression did not improve the explanatory power of the regression, nor substantially change the results reported above, and the coefficient did not have the expected sign (positive). Negative coefficients are counter-intuitive but seem to be not uncommon in the literature (e.g. in Strong 1983, Loomis 1995, and UNEP 1995), Loomis suggesting that perhaps those with higher incomes, have greater opportunity costs of time, and therefore take less trips.

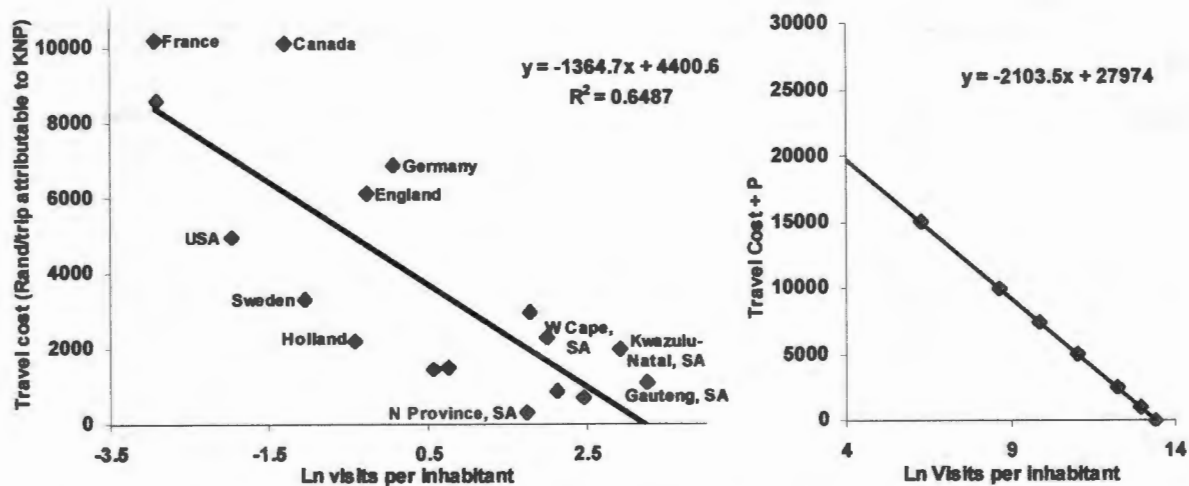


Figure 7.3. Left: visits and travel cost from different origin zones. Right: derived demand curve for visits to KNP.

Change in behaviour due to river quality: contingent behaviour

Tourists' on average said they would spend 24% more time in KNP if river quality improved to a hypothetical best and 29% less time if river quality were in a hypothetical worst condition and the median change in length of stay in both cases was zero.

7.2.2 Conjoint analysis

The responses to the conjoint analysis question were quite consistent (Table 7.2) with fairly low variation in the scores for each scenario. However diagnostics from the subsequent analysis revealed that lower scores were generally more variable than higher scores. This points to one of the reasons for attempting the analysis using ordinal data - people view different ends of the 0-10 scale differently. However, scores were reasonably evenly spread, apart from the expected high frequencies of 10s (for the *status quo* and ideal which were both better than a middle level scenario) and of 0s (for the worst case) (Figure 7.4).

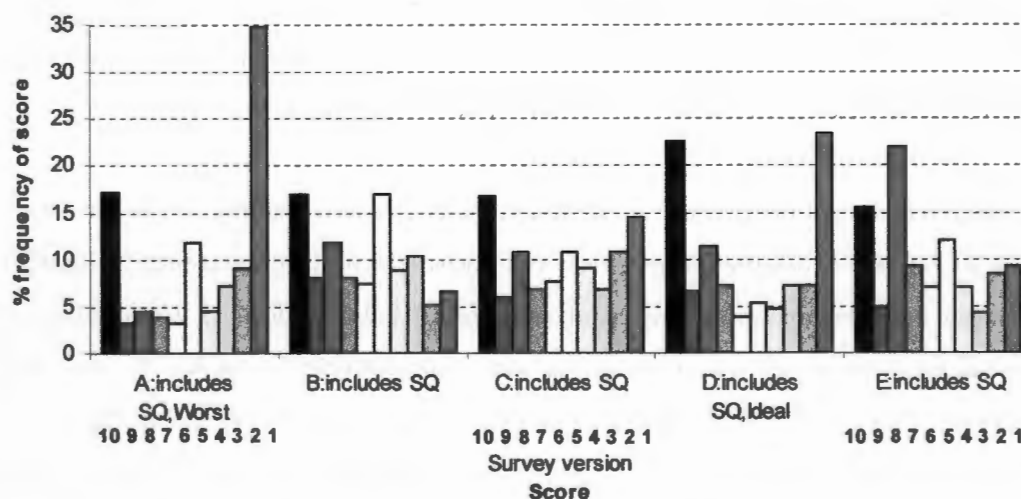


Figure 7.4. Frequency of unadjusted scores in five different scenario versions.

Table 7.2. Survey versions with corresponding average scores, standard deviations and adjusted.

	n	Original average score	Std deviation	Difference of SQ from D1-SQ	Adjusted average scores
A1 SQ	38	9.447	0.921	1.737	7.711
A2		2.921	1.950		1.184
A3 Worst		1.053	0.324		-0.684
A4		4.079	1.880		2.342
B1 SQ	35	8.514	1.574	0.804	7.711
B2		4.514	2.726		3.711
B3		6.357	2.248		5.553
B4		4.900	2.727		4.096
C1 SQ	33	9.242	1.146	1.532	7.711
C2		6.621	1.746		5.089
C3		3.242	2.250		1.711
C4		2.758	1.621		1.226
D1 SQ	38	7.711	1.580	0.000	7.711
D2 Ideal		9.447	1.058		9.447
D3		3.947	2.359		3.947
D4		1.342	0.815		1.342
E1 SQ	36	8.486	1.713	0.776	7.711
E2		6.069	2.462		5.294
E3		3.389	2.499		2.613
E4		6.597	2.219		5.822

The results of the various models (regression coefficients, r^2 and p-values where available) are given in Table 7.3. The number of waterbird species was the most valued attribute, followed by riverscape. The coefficients for Models 1, 2 and 3 are similar, all have the expected signs and relationships, and most are significant (generally $p < 0.001$) for the OLS versions. The amount of variation explained was similar for the two OLS models. Models 1, 2 and 3 showed similar results for predicted scenario scores (Table 7.4). Model3 was least well able to predict the scores of the original 16 scenarios, although this model was particularly robust in being able to predict scenario scores when one survey version was omitted from the regression (Appendix 7B). Model3 was also able to maintain the ‘correct’ preference ordering of coefficients (‘more is better’) in each test (see value functions in Appendix 7A). If assessing the conjoint data on its own, all that is important is that the model is able to accurately predict the order of scores and the size of the gaps between scores. We also required that the models are reasonably faithful to the relative scores of the respondents, in that these are later regressed against their travel costs. In terms of these different requirements it is difficult to choose between Models 1 to 3. However, for Models 1 and 2, survey design is more crucial for adequate results, whereas Model3 was significantly more robust (see predictions with one version missing Appendix 7B). In general, therefore, Model 3 would be the preferred formulation for this type of survey.

However, the model which only used ordinal information, Model4, proved unusable, in that it could not reliably reproduce or predict Z-scores or ordering except directly that imposed by the constraints that $u_{ij+1} - u_{ij} \geq 0.01$. This was because the reduction in level of information from cardinal to ordinal, left the ordinal LP with insufficient information to reconcile variations in responses within survey versions. For example, in version A, most people had the preference order: Sc1 $\approx$ Sc4 $\approx$ Sc2 $\approx$ Sc3. While all version A respondents ($n=38$) agreed that Sc1 was the most preferred, and 30 agreed that Sc4 was next, 13 disagreed with the position of Sc2 and 14 with the position of Sc3. Even though preferences were rather consistent, and for Sc1 and Sc4 generally agreed upon, and even though the deviations in the LP were weighted according to how many people supported a particular ordinal preference (i.e. the more people who had a particular ordinal preference, the higher the weight on the deviation), the LP could not

reconcile contradictions. Of course, if there had been less agreement on the ordering of the scenarios one would not want the LP to ‘reconcile’ these differences. Other forms of CA – conjoint measurement or conjoint scaling – are specifically designed to be able to derive interval scale value functions from purely ordinal judgements. However, these ordinal rankings are given between pairs of attributes, so the informational content is higher than the example described here, but the number of questions required means that these approaches are really only appropriate in direct analyst- respondent interactions.

Table 7.3. Summary of the four conjoint analysis models’ regression coefficients and regression statistics.

Level	#	Model1-OLS. $r^2=0.671, \beta_{u_{ij}}(p)$	Model1 associated score	Code	Model2-XL $r^2=0.684, u_{ij}(p)$	Model3-LP- interval information	Model4-LP-ordinal information only
K		0.207 (0.212)	0.207	K	-0.376 (0.201)	0	0
Croc1	0	0.095 (<.001)	0	U ₁₁	*	*	*
Croc2	1		0.095	U ₁₂	0.583 (.217)	0.01	0.01
Croc3	10		0.950	U ₁₃	1.228 (.056)	0.871	0.02
Croc4	20		1.900	U ₁₄	2.385 (.001)	1.992	2.014
Bird1	0	0.223 (<.001)	0	U ₂₁	*	*	*
Bird2	2.5		0.557	U ₂₂	0.978 (.002)	0.373	0.01
Bird3	5		1.115	U ₂₃	2.005 (<.001)	2.214	0.02
Bird4	15		3.344	U ₂₄	3.659 (<.001)	3.756	2.014
Rscape1		*	*	U ₃₁	*	*	*
Rscape2		0.738 (<.001)	0.738	U ₃₂	0.389 (.466)	0.831	0.02
Rscape3		2.237 (<.001)	2.237	U ₃₃	2.203 (.001)	2.942	2.004
Rscape4		2.373 (<.001)	2.373	U ₃₄	1.983 (.002)	3.232	2.014
Trees1	0	0.063 (<.001)	0	U ₄₁	*	*	*
Trees2	5		0.313	U ₄₂	0.472 (.234)	0.01	0.01
Trees3	15		0.939	U ₄₃	0.918 (.009)	0.627	0.02
Trees4	30		1.879	U ₄₄	2.042 (<.001)	1.02	3.958

* For the categorical variables, the lowest level of each attribute is subsumed in K, the ‘y-intercept’.

Table 7.4. Predicted Z-scores for all scenarios from all models, compared to original (adjusted) scores.

Scenario	Actual (adjusted) Z-score	Model1 OLS, one categorical	dev ²	Model2 OLS- all categorical	dev ²	Model3 LP interval scale	dev ²	Model4 LP ordinal scale	dev ²
A1-SQ	7.71	7.68	0.00	7.63	0.01	8.20	0.24	4.06	13.34
A2	1.18	0.86	0.11	1.18	0.00	0.38	0.64	0.02	1.35
A3-Worst	-0.68	0.21	0.79	-0.38	0.09	0.00	0.47	0.00	0.47
A4	2.34	2.47	0.02	3.07	0.53	3.07	0.52	0.06	5.21
B1-SQ	7.71	7.68	0.00	7.63	0.01	8.20	0.24	4.06	13.34
B2	3.71	3.40	0.10	3.38	0.11	3.20	0.27	2.04	2.78
B3	5.55	4.89	0.44	5.29	0.07	4.94	0.38	4.02	2.36
B4	4.10	4.27	0.03	4.50	0.17	4.20	0.01	2.04	4.21
C1-SQ	7.71	7.68	0.00	7.63	0.01	8.20	0.24	4.06	13.34
C2	5.09	5.55	0.21	5.37	0.08	5.47	0.14	2.06	9.15
C3	1.71	1.35	0.13	1.07	0.41	0.85	0.74	0.04	2.79
C4	1.23	1.98	0.57	1.51	0.08	1.47	0.06	0.05	1.38
D1-SQ	7.71	7.68	0.00	7.63	0.01	8.20	0.24	4.06	13.34
D2-Ideal	9.45	9.70	0.07	9.69	0.06	10.00	0.31	10.00	0.31
D3	3.95	3.55	0.16	3.64	0.09	3.63	0.10	2.04	3.62
D4	1.34	1.70	0.13	1.52	0.03	1.00	0.12	0.03	1.72
E1-SQ	7.71	7.68	0.00	7.63	0.01	8.20	0.24	4.06	13.34
E2	5.29	5.32	0.00	5.17	0.01	5.22	0.00	2.06	10.43
E3	2.61	1.64	0.96	2.10	0.26	2.22	0.15	0.03	6.67
E4	5.82	5.98	0.02	5.89	0.00	6.22	0.16	4.06	3.11
Total			3.74		2.06		5.25		122.29

7.2.3 Relating the CA scores and attributes to travel cost

The model-predicted Z scores for the *status quo*, ideal and worst case scenarios (e.g. 7.7, 9.7 and 0.21 respectively for Model1) were regressed against TC (Rands spent per trip of the median length of 5 days) for these three

scenarios. Ideal river quality implied a stay 24% longer than the *status quo* and the worst, 29% less, based on the contingent behaviour question regarding lengths of stay. The Z-scores were also regressed against total KNP revenues under the three scenarios (Figure 7.5), implying that revenues would fall from R136 mill under the *status quo* (obtained from KNP statistics) to R103 mill with the worst case scenario, and increase to R175 mill for the best case.

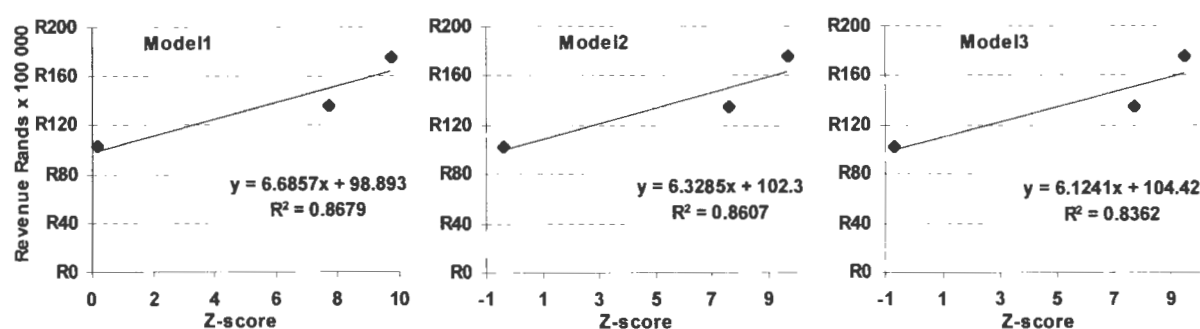


Figure 7.5. Relationship between Z-score and Travel Cost for Model 1, 2 and 3.

Using the relationship between river attributes and utility (Z-score) from the conjoint analysis, and the relationship between utility and expenditure from the regression in Figure 7.5 (based on the contingent behaviour), expected changes in total expenditure could be estimated for any combination of river attribute levels. There is little difference between Model 1, 2 and 3 predicted travel costs with different scenarios.

Additional information

The broader economic valuation project mentioned earlier produced the results summarised in Table 7.5 (Mander *et al.* 2001). These values are presumably based on a change from the *status quo* to a situation where the service were not being supplied at all. However, little comment can be made as to whether these values can be aggregated with the tourism value calculated here, as the assumptions behind the other values were not available.

Table 7.5. Summary of economic values of goods and services supplied by the Crocodile River from Mander *et al.* (2001).

Service	Benefit	Cost
Tourism value to KNP	R18 mill/a (this study, Section 7.2.1, <i>status quo</i>)	
Fish consumed by rural households	R10.2 mill/a saved	
Expenditure on recreational trout fishing	R75.5 mill spent	
Flood regulation	cost savings not calculated	R100 m damage in 2000
River bank maintenance	cost savings not calculated	
Sediment trapping	cost savings not calculated	
Revenue generated by aquaculture	R10.7 mill/a gross turnover	
Industrial and municipal waste treatment	R2 mill/a cost savings	
Benefits of clean water to consumers	not estimated	
Exports to marine ecosystem	Contribution to R25 mill/a Mozambican prawn industry	
Control of pests and pathogens	not estimated	R5.5 mill costs for treatment, R11 mill lost productivity to households

7.3 Discussion

The aim of the study was to estimate the effect of changes in river quality on tourism value, a non-consumptive direct-use value. This value, together with other economic values, would allow for the evaluation of different river

quality scenarios so that the societally optimal scenario could be chosen as a desired management class. A fairly complicated linking of different methodologies (Figure 7.2) was necessary because:

- (a) tourist trips to KNP are often part of multiple-destination trips;
- (b) there are many environmental and other contributions to a KNP tourism 'package';
- (c) the change in environmental quality being measured (river quality) was not functionally independent of the quality of the KNP itself (for the tourist to answer the contingent question, she had to imagine a KNP as it is in all respects except that the river and riparian zone were as described);
- (d) a conventional measure of travel cost (time, petrol), could not realistically be considered as a cost for a CA question, when travelling within the park, even where the specific destination might have been a river.

Other complexities related to what kind of value was elicited or assumed, and some concerned the relationships between the values elicited from different questions.

A primary assumption was that tourism *decisions* are influenced by river quality. Tourism value was assumed to be reflected by visitor expenditure in undertaking the recreational activity – thus an analysis of travel cost formed the basis for the study. This analysis provided both direct expenditures at the destination, the proportion of overall expenditure attributable to the destination in question (in the case of multiple destination travel), and consumers' surplus where the surpluses for each zone are summed to give total consumers' surplus. The next assumption in assessing tourism value is that preferences for different qualities of the environment (as reflected by the CA) would be reflected in economic *behaviour* (and indicated by the CBV) and that changes in economic behaviour should be the basis of societal decision-making. As changes in economic behaviour with changes in quality cannot be observed in the short term and without multiple comparable sites, these have to be ascertained from surveys - in this case, with the CBV. The CBV question only provided an evaluation of three scenarios – the *status quo*, ideal and worst (and it was not possible to include more scenarios within the CBV format), and so the CA was necessary to provide a fuller picture of preferences for additional river scenarios as well as preferences for different attributes of rivers. The *status quo* therefore formed a link between the three methods (TCM, CBV, and CA).

Specifically, it was assumed that we could directly relate the *status quo* travel cost (TC) monetary values to the *status quo* CA utilities (preferences, revealed or stated, are reflected in economic choices). The revealed TC (with consumer surplus) for the *status quo* $TC^{TCM}(sq)$ reflects the utility gained by the tourist through the tourism experience. In the CA we directly estimate $U^{CA}(a)$, total utility from a river health scenario a , and derived $u_i(a)$, the single attribute utilities for each attribute, from the regression. We therefore linked the TC and U values obtained for the *status quo*, assuming that $TC^{TCM}(sq)$ was a linear function of $U^{CA}(sq)$ or $TC^{TCM}(sq) = \alpha U^{CA}(sq) + \beta$. Next, it was assumed that the stated change in length of time in KNP (from the CBV question) would be directly and proportionately reflected in $TC^{TCM}(Ideal)$ and $TC^{TCM}(Worst)$ from the TCM, and in $U^{CA}(worst)$ and $U^{CA}(ideal)$ from the CA. This means that $TC^{TCM}(Ideal) \propto U^{CBV}(ideal) \propto U^{CA}(ideal)$ –making a direct link between river quality, preferences and economic behaviour. The CBV question was pivotal in this study because it linked the TC monetary values to the CA utility values. The form of the question was perhaps naïve, but more detailed questioning was not

possible in a 10 minute in-person survey, and the use of the more conventional change in frequency of visits format was not appropriate in this context. We felt that this question elicited unrealistic responses (i.e. people would *not* actually stay on average 24% longer under an ideal river quality scenario, or 30% under the worst case scenario). The reasons for this were threefold. Firstly, both budget and time constraints would limit the ability to stay longer, and the many international tourists in particular, would have little flexibility to stay longer. Secondly, the river quality changes described were such that the river would not be obviously ‘unnatural’ (e.g. the scenarios did not include visible pollution). Thirdly, the median change in length of stay was zero for both ideal and worst case scenarios (see also Section 9.4.1 and Figure 9.3). If we had used the median rather than the mean, the only conclusion possible in an *economics* framework would have been that *river quality does not affect tourism value*. Intuitively, it seems likely that actual revenues would be little affected (i.e. tourism behaviour would not change as suggested by the median) except with fairly gross changes in quality (e.g. visible pollution), or if the tourist knew that the river were suffering adverse quality effects from upstream abstraction. But, for example, in a semi-arid environment such as that in which the KNP is situated, low-flow in rivers might be perceived to be natural, and therefore without prior information, the tourist would not have less utility than for the truly natural flow situation.

However, it was clear from the CA question that tourists have *preferences* for different river scenarios. In an economics framing, do we conclude that the preferences reflected in the CA analysis are not relevant as they may not actually be reflected in behavioural (i.e. economically relevant) changes as reflected by median values from the CBV? These questions highlight the fact that there are in fact two issues of interest: tourist revenue – a ‘real’ economic value, which might only be affected by gross changes in quality, and tourist experience (utility). The latter might be affected by more subtle changes in quality, or may not be affected by quality (where quality is a synonym for integrity or naturalness), but more by aesthetics and game viewing. High tourist value for aesthetics and game viewing will not necessarily be directly and positively correlated with high river integrity or naturalness. For example, when faced with a choice in a questionnaire the tourist might say she prefers a lush riverine scene with many rhinos and crocodiles, whereas the ‘natural’ riverine scene may be more arid, with fewer species. *In situ*, as it were, the tourist might prefer either scene depending on her perception of which were more natural. If not faced with a choice (i.e. in ignorance of alternatives), she may be quite happy with a uniform river with low biodiversity, as this might appear to be the natural state of the river. In short, tourist preferences may say nothing about environmental sustainability, whereas ironically, tourist revenues might help to ensure conservation. Additionally, the answers to the CBV (most probably) reflected citizen values rather than consumer values, and there is, therefore, in the equating of values from different sections described above, a mixing of citizen and consumer preferences. This is further discussed in Chapter 9.

Further, tourism value is only one of the values needed to assess the desirability of a river health scenario. Within an economic decision-making framework, other relevant values would also need to be included, the goal of Mander *et al.* (2001, see Table 7.5). Many of these values could be found with conventional economic analyses (using e.g. harvesting figures and replacement costs) once the river ecosystem goods and services were defined. Other values,

such as existence value, if required, would also require a survey with respondents comparing different scenarios. With the environmental economics toolbox this would also have required CVM or a combination of CVM and CA.

An important result of the analysis was to show that, although ordinal information might be easier to obtain and not be subject to the problem of different scale use, with the conjoint design employed here, ordinal information would not have been sufficient to predict Z-scores. With a different conjoint design (e.g. pairwise comparison of attributes) ordinal information would have probably been sufficient. However, such a design would have required a much larger sample size, and would have increased respondent burden considerably. From the point of view of practicality of questionnaire administration and usefulness of results, therefore, rating rather than ranking seems preferable, particularly as the respondents had no difficulty giving these scores (i.e. to indicate strengths of preference). The problem of different interpretation and use of the scale remains. A qualitative assessment of the scores given indicated that this was probably not a severe problem in this study. This is clearly an area where controlled experiments would be useful.

A recommendation of this study was that the consumer surplus estimate from TCM should not be used in settings with high variability in tastes and incomes in source zones. Two of these are (a) the extreme sensitivity of the results to organisation of the data, and (b) the heterogeneity of populations within and between zones, especially their incomes would invalidate the aggregation of individual and zonal consumer surpluses (see Sections 3.4 and 9.5.1). Income can be included as a variable in the travel-cost equation, but we found that incorporating income (in the form of mean per capita income per zone) did not add to the explanatory power of the regression (Turpie and Joubert 2001). Ignoring consumer surplus means that we do not have a true measure of WTP (which should include consumer surplus), but rather a straightforward measure of revenue to KNP. From the CBV question, we hoped to get an indication of the changes in revenue to KNP that could be expected with changed river quality. In fact, due to issues discussed further in Section 9.4.1, we probably had an indication of how strongly people felt about changes in river quality, which would not be proportionally reflected in changed consumer behaviour or revenues.

7.4 Summary and conclusions

The tourism value of the sections of the Crocodile river and its tributaries associated with KNP were investigated, together with changes in that value with changed river health. The value for the *status quo* was found using the TCM with travel costs being apportioned from a total value for the southern KNP, to the Crocodile catchment area, and then to the rivers themselves. The TCM also provided estimates of total WTP (price+consumer surplus), although these values were very sensitive to data arrangement and we concluded that the estimates of consumer surplus were unreliable. A CBV question established changes in the length of stay with changes to ideal and worst case river health scenarios. Finally, different combinations of four river attributes relevant to tourists (crocodiles and hippos, waterbird species, riverscape, large trees) were scored by tourists (CA). These scores were regressed against the attributes using four different models. Tourists did not object to giving scores rather than ranks and the model based only on ordinal information proved unsuitable. We concluded that, where possible, scores should be obtained, and adjusted to a common *status quo* to limit the problems of respondents using different ranges. An additional

improvement would be to link more detailed verbal cues to scores. The scenarios included the *status quo*, ideal and worst case, and 13 other scenarios. Thus, any combination of river attributes could be related to travel cost. This information could later be linked to ecological and hydrological examination of different scenarios to assist in choosing a level of river health that maximised social welfare or balanced ecological, social and economic criteria. This stage has not been undertaken. The values for various other ecosystem goods and services from the broader study could not, at this stage, be aggregated with tourism value, as the assumptions on which they were based were not available and they were not based on scenarios (they were only available for an unspecified change from *status quo*).

The values elicited are based on individual preferences, as compared to those in the SMARTx study of the previous study where ‘group values’ were elicited. However, the specific value being investigated, tourism value, could be compared to criteria in the Sand case. For example, in the Sand case revenues from tourism were included in the operating margin criterion, and social values included aesthetics. In addition, ecological criteria indicated additional values which might be associated with the EE concept of existence value or with additional services (indirect-use values) provided by the river.

The following chapter examines two SMARTx studies and two EE studies which used similar approaches to those in this and the previous chapter but in different contexts. This is followed by a comparative discussion of the cases in Chapter 9. Chapter 10 synthesises these discussion with additional insights from Part I of this thesis.

Appendix 7A. Comparison of conjoint models: value functions

CH=Crocodiles and Hippopotamuses, B=Waterbirds, R=Riverscape, T=Trees.

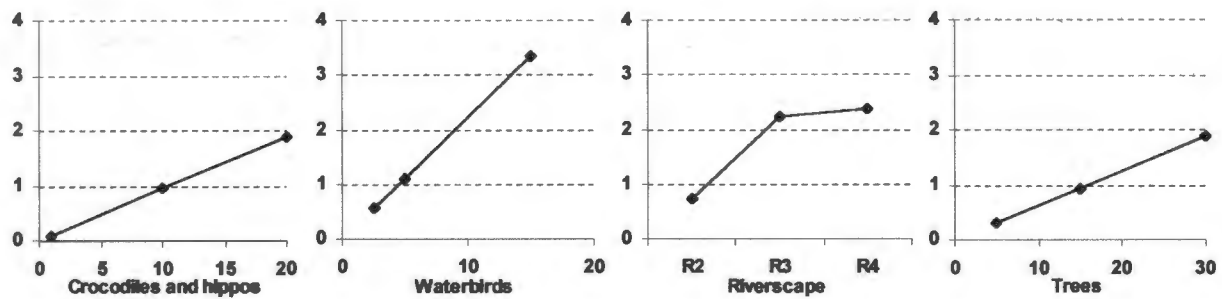


Figure Appendix 7A.1 Model 1 value functions for river attributes

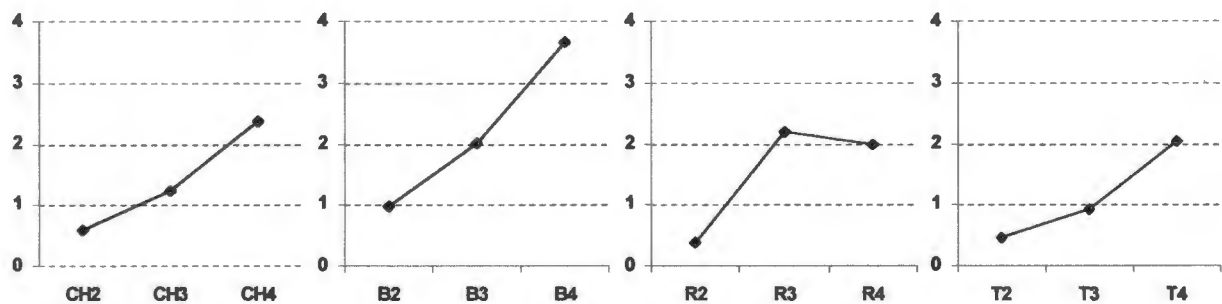


Figure Appendix 7A.2. Model 2 value functions for river attributes

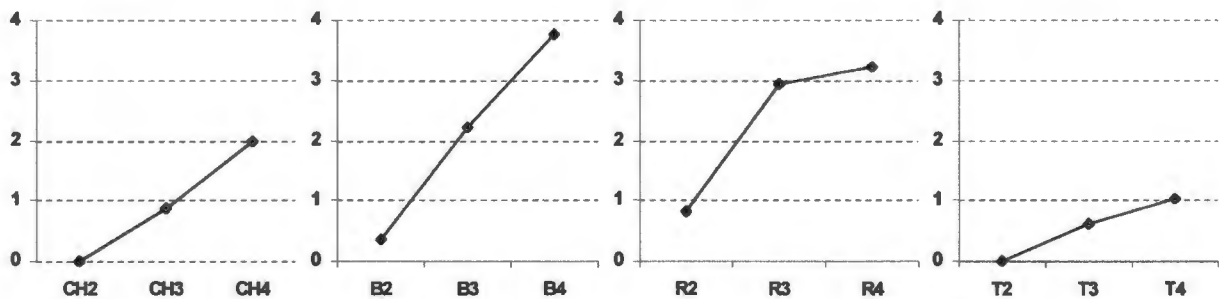


Figure Appendix 7A.3. Model 3 value functions for river attributes

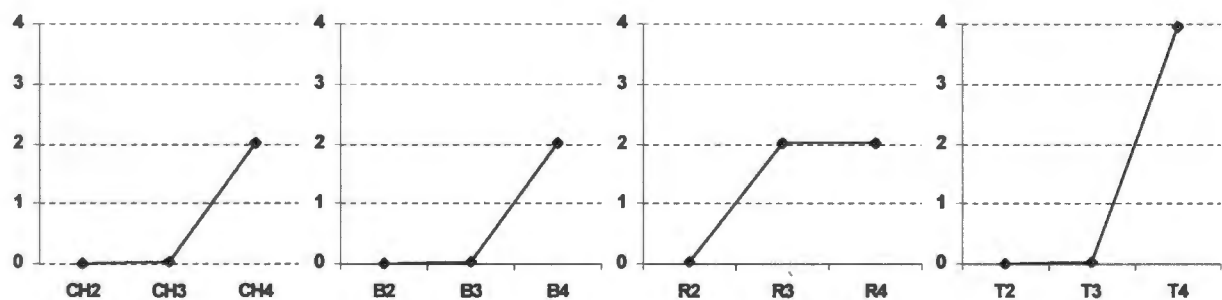


Figure Appendix 7A.4. Model 4 value functions for river attributes

Appendix 7B. Prediction of scenario scores with one survey version omitted from the regression.

Model1

	Average adjusted Z-score	All included	Version A omitted	Version B omitted	Version C omitted	Version D omitted	Version E omitted
A1-SQ	7.711	7.678		48.99	41.43	23.12	46.49
A2	1.184	0.859	1.213	0.932	0.852	0.966	0.531
A3-Worst	-0.684	0.207		0.242	0.199	0.28	-0.09
A4	2.342	2.467		16.2	13.48	7.609	14.92
B1-SQ	7.711	7.678	37.35	48.99	41.43	23.12	46.49
B2	3.711	3.401	3.067	2.15	2.656	3.433	2.572
B3	5.553	4.888	66.77	89.75	74.3	38.72	84.89
B4	4.096	4.265	13.24		14.62	8.32	16.3
C1-SQ	7.711	7.678	37.35	48.99		23.12	46.49
C2	5.089	5.552	15.55	19.27	16.55	10.98	17.91
C3	1.711	1.353	11.41	14.95	12.35	6.488	13.89
C4	1.226	1.979	32.77	44.14		18.48	41.5
D1-SQ	7.711	7.678	37.35	48.99	41.43		46.49
D2-Ideal	9.447	9.703	71.85	95.27	79.64	43.68	90.54
D3	3.947	3.545	13.89	17.96	15.09	8.511	16.96
D4	1.342	1.703	33.17	44.67	36.76	18.83	41.85
E1-SQ	7.711	7.678	37.35	48.99	41.43	23.12	46.49
E2	5.294	5.323	36.06	47.91	39.72	21.85	44.59
E3	2.613	1.635	12.35	16.1	13.32	7.398	14.75
E4	5.822	5.976	37.06	48.33	40.84	22.89	46.55

Model2

	Average adjusted Z-score	All included	Version A omitted	Version B omitted	Version C omitted	Version D omitted	Version E omitted
A1-SQ	7.711	7.631		7.71	7.71	7.71	7.71
A2	1.184	1.184	190.78	1.18	1.18	1.18	1.18
A3-Worst	-0.684	-0.376	0.56	-0.68	-0.68	-0.68	-0.68
A4	2.342	3.073		2.34	2.34	2.34	2.34
B1-SQ	7.711	7.631	7.66		7.71	7.71	7.71
B2	3.711	3.376	3.71	1.25	3.71	3.71	3.71
B3	5.553	5.287	5.45	0.57	5.55	5.55	5.55
B4	4.096	4.504	4.31		4.10	4.10	4.1
C1-SQ	7.711	7.631	7.66	7.71		7.71	7.71
C2	5.089	5.372	5.20	5.09	10.32	5.09	5.09
C3	1.711	1.068	1.26	1.71	-0.50	1.71	1.71
C4	1.226	1.514	1.67	1.23		1.23	1.23
D1-SQ	7.711	7.631	7.66	7.71	7.71		7.71
D2-Ideal	9.447	9.693	9.55	9.45	9.45	14.51	9.45
D3	3.947	3.639	3.95	3.95	3.95	1.61	3.95
D4	1.342	1.520	1.25	1.34	1.34		1.34
E1-SQ	7.711	7.631	7.66	7.71	7.71	7.71	
E2	5.294	5.173	5.29	5.29	5.29	5.29	12.37
E3	2.613	2.101	2.72	2.61	2.61	2.61	-4.58
E4	5.822	5.888	5.72	5.82	5.82	5.82	

Model3

	Average adjusted Z-score	All included	Version A omitted	Version B omitted	Version C omitted	Version D omitted	Version E omitted
A1-SQ	7.711	8.196		8.22	8.04	8.26	8.20
A2	1.184	0.383	0.18	0.60	0.26	0.26	0.27
A3-Worst	-0.684	0	0	0	0	0	0
A4	2.342	3.065		3.26	3.26	3.26	3.26
B1-SQ	7.711	8.196	8.0	8.22	8.04	8.26	8.20
B2	3.711	3.196	3.20	3.45	3.20	3.20	3.20
B3	5.553	4.936	4.94	4.90	5.20	5.20	5.20
B4	4.096	4.196	3.58	4.64	4.20	4.10	4.20
C1-SQ	7.711	8.196	8.0	8.22	8.04	8.26	8.20
C2	5.089	5.468	5.47	5.47	5.03	5.47	5.47
C3	1.711	0.851	0.64	1.05	1.79	1.05	0.47
C4	1.226	1.468	1.47	1.45	2.69	1.06	0.94
D1-SQ	7.711	8.196	8.0	8.22	8.04	8.26	8.20
D2-Ideal	9.447	10	10	10	10	12.05	10
D3	3.947	3.625	3.0	4.0	4.0	4.07	3.0
D4	1.342	1	1.0	1.0	1.0	0.03	1.0
E1-SQ	7.711	8.196	8.0	8.22	8.04	8.26	8.20
E2	5.294	5.224	5.22	5.22	5.72	5.22	4.73
E3	2.613	2.224	2.22	2.22	1.49	2.22	3.10
E4	5.822	6.224	6.22	6.22	6.22	6.22	

Model4

	Average adjusted Z-score	All included	Version A omitted	Version B omitted	Version C omitted	Version D omitted	Version E omitted
A1-SQ	7.711	4.058	9.96	0.09	4.058	4.05	0.11
A2	1.184	0.020	9.86	0.02	0.02	0.02	0.02
A3-Worst	-0.684	0	0	0	0	0	0
A4	2.342	0.060		0.05	0.06	0.05	0.05
B1	7.711	4.058	9.96	0.09	4.058	4.05	0.11
B2	3.711	2.044	9.9	0.05	2.044	2.04	0.08
B3	5.553	4.018	9.92	9.96	4.018	4.01	0.08
B4	4.096	2.044	9.9		2.044	2.04	0.08
C1	7.711	4.058	9.96	0.09	4.058	4.05	0.11
C2	5.089	2.064	9.94	0.07	2.064	2.06	0.07
C3	1.711	0.040	0.03	0.03	0.04	0.03	0.03
C4	1.226	0.050	0.04	0.04	0.05	0.04	0.04
D1	7.711	4.058	9.96	0.09	4.058	4.05	0.11
D2-Ideal	9.447	10	10	10	10	10	0.17
D3	3.947	2.044	9.9	0.06	2.044	2.04	0.08
D4	1.342	0.030	9.87	0.03	0.03	0.03	0.03
E1	7.711	4.058	9.96	0.09	4.058	4.05	0.11
E2	5.294	2.064	9.94	0.07	2.064	2.06	0.07
E3	2.613	0.030	9.87	0.03	0.03	0.03	0.03
E4	5.822	4.058	9.94	0.09	4.058	4.05	

8. Other SMARTx and EE case studies

Two SMARTx and two EE case studies are outlined in this chapter. The same methods, with variations, were used as described in the previous two chapters. These case studies influenced the views and generally supported the discussions and conclusions in Part III, but are not reported in as much detail as the main case studies. Their inclusion is intended to illustrate similarities and differences to the applications in the previous two chapters and to highlight features relevant to later discussions and to raise some additional issues, rather than to provide a full account of the entire case study process and results. Generalisations made in Chapter 9 about SMARTx and EE are therefore drawn from three case studies each. Where possible, the sections, figures or tables which are of most relevance to later comparisons and discussions are mentioned near the beginning of each case description so that the reader does not need to read the entire case.

The applications in this chapter include the evaluation of alternatives (the SMARTx cases in Sections 8.1 and 8.2), the valuation of species and other game reserve attributes (the EE application in Sections 8.3) and the valuation of open-space types (the EE application in Section 8.4). The case in Section 8.3 is similar to other EE-valuation contexts, where a particular habitat or species is valued, whereas the evaluation of alternatives (Section 8.4) is the more conventional domain of SMARTx and EE valuation within a CBA framework (referred to as EE-CBA).

The SMARTx applications were:

- An evaluation of forestry and land-use scenarios in the Maclear district using alternative development (AD) and SMARTx in combination called AD-SMARTx (Section 8.1), and
- An evaluation of water supply and demand management options for the City of Cape Town using SMARTx (Section 8.2),

The EE applications included:

- Valuation of the contribution of the 'Big Five' and other park attributes to the tourism value of a national park using scoring, ranking, the travel cost method (TCM), conjoint analysis (CA), and the open-ended contingent valuation method (CVM) (Section 8.3), and
- Valuation of different types of urban open space within Cape Town using scoring, ranking, TCM, CA, and open-ended CVM (Section 8.4).

8.1 Maclear district forestry and land-use decisions

This case is similar in nature to that in Chapter 6, and the same analyses were conducted (e.g. calculation of trade-offs and indirect compensatory values). Sections of most relevance to later comparison are the criteria that were selected (Figure 8.2), the value functions (Figure 8.3) and the comments in Section 8.1.4. Initial results before completion of the project were published in Stewart and Joubert (1998).

8.1.1 Background

The impetus for the work described here came from a widely felt need for a more streamlined afforestation permit application system at a time of rapidly increasing afforestation in the Maclear district. The Maclear district (ca. 200 000 ha) of the Eastern Cape province of South Africa, stretches approximately 80km along the southern foothills of the Drakensberg mountains (Figure 8.1), in an area which typifies the conflicts which often arise. Approximately 1½ million hectares of South Africa is presently under commercial plantations of non-indigenous trees. Much of this afforestation has occurred on the eastern escarpment at the edge of the inland plateau on land which was naturally afro-montane grassland. The Maclear district is at the southern end of the Eastern Mountain 'hotspot' of plant diversity, one of eight recognised for southern Africa (Figure 8.1). About 30 % of the plant species are endemic and only about 5 % of the 'hotspot' is formally conserved, almost exclusively at the northern end (Cowling and Hilton-Taylor 1994). The vegetation is primarily afro-montane grassland which is threatened by afforestation, overgrazing and increased crop-farming throughout Africa. This led to its identification as one of the three most threatened habitats in Africa, in response to which, the World Wide Fund for Nature (WWF) funded a conservation evaluation of the afro-montane grasslands in the area. This was carried out by the Dept. of Nature Conservation of the University of Stellenbosch (Armstrong 1996, Armstrong and van Hensbergen 1997), who asked us to carry out the MCDA work described in this section to evaluate different forestry expansion alternatives.

By the time of this study (1995/1996), the forestry company had acquired approximately 75 000 ha in the region and had planted mainly pine trees on some 38 000 ha, indicating on average that afforestability of the land was about 50 %. This area under afforestation would be insufficient to support the operation of a pulp-mill, but could support the operation of a fairly large sawmill (output of more than 200 000 m³). Given the uncertainty of future demands and the fact that much of the land already owned by the forestry company in Maclear was not ideal for afforestation (implying larger costs), they were seeking permission to extend afforestation within the district.

Economic pressures led many Maclear farmers to sell at the time of initial forestry expansion in 1989, as only larger farms seemed viable after a prolonged drought. The change from predominantly cattle grazing and a farming community to commercial forestry has changed the economic and social structure of the area considerably. Further relevant factors are that the Eastern Cape has the second highest unemployment figures in the country (around 50%), and the relatively wealthy Maclear district is bordered on the east by the Transkei where population pressures, overgrazing and erosion are more extreme. In addition, political changes have seen Maclear district local councils pass from the control of commercial farmers to the African National Congress, who were seeking upliftment of previously disadvantaged communities.

Commercial forests have the potential to seriously restrict run-off into public streams and rivers both directly and through the invasion of other areas by the exotic species. Landowners are required to apply to the Dept. of Water Affairs and Forestry for permits to plant forests. Although primarily concerned with effects on run-off, representations from affected parties were allowed for and a full impact assessment (IA) could be recommended for

each application in order to assess other impacts as well. This process has slowed down the operating of the permit system and also means that small growers may potentially face the very high costs of funding the IA. For this reason, a government Green Paper identified the need for a more streamlined approach to the issuing of permits, which, while still allowing for participation, does not imply such large costs. Our study was seen as contributing towards this aim.

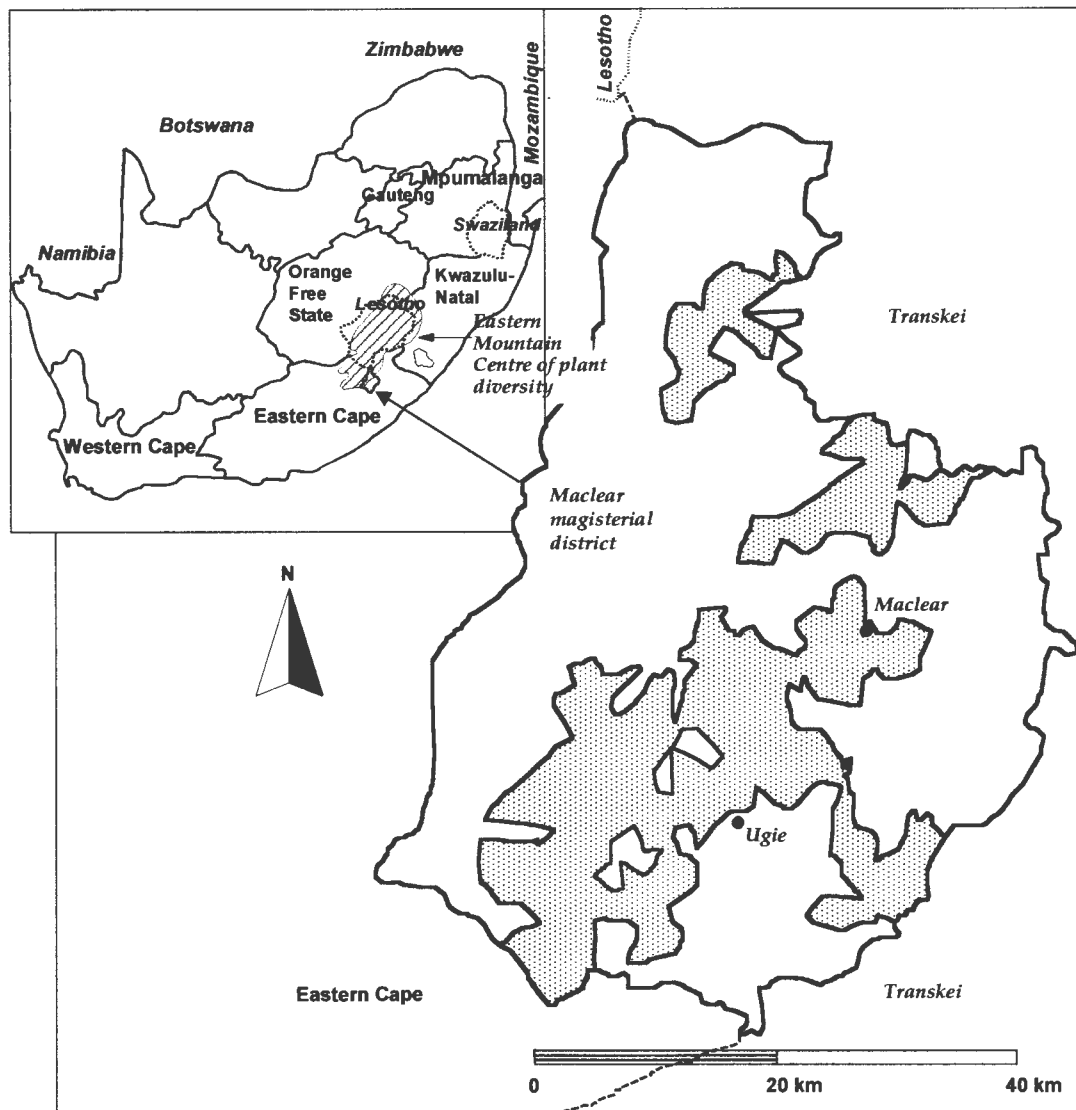


Figure 8.1. Map of the Maclear area of the northern Eastern Cape. The shaded areas are those owned by the commercial forestry company at the time of the study.

8.1.2 Methods

The AD-SMARTx process was similar to that for the Sand case (outlined in Sections 5.3 and 6.1) and took place during four one and a half day workshops. Through an informal process various points of view were identified as relevant to the problem and representatives of these viewpoints were found to attend the workshops or contribute if they could not attend (Table 8.1). These included the forestry company representative, representatives of three relevant government departments, the mayor as a representative of the community, and the ecologists who had conducted the initial land-use evaluations.

Table 8.1. Points of view, stakeholders and their representatives at the workshops.

Point of view / Stakeholder	Representatives
Commercial forestry	North East Cape Forests (the forestry company)
National Forestry planning	Department of Water Affairs and Forestry
Agricultural interests	Dept of Agriculture and Land Affairs, Eastern Cape Province.
Nature conservation	Dept Nature Conservation, Eastern Cape. Dept Nature Conservation, University of Stellenbosch
Social interests	Mayor and Town Clerk of Maclear and Ugie
Hydrology	Environmentek (Council for Scientific and Industrial Research)

Identifying values and appropriate spatial scales for decision-making

The permit system operated on a farm by farm basis, and this was the level which was developed during the first workshop. However, perhaps one of the most useful outcomes of this workshop, was the agreement that this was not an appropriate level at which to make decisions, unless reference could be made to larger scales of decisions, termed the meso-scale (which came to be defined as the Maclear district, with some reference to neighbouring districts) and the macro-scale (which was generally accepted as referring to the national level). It was acknowledged that micro-scale (farm level) decisions would always be necessary, but needed the context of a larger scale in order to avoid sub-optimal incremental decisions. The following three workshops, therefore focused on the meso-scale and re-assessed the criteria of interest, developed hypothetical alternatives or scenarios for evaluation, determined the criteria ranges and evaluated the scenarios at this scale.

Scenario or alternative development (AD)

Using the AD process described in the Sand case, six alternatives were developed, before and during the second, third and fourth workshops, which covered a realistic range of possible future developments at the 'meso-scale' (i.e. the Maclear district) (Scenarios 1 to 6, Table 8.2). This continuous development throughout the process was so that we would *arrive* at a level of detail sufficient for the workshop participants to compare and distinguish between alternatives rather than requiring a specific level of detail up front. The scenario elements were land-use (forestry, grazing or crops) and primary processing of the timber (sawmill or pulpmill). As the scenarios were refined, the impacts or criteria levels were specified in more detail, so that value functions could translate the related quantitative information to a value (e.g. number of land-types preserved), while others were evaluated directly (e.g. housing and services). Those impacts or criteria which were well specified earliest in the process were those which related to work already completed in the district viz. the WWF wildlife indices study (Armstrong and van Hensbergen 1997) and a study of the hydrological affects of afforestation on the quaternary catchments (Forsyth *et al.* 1996).

Some criteria remained were not 'operational' until the fourth and final workshop, specifically those relating to economic impacts. Research between workshops allowed these to be included in the scenario descriptions for this workshop, which meant that the scenarios could be evaluated on the basis of all the criteria specified earlier in the process (Figure 8.2). An issue which arose at the third workshop was the size of the multiplier effect of the forestry primary processing industries. Due to the controversy, therefore, multipliers were included in the scenario descriptions, by including 'sub-scenarios' of a range of possible multiplier effects. It was agreed at the fourth workshop that only local multiplier effects were relevant, with the understanding that decisions on a national scale

would include national level multipliers, either as effects on GDP or as employment multipliers or both. Only evaluations concerning multipliers of all land-uses and processing at 1.2 are shown in this chapter.

The farms of the district were entered into an Excel spreadsheet, designated as either agriculture or forestry and linked to information and calculations relating to area, species present, employment rates, land-types etc. The designation could be changed using Excel's 'scenario' function, allowing all calculations to be updated for different scenarios.

Input from the conservation representatives led to further adjustments during the course of the fourth workshop, namely the addition of a conservation constraint to the effect that no further afforestation should occur on land-types 2, 4 and 9 which had high biodiversity and endemism. This led to the addition of Scenarios 4a, 5a and 6a, which were in all respects the same as Scenarios 4, 5 and 6 apart from this constraint (Table 8.2). As this brought the total number of scenarios to 9, Scenarios 2 and 3 were not specifically evaluated during the workshop, as they were perceived to be not very different from Scenario 4. Where possible, calculations made for these two scenarios subsequent to the final workshop, are included for completeness (Appendix 8A). In summary, seven scenarios were evaluated at the fourth workshop, based on the aforementioned multiplier effects. These were Scenarios 1 (*status quo*), 4, 5, 6, 4a, 5a and 6a.

Table 8.2. Hectares of forestry owned, afforested, primary processing and constraints for the Maclear scenarios.

Scenario	Hectares owned	Hectares afforested	Primary processing	Output	Constraints-forestry free land-types
Scenario1	64 000	35-38000	2 Sawmill	1=200 000 m ³ /a, 1=64 000 m ³ /a	
Scenario2	64 000	44 000	2 Sawmills	total = 330 000m ³ /a	
Scenario3	70 000	50 000	2-3 Sawmills	total = 370 000m ³ /a	
Scenario4	75 000	53 000	2-3 Sawmills	total = 400 000m ³ /a	
Scenario5	"	53 000	Pulpmill in district	300 000 T/a	
Scenario6	"	53 000	Pulpmill outside district	300 000 T/a	
Scenario4a	"	53 000	2-3 Sawmills	total = 400 000m ³ /a	Landtypes 2, 4, 9
Scenario5a	"	53 000	Pulpmill in district, 90 000 T from outside	300 000 T/a	Landtypes 2, 4, 9
Scenario6a	"	53 000	Pulpmill outside district, 90 000T additional	300 000 T/a	Landtypes 2, 4, 9

Selection of criteria and the development of the value tree

Criteria were identified during 'brainstorming' sessions during the workshops. These sessions included an 'electronic brainstorming' session using the GroupSystems software (Ventana Corporation, 1994) in the decision room of the University of Cape Town during the second workshop. In this system, all participants are connected to a small network around a table, can type their ideas at their computer, which then appear anonymously on the screens of all participants. This has the advantage of anonymity, of avoiding dominance by stronger personalities and of allowing the rapid generation of many ideas. The ideas were grouped, categorised and organised into a value tree, which was further developed and refined before and during the remaining workshops (Figure 8.2).

Consequences of scenarios and scoring

New research had to be completed in order to obtain adequate descriptions of the scenarios and criteria ranges for some criteria, in particular the criteria concerning economic impacts. Three criteria were expressed as net present value (NPV). It is worth noting here that the term NPV is usually associated with an 'economic' analysis, but in this

case refers to a 'financial' analysis. A full economic analysis of, for example, forestry income, should include externalities such as social and environmental impacts, the effects of subsidies, price controls and exchange rates etc. In our example, at least some of these are explicitly included in the other criteria considered (e.g. the conservation and social criteria). Already completed research (e.g. Armstrong, 1996 and Forsyth *et al.*, 1996) was used to derive certain ranges for conservation and hydrology criteria.

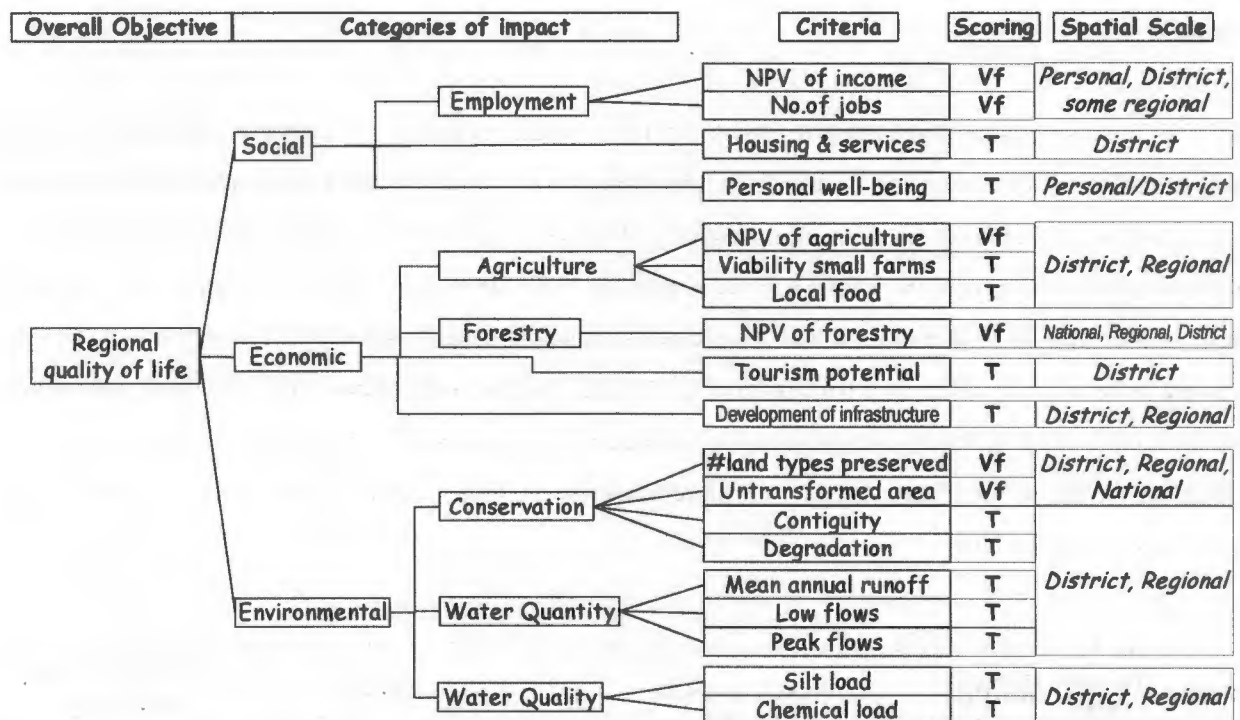


Figure 8.2. Value tree constructed from the criteria and categories defined during the workshops. Vf= Value function scoring, T= Thermometer scale scoring.

Scores on a 0 to 100 thermometer scale were given to the scenarios based on each of the criteria separately using the VISA (1995) software. Linear value functions were used for all criteria relating to NPV, income and employment, and a non-linear value function to relate 'number of land-types preserved' to value (Figure 8.3a,b). Direct scoring was used for all other criteria. Scores were also entered into the Excel spreadsheet.

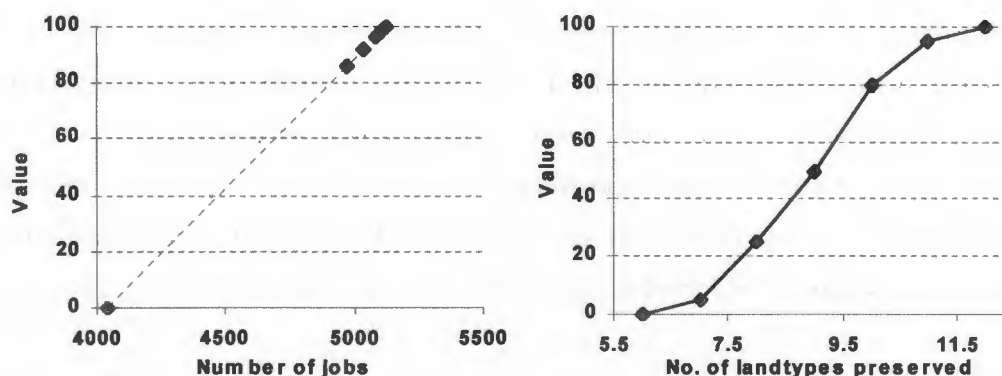


Figure 8.3 (a) The linear value function for employment, and (b) non-linear value function for land-types preserved.

Criteria weights and aggregation of scores

Once scores for the various alternatives were assigned, swing weights were elicited for the criteria. Criteria within a category were first compared, and then the relative importance between the different categories was compared. For example, within the category ‘conservation’, ‘number of land types preserved’ was felt to be most important in some sense and the impact of a swing from the worst level (Scenarios 5 and 6) to the best level (Scenario 1) was perceived to be twice as important as a swing from worst to best on the next most important criterion, ‘contiguity’. These weights were then normalised to sum to one. The relative importance of the three criteria groups (social, economic, and environmental) was determined by the group by discussion. It was agreed that within the decision context, the criteria relating to social issues (specifically employment) were most important and this category was given twice the weight of the other two categories (Figure 8.4). Once weights were assigned to all the criteria, the scores could be aggregated at different levels of the hierarchy, based on a weighted summation (equation (5.4) page 64). The VISA software or Excel spreadsheet automatically completed the aggregations according to the hierarchy of the value tree (Figure 8.2). (Scores, aggregated scores and weights are reported in full in Appendix 8A).

8.1.3 Results

Scenario 5 was generally the most preferred for the criteria relating to employment, and forestry NPV, while Scenarios 1 and 6 were the least preferred from these perspectives. Scenario 1 was preferred for the criteria relating to the environment (both conservation and hydrology) and for NPV of agricultural production. Scenario 6 was least preferred overall (Figure 8.5). Scenario 5a (with conservation constraints) was preferred overall, being marginally preferred to Scenario 5 (no constraints). Scenario 5a is a reasonable compromise, as it is not the ‘worst’ scenario for any category, and is in fact the most preferred option for the social category, the second most preferred for the economic category, and has some conservation value compared to Scenario 5 (Figure 8.5).

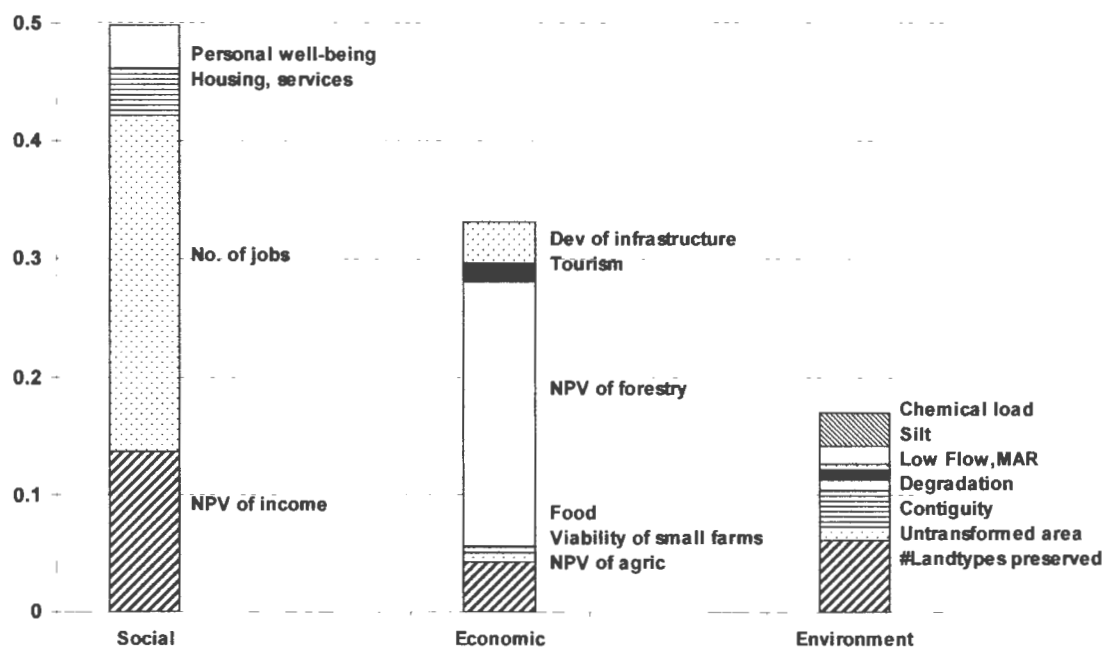


Figure 8.4. Weights of three criterion groups, with contributions by criteria.

However, the economic preferences were not clear-cut due to the conflict between agriculture and forestry. Scenario 5a is in fact the least preferred option from the agricultural point of view. The difference between Scenario 5 and 5a, as far as agriculture is concerned, was based on the argument that, if the forestry company were not allowed to further afforest any of land-types 2, 4, and 9, then they *may* buy farms which are presently used for arable farming and so the NPV of agriculture *may* be reduced. In fact, this may be unlikely as a profitable farm is unlikely to be sold, and so Scenarios 5a and 5 may in fact be equivalent from an agricultural point of view (i.e. both are worst). (The weight assigned to forestry relative to agriculture was based on the range of impacts across the scenarios: this was larger for forestry than for agriculture, and so a larger weight was given.) As an aside, if a pulp-mill were not constructed, another compromise would have to be sought within the original Scenarios 1, 2, 3, 4 and 4a, which do not include pulp-mills. In that case, Scenario 4a would be preferred, and Scenario 1 would be a 'close second'.

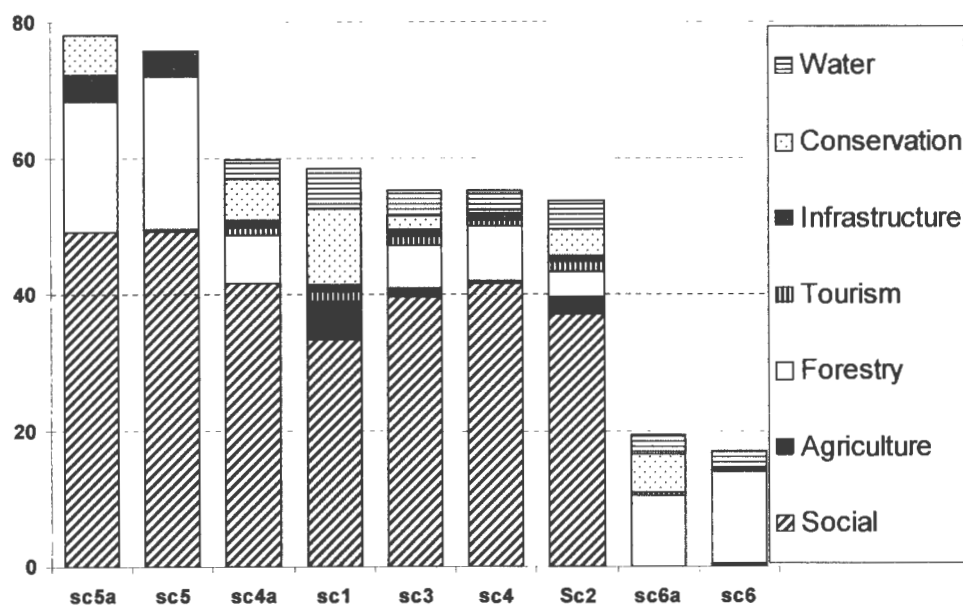


Figure 8.5 Aggregate scores, showing contributions from lower level criteria.

The implication of conservation constraints

Some of the more interesting aspects of the process are revealed by exploring the implicit trade-offs at various levels of the hierarchy. The difference between Scenarios 4, 5 6 and 4a, 5a, 6a is that there are conservation constraints built into 4a, 5a, and 6a (no afforestation on the remainder of land-types 2, 4, and 9). This implies that in order for the forestry company to reach its desired level of afforestation it *may* be forced to afforest on land less suitable for afforestation, which in turn *may* imply increased costs in terms of harvesting, and decreased mean annual increment etc. The workshop participants agreed that the amount of land involved would be approximately 5 000 ha (approximately 10 % of the total afforestation, or a half to two thirds of any new afforestation) and that the mean annual increment could conceivably be reduced from 15 to 12 or 10 m³/ha/a. Considering the worst case of a change to an mean annual increment of 10 m³/ha/a, this loss in production over 25 years would translate to a NPV of about R 16 mill. In order to justify a preference for alternative 5 over 5a, however, the loss in income discounted over 25 years would have to be greater than around R33 mill. Thus, in order to justify *not* adhering to the conservation constraints, the forestry company would have to prove a potential loss of greater than R33 mill. Another way of

considering this, is that the remaining untransformed land types 2, 4, and 9 will have an environmental value of R33 mill discounted over the next 25 years. This value would stem from their present rarity, threatened status, ecosystem services provided, habitat value, existence value, recreation value etc., as indicated by the conservation, hydrology, and social criteria. In the Armstrong and van Hensbergen (1997) study, the importance of these land-types stemmed from the presence of endemics and the rarity of the land-types in the area. As a rough comparison of values, Costanza *et al.* (1997) estimated from various sources, global figures for the value of various habitats in terms of ecosystem services etc. According to them, About 5000 ha of grasslands, at R244/ha/a or R1098/ha/yr, discounted over 25 years, would have an NPV of around R60 mill.

Implied trade-offs

Following the same procedure as in Section 6.3.2 and using NPV of forestry as the standard monetary criterion, the trade-offs and indirect compensatory values shown in Table 8.3 were obtained. The third column from the right shows the value of a 1 point change in the criterion in terms of the NPV of forestry. The last two columns illustrate the value of changing from Scenario 5 (no constraints) to Scenario 5a (with conservation constraints). In this case the NPV of forestry had a linear value function, and so the trade-off does not depend on the attribute level (weights of the criteria were shown in Figure 8.4).

Table 8.3. Monetary value of 1 value point changes in each criterion (indirect compensatory value) and trade-offs between Sc5a and Sc5 (last two columns).

Criterion	Contributing weights			Global Weight	(w1/w2)	Value of 1 point change	Value difference (Sc5a - Sc5)	Rand difference R'000
#landtypes preserved	0.545	0.662	0.169	0.061	0.2713	R 311 970	80	R24 957.0
Untransformed area	0.099	0.662	0.169	0.011	0.049	R 56 670	9	R 510.0
Contiguity	0.279	0.662	0.169	0.031	0.139	R 159 706	25	R 3 992.6
Tourism		0.046	0.332	0.0153	0.068	R 78 139	15	R 1 172.0
							Total gains	R30 632.3
NPV Agric	0.754	0.169	0.332	0.0423	0.188	R 216 455	-8.7	-R 1 882.2
Viability of farms	0.124	0.169	0.332		0.031	R 35 597	-6	-R 213.6
Local food prodn	0.122	0.169	0.332	0.0068	0.0305	R 35 023	-6	-R 210.1
NPV Forestry		0.677	0.332	0.2248	1	R1 150 000	-13.9	-R16 000.0
Degradation	0.077	0.662	0.69	0.0086	0.0383	R 44 076	-10	-R 440.8
							Total losses	-R18 764.7
							Net Gain	R 11 885 6

Sensitivity to weights

Changing the importance weights of some of the criteria also affected the implied cost of not adhering to the conservation constraints as discussed in the previous section. For example, if the weight of the criterion group 'conservation' or the 'environment' category increased, the implied cost of Scenario 5a in terms of lost forestry earnings would decrease. In general the overall preferences, as tested with VISA and Excel, were insensitive to weight changes. However, increasing the weight on the environmental category by 14% made Scenarios 4a and 1 equally preferred (see Figure 8.5 for original scores). Note that the weights of the other two categories would be slightly reduced to compensate as the weights are normalised. If the weight of the economic category were increased by 39 %, then Scenarios 5 and 5a would be equally preferred. Increasing or decreasing the weights on the social category had no significant impact on the overall preferences.

8.1.4 Comments

The AD-SMARTx process appeared to be useful to those involved, and played a role in informing other processes aimed at decision-making at a more 'macro-scale'. The implicit Rand value trade-offs (indirect compensatory values) and the values of scenarios from different points of view were of particular interest. The flexibility of the process was useful, as once a certain level of detail was available, scenarios could be reasonably easily adjusted. Problems encountered in using this approach were more operational than methodological. For example, for various reasons, the four workshops were spread over an extended period of time, and so impetus was lost in-between, and the participants changed. Therefore, not everyone followed the process from beginning to end, and those who did not do so, would be less likely to appreciate the positive aspects. For the most part people were willing to accept others' points of view and direct conflict and disagreement was avoided. Chapter 9 has further discussion of general issues arising from this study.

8.2 Evaluation of supply augmentation and demand management options for City of Cape Town

This SMARTx application is somewhat different to the two AD-SMARTx cases described in Chapter 6 and Section 8.1, in that there were more and predefined alternatives, and we had less involvement in initial problem structuring. Of most relevance to the later comparison of methods are the criteria chosen (Table 8.5) and Section 8.2.1. Predicted water shortages for the greater Cape Town metropolitan area prompted the City of Cape Town municipality to initiate a study in 2000, to identify and evaluate different supply augmentation and demand management options (Table 8.4). As part of this process, people involved in the study (specialists who undertook investigations, stakeholders such as representatives of sub-municipalities and of special interests such as the environment), took part in a series of five day long workshops which made use of SMARTx to obtain an overall 'prioritisation' of these alternatives. The workshops were arranged in five 'themes' with each theme dealing with a particular grouping of issues, namely; yield and technology, financial, socio-economic, public acceptance and buy-in, and the environment. The representatives changed slightly with each theme, with a core group remaining constant throughout all workshops. Another workshop was held a year later to evaluate three additional supply augmentation options. Results were reported in Eberhard and Joubert (2000 and 2001).

The criteria for evaluating the alternatives were pre-determined by the project co-ordinators, but refined and vetted during these workshops (Table 8.5). The participants provided scores (on a 0-100 interval scale) for various levels of the criteria, thereby establishing scoring systems for each criterion, and defining the linear or non-linear nature of the criteria (Table 8.5). The participants then gave scores to the alternatives on the basis of each of the criteria, using the scoring system as a guideline. This was done using the scales as developed in Excel as well as directly in VISA and on flipcharts (the latter proving particularly useful in some cases). The swing weights were determined in a final workshop which included people who had been present at each of the previous workshops, as well as some additional participants.

Table 8.4. Water supply and demand management options

Abbreviation	Description
Options evaluated in 2000	
EerstRiv	A diversion on the Eerste River to create additional water supply.
LourRiv	A diversion on the Lourens River to create additional water supply.
CapeFlats	Extraction of water from Cape Flats Aquifer to create additional water supply. Wellfield would consist of 41 production boreholes and 20 observation boreholes.
PressCont, PressContR	Reduction of pressure in the water distribution systems to reduce water losses, particularly during low water usage periods (at night).
LeakageRep, LeakageRepR	Repair of leaks in distribution system and on private properties in low-income areas.
SubRetrofits, SubRetrofitsR, AFUs	Water efficient fittings/ 'subsidised retrofits' - water service provider provides monetary incentives to households & other consumers for replacing water-inefficient fittings with water-efficient fittings. Removal of automatic flushing urinals evaluated as separate sub-option.
Boreholes, BoreholesR	The more prevalent use of private boreholes, mainly by households, predominantly for the irrigation of gardens.
GreyWater, GreyWaterR	Use of wastewater from the kitchen sink, bathroom, bath, shower, washing machine & dishwasher, predominantly for garden irrigation.
Sew-irrig, Sew-irrigR	The provision of treated sewage effluent to an irrigation area in lieu of obtaining the rights to use the irrigation allocation of fresh water for urban use.
Sew-potable	The treatment of sewage effluent to potable standard for introduction into the potable water distribution network.
Sew-loc/ind, Sew-loc/indR	The supply of treated sewage effluent to local industries and urban irrigation schemes (e.g. sports clubs) for their use as a substitute for potable water.
Tariffs, TariffsR	The adjustment of tariffs by 30% with the view to encouraging greater water conservation. The impact of improved credit control on water use was not evaluated. The impact of metering on water use was not included for inter-option comparison.
UserEd, UserEdR	Impact of user-education on water usage discounting direct effects of above options.
New options evaluated in 2001	
Voëlvlei	Voëlvlei augmentation scheme: Increasing the inflow to Voëlvlei.
TabMtnAqu	Table Mountain Group Aquifer: Extraction of water from the Table Mountain Group (deep) aquifer.
Desalination	Desalination of sea water

Table 8.5. Criteria, criteria levels and associated scores.

Criteria		Description and Units	Levels and Scores				
			Best				Worst
Yield & Technology	Yield (or saving)	Mm3/a	70	50	35	10	1
	Score		100	90	80	40	0
	Time to 50% of yield	Years	1	3		5	7
	Score		100	90		55	0
	Confidence in yield	90% sure that yield will be within x% of estimated yield = X%	5	10	25	50	100
Finance	Score		100	95	79	53	0
	Technological risk/maturity	Probability, type and impact of failure	very low	Low	medium	high	
	Score		100	67	33		0
	URV	Unit reference value (discounted present cost/discounted future sales = R/kl)	0				4.10
	Score (linear)		100				0
Finance	Confidence in cost	90% sure that present cost will be within x% of estimated present cost X%	+ - 10%	+ - 25%		+ - 50%	
	Score		100	80			0
	Capital finance constraint	Is project success likely to be constrained by ability to raise required funds?	Easily obtainable	Obtainable		Constrained obtainability	
	Score		100	67			0
Socio-Economic	Impact on water utility	Tariff increase required for service provider to remain financially neutral = c/kl	-21.5				28
	Score (linear)		100				0.0
	Impact on poor	Impact on monthly expenditure (on < R800/ month households)	positive	neutral	low negative	Moderate negative	
	Score		100		75		0
Socio-Economic	Health	Health	positive	Neutral	low negative	mod neg	
	Score		100	90	30		0
	Impact on employment	Additional permanent employment = numbers	40	30	10		0
	Score		100	75	25		0
Acceptance & Buy-in	Impact on employment	Additional temporary employment = person years	1000	500	200	100	14
	Score		100	49	29	9	0
	Political risk	What is risk project won't receive political support necessary for go-ahead & policy /legislative changes necessary?	none	Low	moderate	high	
	Score		100	75	25		0
Acceptance & Buy-in	Institutional risk	What is risk project will not achieve objectives because of institutional constraints e.g. capacity	none	Low	moderate	high	
	Score		100	75	25		0
	Public acceptance	What is risk that project will not go ahead because of bad image?	none	low	moderate	high	
	Score		100	75	25		0
Environment	Consumer response	What is risk that project will not achieve its objectives because of lack of consumer "buy-in"	none	Low	moderate	high	
	Score		100	75	25		0
	River and groundwater	Nature & significance of impact	positive	neutral	low neg	mod neg	high neg
	Score		100	90	70	40	0
Environment	Estuary	Nature & significance of impact	positive	neutral	low neg	mod neg	high neg
	Score		100	90	70	40	0
Environment	Terrestrial	Nature & significance of impact	positive	neutral	low neg	mod neg	high neg
	Score		100	90	70	40	0

8.2.1 Results

This evaluation provided an initial prioritisation of options (Figure 8.6) which either could be implemented immediately, or warranted more detailed investigations, or which clearly did not warrant further investigation. Nearly a year later, the City of Cape Town wished to evaluate three new supply augmentation schemes (bottom of Table 8.4). These were relatively easily included within the previous analysis by holding a single workshop with representatives of the five themes. At this workshop the new options were given scores by using the previously set up scoring systems as guidelines, as well as the previously scored alternatives. The alternatives included from the previous evaluation were limited to those needed to ensure that there was a 'worst' and 'best' for each criterion. The weights were not adjusted, although the ranges of the criteria changed slightly in the case of the yield, unit reference value and impact on water utility. However, the results were tested for sensitivity to changes in these weights which reflected the change in ranges, and they were found to be insensitive.

The inclusion of the Table Mountain Aquifer option produced the biggest controversy in the study. This was a result of the conflict between those who felt that there was a risk of environmental impact, and those who felt there was low or even no risk. The latter view was based on the fact that there would be a pilot implementation to assess impacts before full implementation, and the former view was based on the paucity of data on, for example, rates of regeneration of the aquifer. Finally, the full range of scores corresponding to these views was included, and the results reflected both ends of these ranges as well as the middle point.

The possibility to examine weight sensitivity was of great interest to the study team. Of particular interest was the examination of the ranking of the options with and without consideration of the group of criteria relating to political and consumer buy-in (Figure 8.7).

8.2.2 Comments

The participants found the process useful, and found it easy to give scores. The elicitation of swing weights was more difficult, but the resulting weights were uncontroversial. The fact that ranges of values could be recorded and included in the analysis of results helped to deflect the controversy over scores for Table Mountain Aquifer, which threatened to derail the last workshop. The City of Cape Town found the process illuminating and have used the results in their plans for future water supply and demand management.

Unfortunately, the decision analysts were only included once the investigations of options were complete or well underway, and the SMARTx process was constrained by the process desired by the City of Cape Town. This led to a number of problems. Firstly, it was sometimes unclear that those who could best report on the particular projects were included, or, on the other hand, why certain other people were present. Secondly, there were more than 15 people in most workshops, making it difficult to run coherent workshops (particular when the reason for some people's presence was not clear). Finally, the options were in fact not mutually exclusive and ideally should have been considered as part of 'portfolios', however this proposal met with strong resistance from the City of Cape Town.

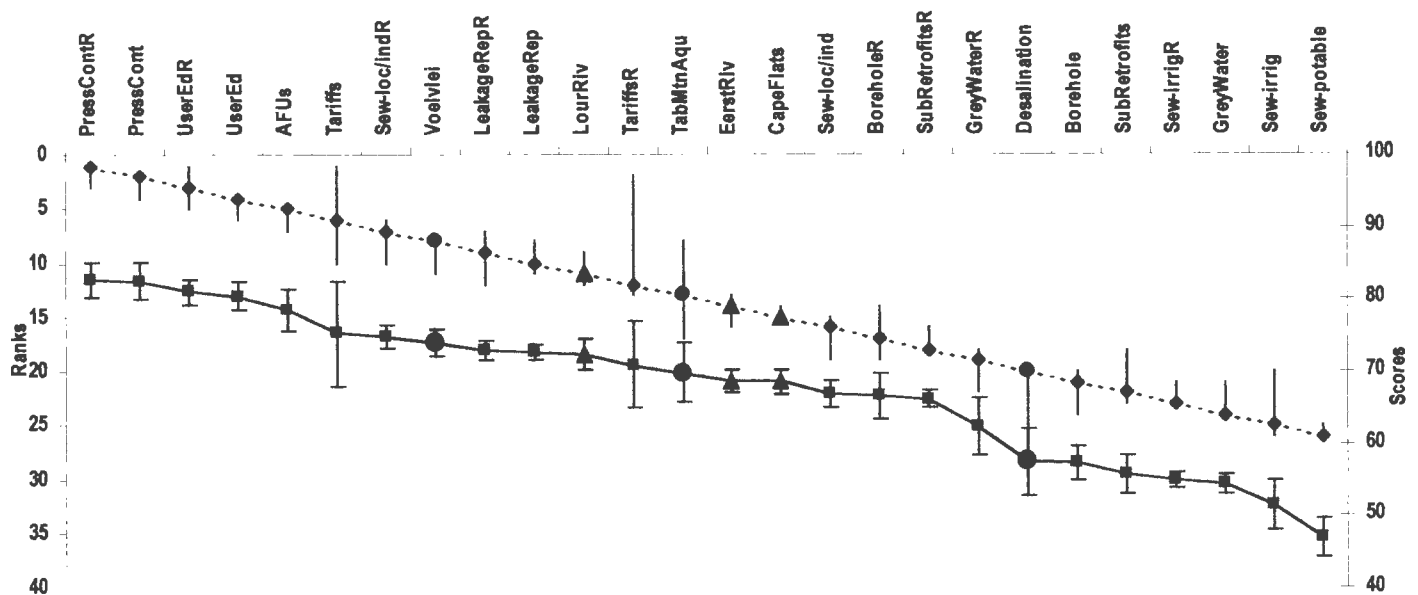


Figure 8.6. Overall scores (solid line) and ranks (dotted line). The standard deviation for scores and range for ranks are shown based on the weights used in the sensitivity analysis. Original supply options are triangles, the new ones circles.

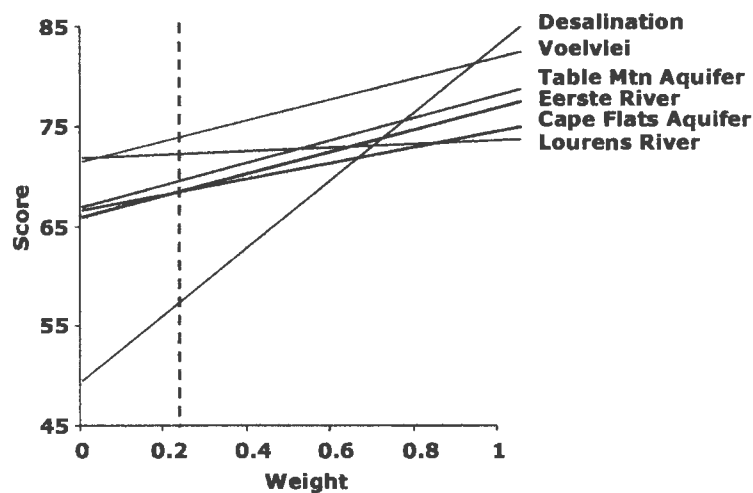


Figure 8.7. Sensitivity of supply option scores to weight on the criteria group 'political and consumer acceptance and buy-in'. The weight given at the workshop is indicated with a dotted line.

8.3 Evaluation of contribution of the Big Five to tourism value of Hluhluwe National Park

This case is different to the other two EE cases in that it is valuing attributes and species, rather than alternatives. However, the methods used were similar to those in Chapter 7. The results for the TCM, CA, CVM and general rating questions are relevant (Section 8.3.1) to the comparison of methods. Game reserve managers have to try to maximise tourist revenues without compromising the ultimate aim of the reserve: the long-term conservation of its biodiversity. The generally accepted view is that only the more charismatic species attract tourists and therefore the revenues necessary to make preservation of biodiversity possible (Daily and Ehrlich 1995). The aim of this study was to examine tourists' reasons for visiting a game reserve and how they apportion the value of their visit to different attributes of biodiversity. The specific aim was to apportion economic value amongst these different attributes including the 'Big Five' (as a group and as separate species), other large mammals and birds, the general diversity of fauna and flora, and the wilderness experience. Many other aspects of this study are not reported here.

The case study was carried out at the Hluhluwe-Umfolozi Park (HUP) in Kwazulu-Natal, an important Big Five game park, renowned for its rhinoceros conservation efforts, and its resultant high concentration of black and white rhinoceros. Section 8.3.1 is of most relevance to the comparisons in Chapter 9.

A combination of methods was used: CA, a zonal TCM, simple ranking and rating, and open-ended CVM. The questionnaire was divided into four sections. Section I of the questionnaire addressed travel costs. Tourists were asked to estimate their total travel costs, the percentage reason for their trip that could be attributable to HUP, and the amount spent in HUP. Total visitor and revenue data was obtained from Kwazulu-Natal Nature Conservation Services. This information was used to calculate tourist WTP (price+consumer surplus) using the zonal TCM method described in Sections 5.2.5, 7.1.1 and 7.2.1).

Section II related to preferences for different attributes using ranking and rating. Tourists were asked to list the three species they most wanted to see (ranked 1 to 3). Next, they were asked to allocate the percentage contribution to their enjoyment of the park of: seeing the Big Five, seeing other large mammals and birds, seeing the diversity of other species, enjoying the scenery and wilderness experience. They were then asked to apportion percentage enjoyment between each of the species of the Big Five (rhino, lion, elephant, buffalo and leopard). If the tourist said that she could identify white and black rhino, she was asked to apportion the importance of each to her visit. Thus, total travel cost could be apportioned between general park attributes, the portion of that attributable to the Big Five could then be apportioned between each species, and where possible the portion attributable to rhino could be apportioned between white and black rhino. These were all analysed according to different groupings: national or international, 'listers' (those who actively mark or take note of which species they have seen) or not. Visitors were also asked if they would visit the park if there were no lion present, and if there were no rhino in the park.

Section III was a CA question which was aimed at further assessing tourists' preferences for game viewing attributes. The CA question consisted of scenarios for evaluation made up of different combinations of animals hypothetically seen. Four 'attributes'; rhinos, lions, other mammals, and birds, were chosen and 3 levels of each were defined except for rhino which had 4 levels. Ten feasible scenarios out of a total of 108 (feasible and infeasible) were defined, and each of three survey versions had four scenarios. Each survey version included the '*status quo*', which was the 'average' tourist sighting suggested by the park conservation officer, and three others. Tourists were asked to give scores to these game-viewing scenarios on a scale of 1 (worst game viewing scenario) to 10 (ideal scenario) (Table 8.6). (A scenario consisting of the numbers of animals actually seen was also included, but is not included in the analyses below.)

Section IV was an open-ended CVM question which asked if the respondent would have been willing to pay more (or would have rather paid less) for entry to the park, considering their actual tourism experience. Then, given the 'best' scenario presented to them in the CA question, they were asked whether they would have been willing to pay more for that scenario.

Table 8.6. The different conjoint analysis scenarios in the three survey versions.

Version	Scenario	Rhino	Lion	Other mammal species	Birds
1.2	Status Quo	5	0	7	50
1.3	Ideal	20	3	15	100
1.4	Intermediate	0	0	2	100
1.5	Intermediate	5	1	7	50
2.2	Status Quo	5	0	7	50
2.3	Worst	0	0	2	20
2.4	Intermediate	10	0	15	50
2.5	Intermediate	10	3	7	100
3.2	Status Quo	5	0	7	50
3.3	Intermediate	0	0	15	20
3.4	Intermediate	20	0	7	20
3.5	Intermediate	20	3	2	20

8.3.1 Results and comment

Section I of the questionnaire provided estimates of revenues, the importance of the visit to HUP to the tourists' trip, and consumer surplus. The total expenditure *within* the park from the sample was R199 940, and estimated for the year (from Kwazulu-Natal Nature Conservation Services) was R33 million. Annual expenditure attributable to HUP was R261 million. This is the total trip expenditure of the sample, adjusted by the percentage importance of HUP to the trip to give expenditure attributable to HUP, this then being extrapolated to an annual figure based on the relationship between the sample and annual HUP revenue: see Section 7.2.1 for method). This gives an indication of the indirect economic effects of HUP in that money is spent in the park as well as in travelling to the park. The TCM analysis (Figure 8.8 shows the initial regression used to develop the demand curve) yielded a consumer surplus of R46.5 million (see 5.2.5, 7.1.1 and 7.2.1 for methods) which is the aggregated individual WTP (or utility) of HUP. As with the KNP study, the consumer surplus calculated was rather sensitive to which countries were aggregated (and therefore also which source populations were relevant).

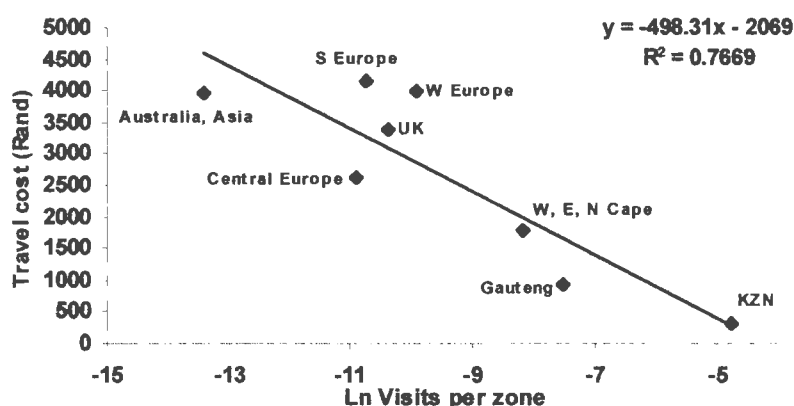


Figure 8.8. Visits and travel cost from different origin zones (used to develop the demand curve for visits to HUP).

Section III of the questionnaire was the CA question. The scores associated with the different CA scenarios were reasonably consistent (i.e. had low standard deviations) (Table 8.7). The average adjusted scores (see adjustment method in Section 7.2.2) were regressed against the attributes using two model formulations equivalent to Model1 (all variables continuous) and Model2 (all variables categorical, therefore using indicator variables) of the KNP study discussed in Section 7.1.3. Table 8.8 shows the regression coefficients from the models, and Table 8.7 shows the scenario scores predicted by the models. Excel was used for the Model1 regression and the generalised linear model module of Genstat (Lawes Agricultural Trust, 1998) (with the normal distribution and identity link options)

was used for the Model2 regression. As expected, the coefficients increase with numbers of animals seen, although level 4 of rhinos is very slightly lower than level 3 (Figure 8.9). This is also not unexpected, because if ‘too many’ of a particular animal are seen, and it is therefore perceived as ‘common’, a species loses its ‘game-viewing value’ (e.g. impala are not highly valued sightings). This analysis therefore suggests that if, on average, tourists saw more than 10 rhino per visit, rhinos would tend to lose their value. The average actual sightings by tourists in this sample was around 5, while the conservation officer also said that 5 would be the expected average rhino sightings. Therefore, the value to tourists would increase if the number of rhino sightings increased from their current level. There were insufficient levels of the other attributes to detect similar decreasing marginal utility trends (Figure 8.9).

The CA was intended to help in determining the importance of different attributes in terms of apportioning total expenditure between them, and of finding the value of increasing the chances of seeing particular species, in particular rhino in this case. As with the KNP study, interpreting this in terms of management actions rests on the assumption that people’s preferences for seeing different numbers and types of species would be reflected in economic behaviour. This would mean, amongst other things, that they would have prior knowledge of what they might see at HUP as well as at other reserves (i.e. substitutes). HUP was well known for its rhino conservation work, so this assumption was reasonable in this respect. However, 93% of respondents said that they would have visited the park if no rhino were present, and 88% said they would if there were no lions present.

Table 8.7. Average and adjusted average from the CA question. Predicted scores from the two CA formulations are shown in the last two columns (these are equivalent to Model1 and Model2 of KNP study).

Scenario	Average Score [Median](StdDev)	Adjusted average	Predicted scores Model1	Predicted scores Model2
1.2-SQ	5.412 [6] (1.743)	5.41	4.99	5.44
1.3-Ideal	9.470 [10] (1.275)	9.47	10.44	9.61
1.4	3.325 [3] (2.229)	3.33	4.08	3.13
1.5	6.742 [7] (1.775)	6.74	5.80	6.72
2.2-SQ	5.348 [5] (1.776)	5.41	4.99	5.44
2.3-Worst	2.652 [2] (1.654)	2.72	3.62	3.04
2.4	6.957 [7] (1.904)	7.02	6.56	7.11
2.5	8.935 [10] (1.932)	9.00	8.28	9.01
3.2-SQ	5.573 [6] (2.388)	5.41	4.99	5.44
3.3	4.656 [5] (2.797)	4.49	5.20	4.36
3.4	6.607 [7] (1.986)	6.45	6.60	6.42
3.5	8.467 [10] (1.941)	8.31	8.40	8.19

Table 8.8. Regression results for conjoint analysis using equivalent of Model1 and Model2 of KNP study.

	Model1 $r^2 = 0.46$				Model2 $r^2 = 0.51$		
	Coefficients	Standard error	P value	Scores for levels	Coefficients	Standard error	P value
Constant	3.262	0.176	<.001	3.262	0		
Rhino1 (0)	0.119	0.012	<.001	0.000	6.951	0.299	<.001
Rhino2 (5)				0.593	8.536	0.324	<.001
Rhino3 (10)				1.185	9.606	0.211	<.001
Rhino4 (20)				2.371	9.605	0.196	<.001
Lion1 (0)	0.803	0.075	<.001	0.000	-2.496	0.211	<.001
Lion2 (1)				0.803	-1.22	0.321	<.001
Lion3 (3)				2.409	0.00	*	*
Other species1 (2)	0.121	0.014	<.001	0.243	-1.324	0.198	<.001
Other species2 (7)				0.849	-0.597	0.211	0.005
Other species3 (15)				1.820	0.00	*	*
Bird1 (20)	0.005761	0.003	0.02	0.115	-0.091	0.197	0.643
Bird2 (50)				0.288	0.00	*	*
Bird3 (100)				0.576	0.00	*	*

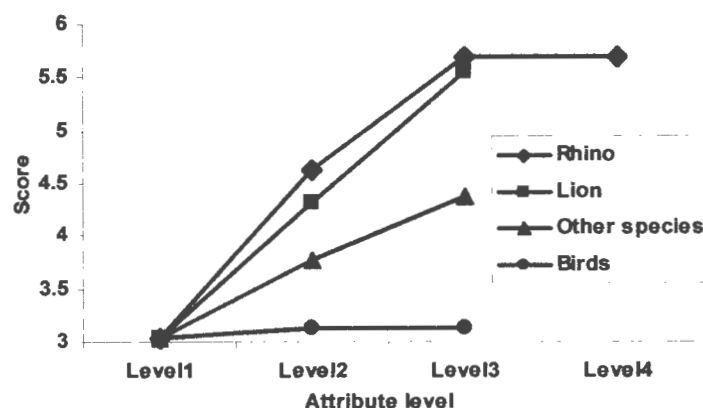


Figure 8.9. Scores for different levels of the four attributes in the CA question using Model2. The scores for levels 2, 3 and 4 are added to the base score or intercept of 3.04. This is found by adding the level 1 score for each attribute in the first column of Model2 in Table 8.8.

The respondents were not particularly comfortable with the CA question, although they did not seem to object to giving scores rather than ranks. In contrast, respondents were more comfortable with the questions where they were simply asked to state the percentage importance to them of different park attributes and species (Section II of the questionnaire). The results of this section were generally consistent with what was expected (Table 8.9), and with the CA results where they corresponded. For example, three of the questions revealed that seeing lions was as important as seeing rhinos to the game-viewing experience. Buffalo, although one of the Big Five species, were given relatively much less value. Overall, the general wilderness experience was the second most important attribute after 'seeing the Big Five'.

Table 8.9. Percentage importance of different attributes from direct questioning and from the CA question.

	Question12			Question13		Question14		Question17 CA (% from range of Model2 coefficients)
	Which animals did you most want to see?			What % importance were the following?		What % importance each of Big Five		
	first	second	third	Average	Median	Average	Median	
Big Five				37.7	30			
lion	28.7	21.5	13.0			23.8	20	38.0
rhino	27.6	14.0	14.3			21.4	20	40.4
elephant	18.1	28.7	18.4			19.2	20	
Leopard	11.26	12.6	16.4			25.3	20	
Buffalo	0	0.68	0.68			10.3	10	
Large mammals &birds				17.2	20			
General diversity				16.0	10			
Wilderness				29.0	25			
Other mammal species								20.2
Birds								1.4

Section IV, the open-ended CVM, revealed that on average tourists would have been willing to pay 40% more (R10) than the current park entry fee, given their actual viewing experience, while they said they would have been willing to pay 90% more (R20) for the ideal scenario (as presented in their particular survey version). These percentage increases are based on their median willingness to pay. Total WTP for the sample would increase from the current total entry paid by the sample (R6 700) to R11 600 (for the actual experience) and to R14 900 for the best scenario (the latter two figures are sums of all the individual WTP offers). The WTP for the actual viewing experience and for the best scenario in each version showed a good relationship to the average score for each scenario (Figure 8.10). The relationship between scores and levels of the attributes (numbers of particular animals seen) and the relationship between scores and WTP could potentially be used by management to inform decisions regarding stocking rates of certain species (or regarding the potential value of invasive bush-clearing projects to improve game-viewing).

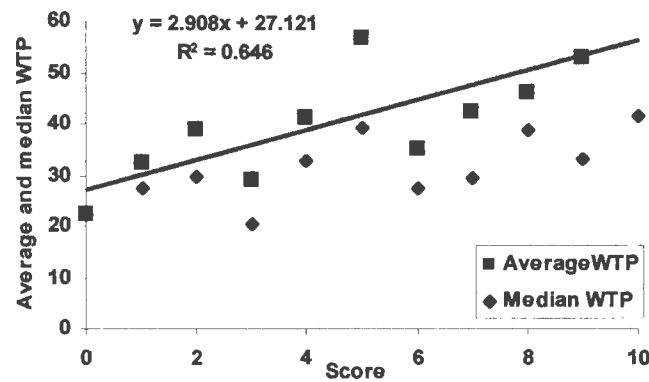


Figure 8.10. Relationship between score for actual game seen, the highest scoring scenario and average and median WTP. The regression line is the relationship between average WTP and the scores.

8.4 Evaluation of different types of urban open space using CA, CVM and ranking and rating.

The City of Cape Town were interested in exploring and demonstrating the use of environmental and resource economics methods for the assessment of the value of open space to assist in their planning and decision-making. Sections of this study are reported here, although much of the detail is omitted. The full report is given in Turpie *et al.* (2001). Section 8.4.1 is relevant to the later comparison of methods. This case study is similar to that in Chapter 7 although the landtypes could not realistically be considered alternatives. Nevertheless, the intention was to compare their relative aggregate values.

Two study areas, Metro South and Metro South-east, were selected by City of Cape Town which contained six different types of open space (parks, sportsfields, areas with natural fynbos vegetation¹⁰, vacant lots, wetlands and rivers, agricultural fields). Three surveys were undertaken, one of which, the general survey, looked at 'total economic value' in both study areas using:

- direct questioning regarding types of values, scores for different types of open space, and rank order of different scenarios of total amounts of open space and rates combinations,
- an open-ended CVM-WTP question for different scenarios made up of a combination of different total areas of open space and different potential rates increases, and
- a CA question regarding changes in use with changed levels of crime and cleanliness.

A survey with the same format as the general survey was undertaken on a specific wetland area, Kuilsriver, in Metro South-east, and another survey undertook a TCM analysis of Zandvlei which is a wetland within the Metro South study area. The first enabled us to look at 'embedding' and 'scale' effects (Section 5.2.4), while the second enabled us to compare values from different methods. The comparison of values from different methods is not discussed here.

The source populations included a wide range of race, income and education. The respondents found the CA question difficult to process and answer reliably (and for this and other reasons this question is not further discussed

here), but were reasonably comfortable giving ‘scores’ to different types of open space. We therefore concentrate here on the CVM question. This proposed a scenario in which the City of Cape Town were reaching a stage where they could no longer afford to retain the current amount of open space areas (of the types mentioned above), and therefore, with no cash injection, they would have to sell or allow development (i.e. building) on 50% of the total open space area. One alternative to this would be to increase the rates paid by home owners, and use the rates increase to help retain open space. The amount that could be retained would depend on the amount by which rates were increased. Residents were asked if they would be ‘concerned’ if 50% of open space were to be built on. They were then asked if they would be willing to pay more rates to avoid this situation and retain different amounts of open space. The options presented to them were Option C (retain 75% of open space and pay a rates increase), Option B (retain 90% of open space and pay a larger rates increase), and Option A (retain the full amount of open space and pay a larger rates increase). They were also asked to rank the four options in order of preference (this included Option D which was to retain only 50% of open space and therefore have no rates increase).

8.4.1 Results and comments

There were several inconsistencies in the responses to the CVM related questions. Firstly, of the 68% that said they would be concerned if City of Cape Town were to choose option D (half of open space put to other use therefore no increase in rates needed), 40% still chose option D as their preferred option. This probably indicates that people disapproved of the notion of increased rates (about which there was a politicised and vociferous debate at the time), rather than disapproved of open space. Secondly, although most people gave WTPs that increased with increasing amounts of open space, 39% gave lower WTPs for more open space. The reasons could also be related to rates rather than open space (i.e. more open space was equivalent to more rates). These first two issues relate to the well-known CVM problem of ‘vehicle bias’. However, other ‘vehicles’ such as a payment in the form of a donation seemed unrealistic in this setting. Thirdly, 6% of respondents specifically gave $WTP_A < WTP_B$, or $WTP_B < WTP_C$, yet said that $A \succeq B$ or $B \succeq C$, respectively. Fourthly, there was no relationship between given WTP and given income. Although this might be argued to disprove one of the arguments against using WTP, namely that WTP reflects ability rather than willingness to pay, it is unlikely to be ‘real’ in this case, given the generally high poverty levels, and highly skewed income levels within and between the study areas. Fifthly, in Metro South, a relatively wealthy area, 40% preferred option D, whereas in Metro South-east, a relatively poor area with housing shortages, 59% preferred option D. This seems to contradict the previous point.

Therefore, in combination, these issues mean that we are unable to tell if differences in preferences or WTPs given are because those in Metro South-east (a) are unable to pay higher rates to maintain the current amount of open space because of budget constraints (they are poorer), (b) object more strongly to the notion of rates hikes (the rates issue was strongly to do with poorer areas paying inequitable rates), (c) are more concerned with having areas made available for housing (this area is overcrowded and has large squatter areas), (d) are more concerned about crime associated with open space, or (e) are less interested in open space. The answer is probably a combination of all of

¹⁰ Fynbos, literally ‘fine bush’ is the indigenous vegetation of the Western Cape, part of the important Cape Floral Kingdom.

these factors. It is also noteworthy that during this project several open space areas in Metro South-east were “invaded” by residents of Khayelitsha. This was due to the general overcrowding in Khayelitsha, exacerbated by flooding. This situation would tend to make people even less sympathetic towards the idea of a rates hike.

Note that the median WTP was zero in all cases (except for a median of R5 for Option C in Metro South). Some economists or CVM theoreticians and practitioners prefer the use of the median rather than the mean WTP (see also Section 9.4.1 and Figure 9.3). By implication, this would have meant that loss of 50% of open space was of ‘no value’ to respondents in these two areas.

Another ‘classic’ CVM problem is that of embedding or scale. This arises when an environmental good is part of a larger good. This was illustrated in the valuation of the Kuilsriver wetland, which is in Metro South-east. The average WTP to retain all of Kuilsriver wetland when evaluated in a separate survey, was 60% of the WTP for all of open space in Metro South-east obtained from the general survey (Figure 8.11). The average WTP to retain 50% of Kuilsriver wetland was 86% of the amount to retain 50% of all open space. This is in contrast to the fact that in the general survey of Metro South-east residents had indicated that wetlands contributed between 11% and 20% of the total value of open space as indicated by scoring the different types of open space directly (in three different questions, results not shown). In our case this means that the apportioned mean WTP of wetlands from the general survey of Metro South-east is between R19 and R32 as compared to the survey dedicated to Kuilsriver which found a mean value of R109. The embedding effect is often reported as having effects that change WTPs by an order of magnitude (e.g. Kahneman and Knetsch 1992).

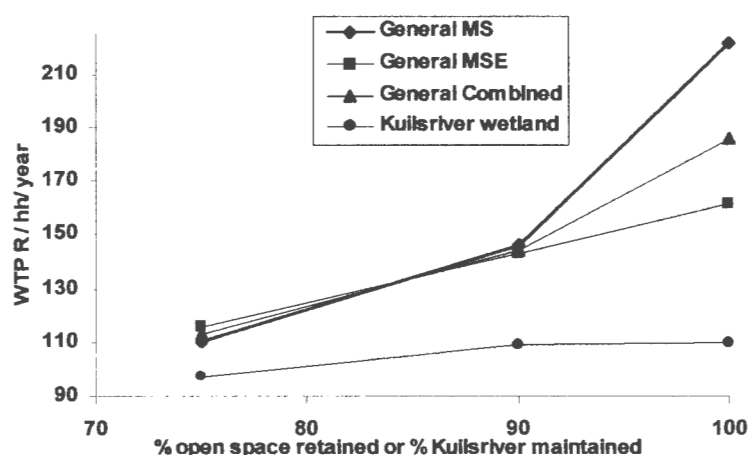


Figure 8.11. Average WTP for different amounts of open space from the general survey for Metro South and Metro South-east and for different amounts of Kuilsriver wetland area within Metro South-east.

The fact that the study could not resolve all of these question was largely to do with its nature as a ‘pilot’ study, which the City of Cape Town wanted to use for a multitude of general questions and in a general sense to illustrate the potential use of EE. Therefore the survey included a number of complex questions, each one of which could have been a survey on its own. However, inescapable context was that the issue of rates and rates increases was highly politicised and emotive, and there is a severe housing shortage in Metro South-east. Thus, people would react strongly against the prospect of rates hikes particularly when no specific piece of open space was being ‘threatened’.

More focused questionnaires might have allowed for questions regarding the trade-offs between open-space of specific types and housing, perhaps using conjoint analysis. Secondly, with the severe housing shortage, many expressed the view in various ways that they would be prepared to give up some open space in order to have land available for housing, but the survey did not offer a trade-off between specific open space and housing. These issues were pointed out to the City of Cape Town before the study as they placed severe constraints on survey design.

8.5 Summary and Conclusions

The Maclear study illustrated the use of AD-SMARTx in a similar context to that of the Sand case in Chapter 6, with similar possibilities for analysis of results. Problems with this study related mainly to the fact that the workshop participants were not constant throughout the study. The analysts were part of all phases of the AD-SMARTx process. Importantly, the process allowed for the progressive development of alternatives and construction of values and preferences. The second study illustrates the use of SMARTx to evaluate water supply and demand management options. The problems were also related to participants in the workshop, but in this case, there were too many and their roles or constituencies were not always clear. The analysts were not involved in the full problem structuring phase of this application, and we were essentially only included in the evaluation workshop phase.

The first EE case valued different attributes of the Hluhluwe-Umfolozi game park. As was the case in the Kruger National Park study, the TCM was sensitive to data arrangement. The second EE case evaluated different types of open space in the City of Cape Town. The problems related to vehicle bias, embedding or scale bias, and to a lack of focus in the survey on specific questions allowing for respondents to make 'real' trade-offs. Median WTP was generally zero, but there is no straightforward interpretation for this or any of the other results due mainly to the problems of vehicle bias. There were also many problems with the complexity of the questions, as many responses showed a lack of understanding of the question. In both EE cases, the respondents were comfortable with scoring and ranking options, but found the CA questions rather difficult.

In Chapter 9 the two main case studies are compared with reference also to the supportive or contradictory information arising in the four cases described in this chapter.

Appendix 8A. Scenario descriptions, consequences and scores for Maclear forestry case study

Economic criteria

NPVs were calculated for the 25 years 1997 to 2021 (only those for discount rates of 6% are shown). 'Scores are values given on a 0-100 scale, either directly or via value function. Where a value function was used the original data is shown.

Category		Agriculture			Forestry		Tourism Potential	Development of infrastructure
Criterion	NPV of agricultural production		Viability of small farms	Local food production	NPV of plantation & primary processing			
Scenario	Criterion level (Rx10 ⁶)	score	score	score	Criterion level (Rx10 ⁶)	score	score	score
Scene 1	98	100	100	100	79	0	100	25
Scene 2	88	57	?	?	98	17	?	?
Scene 3	81	26	?	?	111	28	?	?
Scene 4	77	9	6	6	119	35	65	30
Scene 5	77	9	6	6	194	100	0	100
Scene 6	77	9	6	6	148	60	10	0
Scen 4a	75*	0	0+	0*	115++	31++	80	30
Scen 5a	75*	0	0+	0*	178++	86++	15	100
Scen 6a	75*	0	0+	0*	132++	46++	25	0
Weight		0.754	0.124	0.122				

* Not actually determined, but assumed that some more profitable (arable) land may be used for forestry if forestry constrained not to be on land-types 2, 4 and 9.

+ Similar assumption to *.

++ A calculation made after the fourth workshop, which assumes that the MAI, on 5000 ha is reduced from 15 to 10 m³/ha/a, as forestry is constrained not to be on land-types 2, 4, and 9. This means that forestry *may* go to land that is less favourable and MAI *may* consequently be reduced. These values are used in all other analyses.

Criterion group	Agriculture	Forestry	Tourism Potential	Development of infrastructure	Overall-Economic
Scene 1	100	0	100	25	24.2
Scene 2	48.76	16.52	90	26	26.37
Scene 3	23.35	27.83	85	30	29.93
Scene 4	8.03	34.78	65	30	31.14
Scene 5	8.03	100	0	100	79.86
Scene 6	8.03	60	10	0	42.44
Scen 4a	0	31.30	80	30	28.11
Scen 5a	0	86.09	15	100	69.77
Scen 6a	0	46.09	25	0	32.35
Weight	0.17	0.68	0.046	0.108	

Social criteria

Criterion	Employment				Aggregate Employment	Housing and services score	Personal well-being Score	Overall - Social
	Employment numbers in 2021	Remuneration NPV						
	Criterion level (#)	Score	Criterion level (Rx10 ⁶)	score				
Scene 1	4974	86	267	20	64.7	70	90	67
Scene 2	5034	92	274	32	72.5	?	?	74
Scene 3	5082	96	280	43	79	?	?	79.5
Scene 4	5109	99	283	48	82.4	80	100	83.5
Scene 5	5123	100	312	100	100	100	80	98.5
Scene 6	4043	0	256	0	0	0	0	0
Scen 4a	5109	99	283	48	82.4	80*	100+	83.5
Scen 5a	5123	100	312	100	100	100*	80	98.5
Scen 6a	4043	0	256	0	0	0	0	0
Weight		0.677 [#]		0.323 [#]	0.846	0.081	0.073	

* It is assumed that the mills create extra housing and services to more than compensate for losses through displacement from farms taken over for forestry.

+ Saw mill has greater potential for local training, and wood is more obviously a benefit locally than pulp. Pulp mill will also create benefits through training, but will be more likely to use trained people from outside the area.

[#]The scores for numbers employed and remuneration are summed with these weights to give an overall 'employment' score.

Environmental criteria group

Conservation Criteria group

Criterion	Number of land-types preserved Score	Untransformed area Score	Contiguity Score	Degradation Score	Aggregate Conservation
Scene 1	100	100	100	100	100
Scene 2	?	?	?	?	34.6
Scene 3	?	?	?	?	17.7
Scene 4	5	0	0	10	3.5
Scene 5	0	0	0	10	0.77
Scene 6	0	0	0	10	0.77
Scen 4a	85	9	25	0	54.2
Scen 5a	80	9	25	0	51.5
Scen 6a	80	9	25	0	51.5
Weights	0.545	0.099	0.279	0.077	0.66

Criterion level for Number of land-types preserved (Frequency of occurrence-types 1-10)										
Land type #	Land type	Scene1	Scene2	Scene3	Scene4	Scene5	Scene6	Scene4a	Scene5a	Scene6a
1*	Mll	72	61	58	58	58	58	Not specifically determined		
2*	Mlh	6	5	4	4	4	4	6	6	6
3	Ell	32	25	18	18	18	18	Not specifically determined		
4*	Elh	9	6	3	3	3	3	9	9	9
5	Eml	11	11	11	11	11	11	Not specifically determined		
6	Emh	15	14	13	13	13	13			
7	Ehh	7	6	6	6	6	6			
8*	Chh	14	14	14	14	14	14			
9*	Bhl	9	9	8	6	6	6	9	9	9
10*	Bhh	19	18	17	14	14	14	Not specifically determined		

Criterion level for Untransformed Area of Land-types 1-10: ha [Percentage]										
Land type #	Land type	Scene1	Scene2	Scene3	Scene4	Scene5	Scene6	Scene4a	Scene5a	Scene6a
1*	Mll	49905 [60.3]	43172 [52.1]	40064 [48.4]	40064 [48.4]	40064 [48.4]	40064 [48.4]	Not specifically determined		
2*	Mlh	2763 [56.2]	2386 [48.5]	2098 [42.7]	2098 [42.7]	2098 [42.7]	2098 [42.7]	2763 [56.2]	2763 [56.2]	2763 [56.2]
3	Ell	17821 [41.9]	14409 [33.9]	10795 [25.4]	10795 [25.4]	10795 [25.4]	10795 [25.4]			
4*	Elh	4156 [58.3]	2559 [35.9]	1124 [15.8]	1124 [15.8]	1124 [15.8]	1124 [15.8]	4156 [58.3]	4156 [58.3]	4156 [58.3]
5	Eml	6009 [96]	6009 [96]	6009 [96]	6009 [96]	6009 [96]	6009 [96]			
6	Emh	8548 [87.3]	7666 [78.3]	7185 [73.4]	7185 [73.4]	7185 [73.4]	7185 [73.4]			
7	Ehh	3557 [57.1]	2920 [46.8]	2920 [46.8]	2920 [46.8]	2920 [46.8]	2920 [46.8]			
8*	Chh	11217 [95.5]	11217 [95.5]	11217 [95.5]	11217 [95.5]	11217 [95.5]	11217 [95.5]			
9*	Bhl	7800 [70.6]	7800 [70.6]	6961 [63.0]	4560 [41.3]	4560 [41.3]	4560 [41.3]	7800 [70.6]	7800 [70.6]	7800 [70.6]
10*	Bhh	16797 [89.6]	16175 [86.3]	15663 [83.6]	12071 [64.4]	12071 [64.4]	12071 [64.4]			

Hydrology criteria group

Criterion	Water quantity				Water quality			Hydrology aggregate	Environment total (Water+Con- servation)
	Reduction in MAR Score	Reduction in low flows (%)		Water Quantity aggregate	Silt load Score	Chemical load Score	Water Quality aggregate		
		Criterion level	Score						
Scene 1	100	6	100	100	100	100	100	100	100
Scene 2	?	9	92	90.676	?	?	66.34	72.33	47.34
Scene 3	?	11	87	82.366	?	?	56.34	62.74	32.94
Scene 4	60	12	85	68.45	0	63	41.706	48.29	18.63
Scene 5	0	20	0	0	0	0	0	0	0.51
Scene 6	40	12	66	55.21	0	66	43.692	46.53	16.24
Scen 4a	60	12	85	68.45	0	63	41.706	48.29	52.19
Scen 5a	0	20	0	0	0	0	0	0	34.07
Scen 6a	40	12	66	55.21	0	66	43.692	46.53	49.8
Weight	0.662	0.338			0.338	0.662			
Weight							0.745	0.34	

PART III - DISCUSSION AND SYNTHESIS

9. Comparison of the methods applied in the case studies

"The notion of 'revealed preference' is unclear and should be abandoned. Defenders of the theory of revealed preference have misinterpreted legitimate concerns about the testability of economics as the demand that economists eschew reference to (unobservable) subjective states."
Hausman 2000, p. 99.

"As the choice of embedding structure is arbitrary, the estimates of value obtained from CVM surveys will be correspondingly arbitrary.... The amount that individuals are willing to pay to acquire moral satisfaction should not be mistaken for a measure of the economic value of public goods."
Kahneman and Knetsch 1992, p. 68.

MICDA and EE valuation tools have been widely used in Europe and the USA in water resource management. There has been more limited or experimental use in developing countries, but EE, in particular, is becoming increasingly popular. The potential future formal and informal use of these techniques in South African water management is being investigated in several research programs, the two main case studies having contributed to and been funded by two of these. The two main cases provided realistic management context examples, as both were conducted in very real world situations, where time and budgets were limited, but time in particular was a constraint. The first case (the Sand study) took place in the Sand River catchment which is a tributary of the Sabie river which is, in turn, a tributary of the Crocodile River. The section of the Crocodile river and its tributaries flowing through and along the southern Kruger National Park was the setting for the second study (the KNP study) (Figure 9.1). This chapter compares the case studies discussed in Chapters 6 and 7, with reference also to those in Chapter 8, in terms of the overall framework and the valuation methods applied. The aim is to discuss the issues arising in the case studies in the light of the relevant theory, literature and background given in Part I. A general framework for the methodologies is provided in Sections 9.1 and 9.2 followed by brief summarising comments on the SMARTx (Section 9.2.1) and EE (Section 9.2.2) studies. The metacriteria for comparison are introduced in Section 9.3. Each of the sections that follow compares the methods on the basis of these metacriteria and ends with brief recommendations arising from that section. The discussions and recommendations are summarised in Section 9.8.

9.1 Social choice and environmental decision-making

Overall, a population with a certain level of social welfare and environmental quality, has unknown preferences for environmental quality and equity, etc. Their preferences for various goods and services may be revealed through their repeated choices in the market and through voting, or may be constructed in debate (Figure 9.2). Faced with alternative projects with benefits and costs, we want to know how to compare them in order to choose one which, overall, will improve social welfare. Note that the definition of welfare is rather blurry depending on the application or disciplinary background – choice, preference, WTP, utility and welfare are often used interchangeably (Sen 1997).

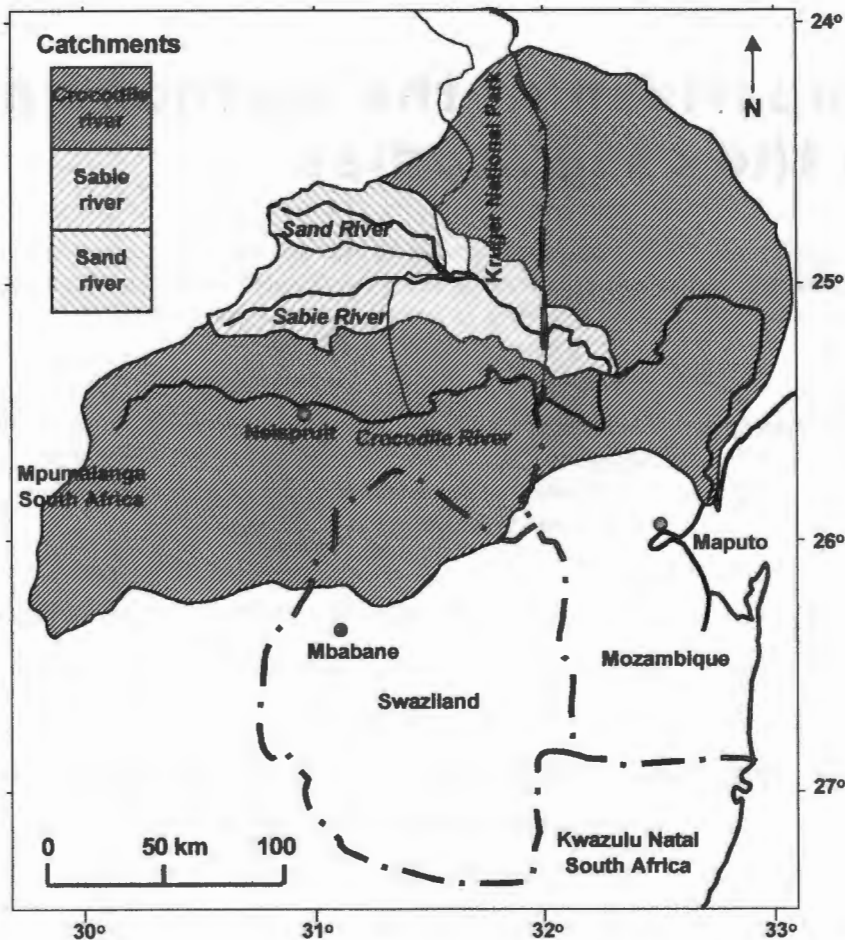


Figure 9.1. Inkomati / Crocodile catchment showing the study areas for the Sand and KNP studies.

CBA and SMARTx are practical implementations of social choice theory, providing different responses to a 'simple' problem: when one party benefits, but another loses, can we test to see if society as a whole is better off? In other words, if the Pareto criterion is not satisfied because there are those whose welfare is diminished, how do we make the interpersonal utility comparisons necessary to tell if welfare as a whole has improved. In Chapter 3 we mentioned the search for appropriate criteria: this search is pursued differently with EE and MCDA decision-making frameworks, with particular implications for developing world settings. Both are rooted in utilitarian precepts aimed at the implementation of the pre-defined moral or ethical goal of maximisation of social *welfare*. Both frame this choice in terms of choosing the option that maximises aggregate *utility* (defined in some way, e.g. choice or WTP, and perhaps subject to some constraints). To this extent, they therefore have a common philosophical basis, but at this point, the two paradigms take fundamentally different routes in terms of implementing this philosophy. (Note that the term utilitarian is used broadly and inclusively here without implying any particular of the different forms of utilitarianism and is not limited to the notion of the selfish utility maximiser dubbed *homo economicus* commonly assumed in economics (and in fact assumed by Bentham)). The simple notion of utility maximisation is one with which philosophers, economists and others have found cause for concern, leading to Sen's critique of the one-dimensionality of 'welfarism' (Graaff 1971, Sen 1979, Massam 1999). A reductionist approach to decision-making is clearly necessary, whatever form it may take, due to the limited human capacity for processing information simultaneously (Miller 1956). The questions remain: how is the reduction of dimensions to occur (what algorithm, what process, what framework), who controls the reduction (an analyst, politicians, stakeholders, lobby groups), and

how far is it necessary or desirable to reduce the complexity of a problem. Pearce (1983) simplified this to two questions: whose preferences should count and how should preferences be weighted.

One may feel that, generally, monetary value or WTP (a measure of consumer surplus) reflects potential impacts on people's preferences or utility on comparable scales, which when aggregated reflect social welfare and that therefore CBA is appropriate. We then need to measure all impacts in monetary terms. Some values of environmental goods and services may be revealed by the process of repeated choices in the market (or by voting), for example, tourism preferences (which might imply something about environmental quality) (Figure 9.2). EE surveys sample the population, and provide hypothetical markets where necessary in order to estimate the unknown preferences (e.g. CVM), reflected by WTP, or use revealed preferences to estimate WTP (e.g. TCM). The various values are added and the B/C ratio calculated. If the B/C ratio is greater than one then a single project can get the go-ahead, because this will pass the Hicks-Kaldor criterion (those who benefit can compensate those who lose and have some benefit left over). If there are several projects, then one needs to look at the relative sizes of the B/C ratios and / or NPV values. At this point, the relative sizes of the B/C ratios supply a rank order of projects. The weighted summation of CBA is therefore the equivalent of value-measurement with mutual preferential independence, which produces multi-attribute value functions interpretable on ordinal scales.

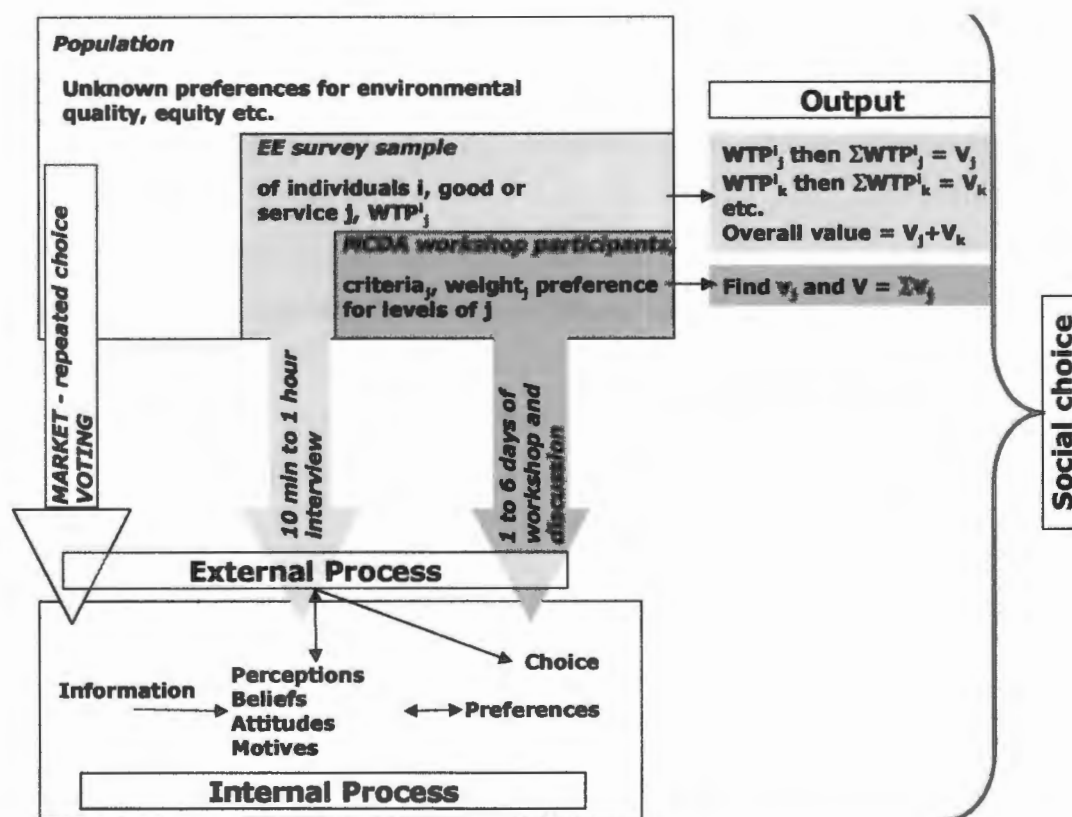


Figure 9.2. Social choice through the EE valuation and MCDA paradigms (the lower left box is adapted from McFadden 1999).

On the other hand, one may not want to limit the criteria to those measurable on a monetary scale and may use SMARTx instead. Rather than finding value revealed in the market or hypothetical markets with surveys, SMARTx uses a small number of participants in workshops to investigate values or utility directly (Figure 9.2). The weighted summation gives the overall relative worth of the projects, and depending on the degree of independence of partial

values this may be on an ordinal scale (as with CBA) or interval (Section 2.2.5). Depending on the formulation of criteria, results could also be expressed as B/C ratios.

The brief introduction here and the framework in Section 9.2 give the impression that there are perhaps only minor differences between the two approaches. However, ‘the devil is in the detail’ - we are now ready to analyse the differences and similarities in more detail.

9.2 Framework for and summary of methods

The Sand study used an AD-SMARTx framework, and the KNP study used a combination of EE tools, namely the travel cost method, contingent behaviour valuation and conjoint analysis. Both studies were based on the assumption that overall value or utility of an alternative river health scenario or management class a is a function of its attributes and our interpretation of or relationship to these attributes. Within a CBA framework, the overall value of river quality a (or management class a) would be:

$$U(a) = \sum_{j=1}^n B_j(a) - \sum_{k=1}^m C_k(a), \quad (9.1)$$

where B_j and C_k are monetised benefits and costs, derived from the utility of each (e.g. U_{B_j} is the utility of benefit j).

The value provided by tourism U_{B_T} would be a function of attributes relevant to the tourist i :

$$U_{B_T}^i = f(Q, K^i, Y^i), \quad (9.2)$$

where Q is the set of environmental attributes, K^i the individual's attributes and Y^i the individual's income.

Ultimately these values are aggregated to give overall tourism value (a welfare measure of B_T):

$$B_T = f(U_{B_T}^1, U_{B_T}^2, \dots, U_{B_T}^n), \quad (9.3)$$

In the case of the KNP study, we used the travel cost method where the $U_{B_T}^i$ are reflected in consumer's surplus (i.e. willingness to pay of individual i , WTP^i) and are summed. In other words, the use-value part of the tourism value of a river health scenario is the sum of WTP^i or $U_{B_T}^i$. Scenarios other than the *status quo* were valued via the contingent behaviour valuation and conjoint analysis questions. Thus, equations (9.1), (9.2) and (9.3) apply to the KNP study. In the open-ended contingent valuation method (the two EE studies in Chapter 8), $U_{B_j}^i$ is also assumed to be reflected directly by WTP and summed.

In the case of the Sand study, the value of a river health scenario a (resulting from a land- and water-use scenario) is given by the SMARTx equation:

$$V(a) = \sum_{j=1}^n w_j v_j(a), \quad (9.4)$$

Where each v_j is a function directly of an attribute x_j or indirectly of one or more attributes X : $v_j(a) = f(X(a))$. Equation (9.4) applies also to the evaluation of alternatives in the two SMARTx cases in Chapter 8. (Note that $w_j v_j$ are now group values, originally expressed as $k_j g_j$ (see Sections 4.1.3 and 5.3.3).

The two main cases were specifically aimed at evaluating alternative river health scenarios to maximise aggregate societal utility, either represented by equation (9.1) (EE) or equation (9.4) (SMARTx). In the Hluhluwe study

(Section 8.3) a particular U_j was being examined (equations (9.2) and (9.3)) which could be used to inform decisions about stocking and exotic bush-clearing. In the open space study (Section 8.4) several U_j were examined (not all reported here), which could inform decisions about which types of open space might be developed or should be retained (based on whether total economic value of an open space type were greater than its real estate price) (therefore equations (9.1), (9.2) and (9.3) apply).

Summaries of the essential features of the applications are given below, followed by a general discussion and comparison of the main issues in the light of the background in Part I.

9.2.1 Evaluation of alternatives with AD-SMARTx and SMARTx

In the MCDA case studies the AD-SMARTx or SMARTx interventions consisted of:

- Workshop facilitation during a series of four or five workshops,
- Providing a relatively structured way of designing alternatives and of selecting and structuring criteria (and general problem structuring),
- Providing a structured approach to evaluating alternatives on the basis of each of the criteria separately [$v_j(a)$ or $v_j(x_j(a))$], with the understanding that values and preferences are both ‘constructed’ and ‘revealed’ and in this case, developed through the group-value sharing model,
- Providing a theoretically valid approach (if the assumptions are met) to the aggregation of the separate criteria evaluations (simple weighted addition in the form of equation (9.4)) to obtain an overall ranking or rating $V(a)$,
- Data analysis to determine overall scores, sensitivity to weights, trade-offs and indirect compensatory values, and
- Provision of feedback to participants.

Some of the assumptions or requirements of the SMARTx applications were that:

1. The criteria were at least mutually independent and perhaps difference independent (see Section 2.2.5),
2. That the group values found by applying the group-value sharing model could be interpreted as social values,
3. That the criteria are compensatory. The additive model meant that explicit relative trade-offs between criteria could be determined and the value of one criterion could be expressed in terms of an underlying attribute. Indirect compensatory values and the implied monetary value of changing from one scenario to another could also be calculated (see sections 6.3.2 and 8.1), and
4. That appropriate non-linear value functions would more accurately reflect values than linear value functions.

Questions or issues arising directly from these assumptions within the applications were:

- Whether the discussions and debate would mean that the team could represent societal values in terms of scores and weights (Section 9.4.2).
- Whether a compensatory model is appropriate in environmental decision-making (Section 9.6).
- Whether the indirect compensatory values could be considered as a SMARTx version of WTP (Section 9.6).
- That non-linear value functions for monetary criteria substantially affected implied values and interpretation (Section 9.6).

- Whether more accessible methods could be found for finding trade-offs weights (Section 9.7).

9.2.2 Valuation of alternatives or attributes using EE

The EE-CBA or EE analyses consisted of:

- Developing an overall method for evaluating, comparing and choosing alternatives,
- Determining appropriate attributes and attribute levels for CA questions,
- Designing TCM, CA, CBV and CVM questions and combining them in coherent 10 minute in-person surveys (KNP and Hluhluwe) or 30 minute in-person surveys (open-space study), with the understanding that preferences are revealed (TCM) or stated based on existing values (CA, CBV, CVM),
- Organisation of survey administration, and
- Data analysis.

Basic assumptions or requirements of the EE applications were that:

1. Monetary values obtained for use and non-use values could be aggregated in a cost-benefit equation,
2. Aggregate individual consumer values as measured by WTP would reflect society's preferences,
3. TCM, CVM, CBV and CA supply these values, and
4. The criteria are compensatory.

Questions arising directly from these assumptions within the applications were:

- Whether consumer preferences are appropriate in environmental decision-making (Section 9.4.1),
- Whether CVMs supply consumer values, and considered together with the other known biases of the method, whether CVM values are therefore acceptable values in these decision-contexts (Section 9.4.1),
- Whether the mean or median WTP is the appropriate measure to use (Section 9.4.1),
- Whether the TCM is appropriate in the contexts in which it was applied (Section 9.5.1), and
- Whether WTP values are appropriate measures of citizen or consumer preferences in environmental decision-making, in terms of the inclusion of equity and sustainability criteria (Section 9.6).

9.3 Metacriteria for comparison

A number of the issues which arose in the case studies are important in revealing the differences and similarities between the two paradigms and in clarifying which approach (or combination) would be appropriate in similar decision-contexts. These are elaborated here with emphasis on the two main cases and with reference also to the cases in Chapter 8.

Within the context of environmental decision-making in South Africa, choice of a decision-making framework and specific valuation or evaluation methods would need to refer to the socio-political context as well as issues of theoretical validity and practicality. The approaches were therefore compared in terms of the following metacriteria:

1. the resonance of the underlying philosophy of the method with the decision context and prevailing political milieu, and the consequent validity of the methods in this context (Section 9.4),
2. general methodological reliability and validity outside of the political context (Section 9.5),
3. the ability of the method to adequately include efficiency, equity and sustainability criteria (Section 9.6), and
4. the practicality of the method in terms of provision of decision-aid (Section 9.7).

The term validity usually refers to how “well-grounded” in principles, evidence and / or theory the results of a study are (Adelman 1991). There are several types of validity defined in the literature: the two that seem relevant here are theoretical validity (Hobbs 1979 in Goicoechea *et al.* 1982) and external validity (Adelman 1991). External validity refers to the degree to which the results of a sample or study can be generalised to the population, other populations or other settings (Garson 2002). There are two aspects to theoretical validity of MCDA and EE methods. The first is how well the assumptions of the method (e.g. those required for an additive model) correspond to the actual situation (this may correspond to what is elsewhere termed construct validity (Garson 2002)). The second is the ability of the method ‘correctly’ to choose A over B when the decision-maker prefers A to B (these are treated as one in Goicoechea *et al.* 1982). Reliability relates to whether a study can be repeated or replicated producing the same results.

The metacriteria are not disjunct: methodological problems may arise because of a dissonance with the political and decision context milieu, or a methodological problem may have interpretational, philosophical and practical implications. There are therefore numerous cross-references throughout the following discussions. First we outline the political and decision-context of the case studies (Section 9.3.1), and examine which approach resonates better within this setting (Section 9.4). We then discuss implications for the validity of EE (Section 9.4.1) arising from the requirements of the prevailing political and decision-context (consultation with citizens), as compared to that required by neo-classical economic theory (consumer preferences). Methodological issues arise due to the mixed nature of the values obtained in EE. This in turn has ramifications for the aggregation of different values, and for the interpretation of results. The validity of SMARTx in the context of the Consultative Model is examined in Section 9.4.2. Reliability and validity issues other than those arising from the prevailing political and decision context are discussed in Section 9.5. The ability of either method to adequately include equity and sustainability criteria is examined in Section 9.6. Finally, practical aspects relating to time, cost, transparency and accessibility are discussed in Section 9.7. Summaries and proposals are given at the end of each of these sections, where appropriate, and finally, Section 9.8 provides a summary of the methods in terms of the four metacriteria (Table 9.2). Examples from Chapters 6 to 8 are highlighted in boxes (where possible). However, the boxes are not ‘stand-alone’ as the text is continuous through them.

9.3.1 Political and decision-making context

The two main studies were carried out in the context of the National Water Act (NWA, Act 36, RSA 1998): finding a river health scenario which balances social, economic and ecological needs. The NWA promotes the integrated management of rivers, while falling somewhat short of full integrated catchment management (mainly because land-

use and practices directly impacting on river management are under the jurisdiction of other government departments). The NWA requires that the responsible water authority designate a management class for each part of the resource (e.g. river reach, quaternary catchment, estuary). Management classes broadly indicate the desired quality of the resource and therefore imply different levels of protection or allow different levels of use. In order to achieve the designated management class, the authority determines the Reserve. The Reserve is the quantity and quality of water appropriate for the level of ecosystem functioning indicated by the class, together with basic human needs. The authority also develops catchment management strategies, allocation plans and principles for the granting of individual licences. The NWA explicitly requires the consideration and balancing of ecological, economic and social issues and the participation of stakeholders to varying degrees in these different stages. While the other case studies were not conducted directly under the NWA, indirectly this act was also applicable to the Maclear study (together with the National Forests Act No. 84 of 1998) and the City of Cape Town water study (together with the Water Services Act No. 108 of 1997). Other national legislation was relevant (e.g. the National Environmental Management Act No. 107 1988) to CMOSS, also requiring participation and the consideration of social, economic and environmental issues. No particular policy framework was relevant to the Hluhluwe case, although legislation regarding exotic control, means that responsibility has devolved to the land-owner.

The relevant national legislation therefore (a) specifically requires that stakeholders be consulted or that participatory processes be used, and (b) explicitly requires the consideration of criteria such as equity, efficiency and sustainability. The two main cases are compared in the context of their contribution to choosing an appropriate level of protection and development (i.e. river health) for a river, given this policy framework (note that the term river quality includes the quantity component). The four other case studies supply contrasting or supporting contexts and results.

The second requirement (inclusion of equity, sustainability and efficiency criteria) is discussed in Section 9.6. With regards to the first requirement, the desire for consultative and participatory processes has been a strong theme in South Africa since the transition to democracy in 1994 as seen in the legislation mentioned (and as exemplified by the transition itself). We will refer to this model of political decision-making as the Consultative Model. On the other hand, the government has embraced the market economy and the need to base decisions on economic efficiency, sometimes to the apparent exclusion of other criteria (e.g. the strong push for privatisation of communication services and public transport, despite the potential job losses implied). We will refer to this latter model of political decision-making as the Market Model. The resonance of EE and SMARTx with these models of political decision-making are discussed below, together with related issues of internal theoretical validity.

9.4 Resonance with the prevailing political milieu and policy context

The Consultative Model gave rise to the relevant legislation. The Consultative Model and the wording within the legislation strongly suggest that an MCDA (although we limit ourselves to SMARTx) approach would be appropriate as a consultative process is followed. The Consultative Model also suggests that ‘citizen values’ are the

values which should be referenced, rather than only the values of the citizen as consumer. Citizen values are developed through debate, education and interaction, again suggesting that SMARTx might be appropriate as it is a discursive process. Accepting the Consultative Model, there is a legislative directive also to consider economic efficiency, and therefore appropriate economic measures are required. On the other hand, acceptance of the Market Model would obviously suggest the use of EE. There is also an inverse argument prevalent: that economic theory requires a certain type of value (of individual consumers), and that therefore these are the values that should be found (regardless of the prevailing Consultative Model or the need to consider equity and sustainability criteria).

Therefore it would appear that SMARTx might be a more appropriate overall framework than EE, but that EE is needed to provide economic efficiency measures. However, the evidence for the former statement is inconclusive at this stage (besides the discussion below, further support is proposed in Section 10.3). Furthermore, as discussed below, there are shortcomings in the use of SMARTx as an implementation of the Consultative Model, while, ironically, problems arising with some of the EE methods mean that conclusions drawn about economic efficiency may be invalid. The prevailing Consultative Model, and the implied need for citizen values, has various ramifications, particularly with regards to EE. But even if one would rather use the Market Model (and therefore EE), there would remain problems of internal validity in EE (Section 9.4.1). These relate to which values are required by economic theory, which values are obtained by the EE valuation, and therefore how theoretically valid they are and how to interpret them. These are discussed below, roughly in that order, the bulk of the debate arising directly out of the need for both ‘citizen values’ and economic efficiency criteria. Finally, in this section (9.4.2), the implications of the need for citizen values with respect to SMARTx are discussed.

9.4.1 Theoretical validity of EE

Many have argued that there are situations in which it is not ethically defensible to let the market decide (e.g. Cowen 1993), but the neo-classical paradigm and market mechanism of EE mean that it is the values of the consumer which are of interest. EE studies therefore take a sample of individual consumer or WTP values to reflect consumer preferences of the population (i.e. the Market Model framework). There are five debates around this point, which are discussed in Brown (unpubl.), and summarised here. The first relates to whether consumer preference is the appropriate basis for societal (in particular environmental) decision-making. The second debate is about whether the WTP values given in CVM or CA surveys actually reflect consumer preferences, or whether in fact they are some number which indicates their preferences as *citizens* (e.g. Brown unpubl, Sen 1977, Kahneman and Knetsch 1992, Carson 2000). In other words, do the WTPs given include issues relating to society’s welfare, rather than just the individual’s own self-interest and, depending on the answer, can they be used (and is there double counting). The third relates to whether consumer and citizen expressions of preference through WTP can be aggregated (for example in a CBA), and whether values found by different EE valuation methods can be aggregated and how to interpret results if they are. The fourth relates to potential biases that arise through the reliance on monetary measures (discussed mainly in Section 9.6). The fifth debate intermingles with all of these, and relates to the meaning and measuring of utility (and its relationship to preference and welfare) (Section 9.5.2).

EE: What does economic theory require?

Relating to the first point, economics theory since Samuelson has suggested that relevant choices are revealed in market choices. But within EE, stated preference techniques (primarily CVM, but the trend is towards CA) must be used for values such as existence value. These techniques, nevertheless, still rely on the construction of hypothetical markets, and so it is still the market mechanism that determines the measurement of value. Economists disagree on the validity of both the revealed preference and stated preference approaches (e.g. see the quotes at the beginning of this chapter).

The survey based EE methods used in the cases reported here were the TCM (a revealed preference approach), CBV, CVM and CA (stated preference techniques). TCM gave current preference, revealed in economic behaviour, CBV and CVM gave hypothesised (stated) changes in economic behaviour, and CA gave stated preference (linked to the travel-cost and CBV WTP in the KNP study).

Diamond and Hausman (1994) argue that stated preference methods have to elicit purely selfish values if they are to be consistent with economic theory (in Brown unpubl.) – by implication only consumer choices are relevant. Arrow and others disagree that only selfish values are relevant, their view being that it is perfectly acceptable that “...the individual orders all social states by whatever standards he deems relevant” (Arrow 1963, p. 17). On the other hand, Brown (unpubl.) feels that CVMs *should* be producing citizen values, and that only stated valuation techniques can provide collectively based, citizen values. In fact, values which include some social conscience may produce ‘better’ decisions. Sen (1974) showed with a prisoner’s dilemma game that decisions which include some ‘moral’ considerations or socially accepted norms (e.g. ‘fair-play’) would make everyone better off than if everyone acted as individuals in a market. The issue is then around finding decision mechanisms that reflect these norms and institutional mechanisms for implementation, rather than about the “superiority of one preference map over another” (Brown unpubl. p. 14). Blamey *et al.* (1995) argue that CVMs should not be valuing the project but finding the environmental impacts of the project on respondents personally isolated from other aspects of the project. Thus, EE and the Market Model taken together strongly imply that individual consumer values are needed, yet economists do not agree on this point. However, the practical applications of the theory (i.e. the valuation methods) are formulated using individual WTP and the use of CBA.

EE: What values do stated preference methods produce?

This leads to the next two issues: whether the stated preference techniques actually produce consumer preference information or citizen preferences, and whether, if the ordering produced is other than a consumer ordering, it can be combined with values produced from methods which produce consumer orderings.

This discussion is relevant to the results of the CBV question in the KNP study (as a type of CVM), and the open space study CVM. Besides the usual survey biases, the results are subject to another potential ‘bias’ because the responses potentially reflect citizen values rather than consumer values. As we stated, we felt that the responses to the KNP CBV question were ‘exaggerated’ (see Section 7.3 for the reasons)

and that it was unlikely that tourists would stay 24% longer if the river reflected the ‘ideal’ scenario we portrayed, or 30% less if it stopped flowing¹¹. This may have been due in part to the question format, but in addition, in the KNP study, respondents knew that the ‘ideal’ scenario presented was ‘better’ than the *status quo* and the ‘worst’ case scenario was much worse. Also, at this stage in the questionnaire, although not told about the purpose of the questionnaire, respondents knew that the survey was about ‘the environment’. It is very probable that responses included an element of ‘indignation’ at the thought of the river not flowing or ‘pleasure’ at the thought of upstream effects being reduced to improve environmental conditions thereby producing the ‘ideal’ scenario. In other words, the length of their stay at KNP would be unlikely to change much, but they unconsciously wanted to indicate their feelings about river management and quality and the environment as citizens. In the open space study, preferences as indicated by rank orders, were sometimes inversely related to WTP, while on the other hand, 35% of respondents gave WTP values that were more than 10% of their income. This supports the view that WTPs indicate non-consumer preferences.

It has been fairly conclusively shown that CVM WTP values tend to include ‘responsible citizen’, “warm glow” (Arrow *et al.* 1993) and “moral satisfaction” (Kahneman and Knetsch 1992 p. 57) values, and that they do not necessarily reflect consumer choice at all. The intuitively appealing view is that several ‘objective functions’ or ‘preference orderings’ exist –one may apply to choices in a consumer context and another to choices in a citizen context, and these two are quite likely to be different (Brown unpubl, Sen 1977). When faced with hypothetical environmental decisions, respondents might switch to the citizen preference ordering which might be “ethically superior” to the consumer preference ordering; when faced with an out of pocket payment, one might switch to the consumer ordering. An example given in (Brown unpubl. p. 8) is of a survey asking whether people would be willing to pay more than their current electricity bill in order to have ‘green’ power. As a citizen, one feels that green power is a good thing, especially if everyone agreed with it. Many people would respond to a hypothetical survey to reflect this feeling that green power is a good idea, by offering some positive WTP. However as a consumer faced with actually choosing which electricity to buy, many people would choose whichever were the cheaper power option. Even if they felt strongly that green power were ‘good’, their budget constraints might not permit them to buy it. (In European countries there is strong support for green power options and demand exceeds supply in some cases. This situation is unlikely to arise in developing world countries unless green power were cheaper.) On reflection, it seems clear that many CVM WTP surveys would elicit values which would never be revealed in a market choice, but which directly reflect the respondents’ preferences. This was most likely the case with both the KNP and open space surveys. The utility function of the individual may therefore also include the utility of others (Brown unpubl, Becker 1974, Sen 1977) –resurrecting the Edgeworth problem of nearly a century before of including other’s utility in our utility (Section 3.2).

Thus, one may find that the application of the methodologies has produced citizen values (not fully consistent with the Market Model), but one may feel that it is, in fact, more appropriate to include citizen values for environmental

¹¹ This scenario was not unrealistic as the river had stopped flowing for the first time in recorded history in a recent drought.

decisions (which would also be more consistent with the Consultative Model). CVMs that elicit citizen responses are regarded as closer to a referendum (as they may include ‘moral’ considerations) than they are to market consumer choices (Mitchell and Carson 1989, Arrow *et al.* 1993). Some therefore conclude that the results should be treated as a surrogate referendum. There are several questions that arise:

- Whether the mean or median WTP is appropriate,
- How to interpret combined citizen and consumer values, and
- Whether conclusions can be drawn regarding economic behaviour from citizen values.

EE: Interpretation – mean or median

Generally, there does not seem to be agreement about whether the mean or median WTP is the appropriate measure. Some regard mean WTP as appropriate for CBA (implying the use of the Hicks-Kaldor compensation rule), while median WTP is the standard public choice criterion (Carson 2000, Jakobsson and Dragun 1996), where the median corresponds to interpreting a CVM as a referendum and to majority voting. On this basis, the median WTP from CVMs should be used. However there appears to be no real reason why there should be different criteria applied in projects with environmental impacts (where a CBA is regarded as a means to finding social choice), and some argue that the median should be used in either case (e.g. Hanemann 1989) or that either correspond to the equal weighting of individual utilities implied by conventional CBA (e.g. Hanley 1987). Thus, whether the mean or median is used seems to be an ethical or political judgement that may be left up to the analyst, as the decision-maker or decision-ratifier may not be in a position to make a choice between these two or even be offered the choice. However, any conclusions drawn may depend fundamentally on which measure is chosen.

In both the KNP and open-space studies we found that although mean WTP revealed positive and significant WTP (in terms of length of stay for KNP and rates increases for open space), median WTP was zero (Figure 9.3). In an economics framing, using the median would lead to an interpretation that the majority of respondents were indifferent to changes in river quality from the *status quo* to the ideal or worst (KNP), or loss of 50% of open space (open-space study). Specifically, in the KNP study there would be no change in economic behaviour as a result of river quality changes. Note that vehicle bias probably affected people’s WTP (besides other potential survey biases, Section 5.2.4) as there was a slightly unrealistic vehicle in the KNP study (length of stay) and a politically very unpopular vehicle in the open space study (rates). However, in both cases, it was evident from other questions that respondents *were* concerned about these potential changes and had preferences for better river quality or more open space. This also pointed to the different types of value which could have been elicited, as discussed above, namely both individual consumer values and individual citizen values.

On the other hand, in the Hluhluwe case, mean and median WTP (in increased entry fee) were both positive (the mean increase was R20 and R37 and median increase was R10 and R20 for the current and ‘best’ scenarios respectively). Two factors might have contributed to this positive median WTP as compared to the median WTPs of zero for the other two studies. Firstly, the payment vehicle was realistic

as it was associated with the entry fee, an uncontroversial and familiar vehicle, compared with rates and length of stay. Secondly, a positive WTP for a recognised good was already established as people already paid an entry fee for viewing game. In the other cases, river health and open space were not things that people were familiar with paying for i.e. a market was not established.

If, as was the case in the KNP, there are many other computations associated with whether one chooses the mean or median, the potential for skewed results is large and conclusions dramatically different (see the next section).

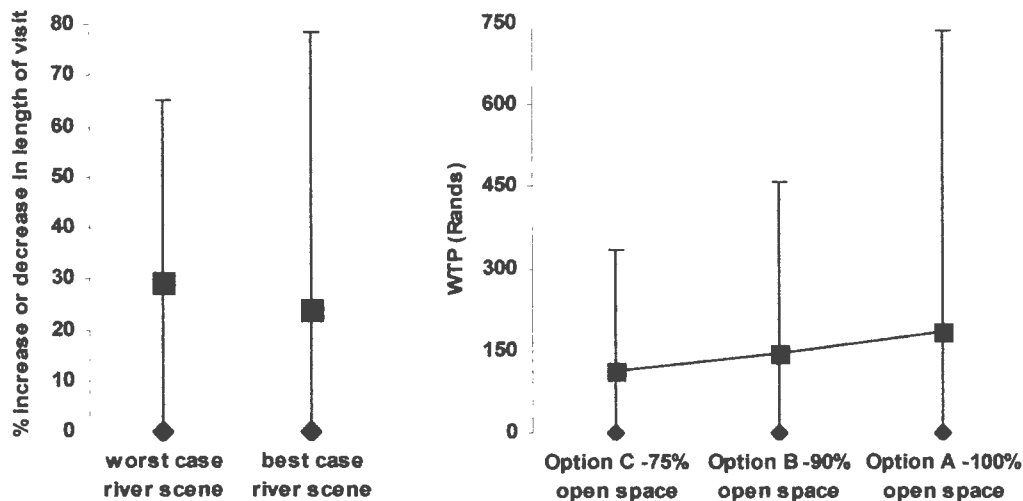


Figure 9.3. Mean (square), median (diamonds) and standard deviation (truncated at zero) for changes in length of stay in KNP with worst and best case scenarios and for WTP for open space in Cape Town.

EE: Interpretation – economic behaviour and/or preferences

That the CBV and CVM in the KNP and open space studies may have produced citizen values is not in itself problematic, but inferring economic behaviour changes (e.g. changes in consumption and actual revenues) is likely to be. This issue is perhaps exaggerated in the KNP study, because attempts to deal with the complexities (of multiple trips, of no convenient payment vehicle, of embedding of river value in game park value, etc.), led to a complex study design. However, it is likely to be relevant in any situation where value derived from CVM or CA are taken to indicate market or consumer values. Also, combining these values with market-based values is likely to create problems in interpretation (see next section).

In the KNP case, the indicated changes in length of stay (whether median or mean) from the CBV, which I have argued were citizen values, were taken to directly reflect changes in *consumer* value as revealed by the TCM (examples in the literature do the same thing e.g. Freeman 1993). For example, if someone said she would stay 20% longer with the ideal scenario, we assumed that her on site costs would increase by 20%. This is likely to have exaggerated consumer effects. Furthermore, the CA question in this case study was aimed at eliciting individual but non-market or non-consumer based preferences, and these

were also related via the CBV to changes in the consumer value reflected in TCM. Thus, we used citizen and individual non-consumer values to indicate changes in consumer value.

EE: Interpretation - values aggregated from different methods

For the broader study of which the KNP formed a part, the intention was to aggregate the various types of value for different river health scenarios (equation (9.1)). However, due to the high data requirements and expense, other values were not obtained for scenarios, but rather as one-off values (see Table 7.5 and Mander *et al.* 2001). These could not be translated to any particular change in river quality, and thus could not really be included in our decision-context (but were informative). Besides the expense of finding scenario-based EE values for the broader KNP study, the problems discussed above would apply to many of these values. In particular, citizen and consumer values (a problem if it is not clear that they are different) would be aggregated, encompassing widely different levels of 'accuracy' or 'reliability'. One would also have to consider whether to weight different types of value differently, based on who received the benefits or costs (e.g. tourists versus riparian residents). Interpretation and decision-making based on these values would be difficult to say the least.

In the open space study, the CVM values were added to values from a hedonic pricing study (assessing the effect of different types of open space on property prices), and values from a replacement cost study (estimated the cost of engineering alternatives to the ecological function of certain open space types). The standard deviations of the CVM value were large, and were not reported for the hedonic and replacement cost studies. For hypothetical decision-making purposes, the summed values were compared with the real estate values of the land. Unsurprisingly, the conclusion was that open space categorised as 'vacant lots' had least value and could be developed, particularly in Metro south-east. Other open space types had aggregate WTP values that were only slightly more than their real estate values. Given the wide standard deviations and issues of scale insensitivity any conclusions drawn from these relative values would be subject to a multitude of criticisms. The recommendation was that given the pilot nature of the study and the small sample sizes the values could not be used in the ways suggested hypothetically, but that more detailed studies would be needed, in particular related to decisions about particular sites.

Given the wide range in accuracy of values arising from different methods for use and non-use values, and given the potentially different types of value (citizen, consumer, individual), there is potential for error in any conclusions drawn from aggregated values. Whether one has intended to produce citizen values from a CVM or accidentally elicited them, one may end up with both consumer and citizen values from the application of several methods within EE. Ignoring all the arguments above and accepting both revealed and stated (or mean, median, referendum, market) WTP values within a CBA, the question remains as to whether a 'citizen Rand' is equivalent to a 'consumer Rand', and therefore whether they need to be weighted differently in some way when aggregated. Whether they are explicitly weighted or not, how does one interpret the 'final answer' and form any conclusions? The issue of

interpersonal comparisons of utility also arises in the EE studies because individual values were aggregated: individual consumer surpluses in the TCM, individual WTP from the CVM, and individual preference from the CA.

9.4.2 SMARTx: Are group-values also citizen values?

The SMARTx cases were run in workshop settings. Specialists represented the broad themes of ecology, society and economy, and sometimes implicitly represented associated constituencies. For example, although no Muslim participated directly in the City of Cape Town water study, issues relating to Muslim views on recycled water were included in the discussions around ‘public acceptance’ – and consideration of the potential problems was one of the main reasons for the inclusion of this criterion. Also, although no ‘subsistence user’ was included in the Sand case, their interests were one of the focuses of attention because of the specialists’ experience in that catchment. The representation of constituencies in this way, only really occurred for criteria within the social group of criteria as other criteria groups had more ‘discipline’ based criteria.

The SMARTx process applied the group-value sharing model. The model was applicable as the broad objectives for the projects were established already (by policy, by the initiation of a project, by the terms of reference of the study team). Though vague, these were sufficient to allow people to sit down together and begin to tackle the problem. Also, in the Sand and water supply cases, the teams had official mandates to undertake the task, and were all very familiar with the range of issues.

The group value-sharing model, by means of workshops, applied ‘rule *d*’ (Section 5.3.3) to select criteria and transform values to group values. The group-value sharing model applied meant that the partial values found were group values measuring the impacts of alternatives projects on criteria (not necessarily on different groups of society). Weights did not directly trade-off ‘winners’ against ‘losers’, but rather improvements in one criterion against reductions in another. Normative decisions about weights etc. were made by the participants as part of the ‘citizen analysis’, were a focus of attention and were available for perusal by the final decision-makers (e.g. Dept. Water Affairs and Forestry, Dept. of Agriculture, City of Cape Town).

These cases were all ‘Phase 1’ projects, to be followed either by broader participation or more in-depth investigations. In this context, the group-value sharing model seems appropriate and corresponds with the requirements of the NWA. It also corresponds with the Consultative Model to the extent that the values arose through consultation across various disciplines representing different interests. The group-value sharing model should reduce problems of interpersonal comparisons, as the group discussions concentrate on the *relative* effects on different criteria. Although we felt that concentrating on valuing impacts on criteria rather than on groups of people was a strength of the approach, in some cases this might be considered a weakness. For example, one might want to keep an explicit account of who benefits and loses and by how much.

In any case, although the group-value sharing model was explicitly applied in Phase I projects, there was perhaps an underlying assumption that the group process produced ‘citizen values’, or that the derived group values could be extended to imply social value. However, in the applications in this study, it was the specific aim of the exercises and the given responsibility of the participants to produce an overall assessment which would broadly represent the affected population. De Montis *et al.* (undated) say that no “social function” is built using MCDA, but base this on the use of the group-value aggregation (rather than sharing) model. However, although we feel de Montis’ statement is too strong, and they do not say on what basis they make it, the assumption that the group-value sharing model produced ‘citizen values’ may be questionable because the number of people who can be involved in the workshop process is limited. So the problem may arise that not all groups feel that they are represented or that only one representative can attend from each group. This in turn means that individual representatives may have different abilities to adequately represent their constituencies. In effect, the process has increased depth of consultation, but less breadth of input in comparison to EE which surveys 100s or 1000s of people. In summary, there is no available theory that says that group values will correspond with the maximisation of social welfare. However, the only real alternative to the group-value sharing model is the group-value aggregation model, and this model was inappropriate given the political and decision contexts. We were specifically *not* interested in individual utilities, but trying to solve a social problem (improve social welfare) by referring to impacts on criteria, rather than by referring specifically to how individuals felt about them (maximise an aggregation of individual preferences). The former is perhaps a more paternalistic view than the latter, consumer sovereignty, view.

This shortcoming in terms of the breadth of representation in SMARTx has implications for its validity in terms of the political level Consultative Model. It also has implications for equity in terms of equitable representation in the decision-making process and for the acceptable inclusion of equity and sustainability criteria in terms of both selecting criteria and in terms of assessing trade-offs (Section 9.6). Generally, the shared understanding of the implied trade-offs developed during the workshops meant that weights were uncontroversial, but there may be situations where this does not always apply (and conflict resolution methods (Section 4.1.3) may apply). There are also practical implications due to the participant burden in terms of representation and commitment (Section 9.7).

Note that MCDA *theory* in contrast to economic theory, does not have anything to say about what type of values (consumer or citizen) should be included in an analysis, nor whether the values should be revealed or stated (although von Neumann and Morgenstern were interested in revealed preferences). Although some of the criteria included in the SMARTx cases were economic criteria, they did not reflect economic behaviour of consumers so much as economic impacts on both consumers (e.g. in their income earned) and producers (e.g. in their profits). As preferences for these attributes are monotonically increasing, their value is ‘revealed’ by the value function. So the *measurement* of value does not depend on the market, although effects on the market may be measured.

9.4.3 Summary discussion and proposal for Section 9.4

Thus far, we have seen that the CVM method may produce either citizen and consumer values and this has various ramifications in terms of the validity of this method within either the Consultative Model or the Market Model. There are also implications in terms of the validity of any conclusions drawn which have combined CVM values with other values, or inferred economic behaviour changes from the responses. Thus, the confounding of citizen and consumer values creates a major stumbling block to the use of CVM. We questioned the assumption that the group-value sharing model would be adequate for social choice, although it appeared to be appropriate for ‘phase 1’ projects. We propose that, where appropriate, SMARTx studies could relatively easily be augmented by a survey or surveys of the broader population. Questions in such a survey might apply to the choice of criteria, alternatives, scoring and weighting (other approaches to dealing with this shortcoming are discussed in Chapter 10).

To a large extent the experience with CBV and CVM in the case studies has supported the view expressed by Niewijk (undated) that “Under the glare of that scrutiny, [CVM] looks dreadful: experimental evidence suggests that, at best, [CVM] grossly overestimates the actual values that people hold for a resource, and at worst, it does not measure those values at all. In the light of these results, [CVM] should not be used in either cost-benefit public policy analyses or natural resource damage assessments”. In the case studies described here, WTP values were given that were in some cases clearly beyond the means of the respondent (e.g. the open-space study), which contradicted their preferences as given by simple ranking (e.g. the open-space study), and which were probably exaggerated (e.g. change of length of stay in the KNP study).

Although the majority of economists agree that CVM is a severely flawed method, they usually claim they have no alternative for measuring existence value. However, while there may not be an alternative for estimating compensation in resource damage assessments, SMARTx provides a realistic alternative in the case of project or policy evaluation. In fact, SMARTx would appear to be a superior alternative, rather than a poor relation. From the vast literature available, and the examples in this thesis, we have to conclude either that (a) the results of CVM can only be used to indicate relative citizen preferences, or (b) that SMARTx-type methods should be used instead. If (a) is concluded, then the use of SMARTx to indicate relative preferences would probably be simpler, and so we conclude (b) anyway. The concept of existence value can be captured either by using several criteria within SMARTx which contribute to existence value (e.g. biodiversity, aesthetics, ecosystem health), or by including an explicit existence value criterion.

SMARTx relies very much on *procedure* or *process* (procedural rationality) to obtain trade-offs which reflect the group’s or society’s values. In the context of the Consultative Model, the relative transparency of the SMARTx process is important. Following the steps of SMARTx and the prescriptions of value measurement theory, is a process of rational decision-making, and the group-value sharing model extends this to a group. However, there are many influences on these steps, particularly problem structuring, which might influence the results. So, the values and weights resulting might be ‘wrong’ for a number of reasons: the ‘wrong’ people were there, some people were more eloquent or more comfortable with the process and therefore were able to sway the argument a particular way

(without necessarily having any sinister intentions). Of course, even if these problems were absent, decisions might still be taken which are ‘wrong’: the emphasis of SMARTx (and MCDA in general) on procedural rationality means that at the very least, the reasons for such a decision can be relatively easily traced.

9.5 General validity and reliability

Additional problems of reliability and validity (other than those arising out of the policy and Consultative Model setting) arose with respect to the consumer surplus calculated using the TCM, the scales of measurement implied and used, preferential independence, survey biases, weights and uncertainty.

9.5.1 Consumer surplus and the travel cost method

The idea of aggregating individual consumer surpluses has been debated since Edgeworth, around the turn of the century. On the one hand Staley (1989) says that “...so thoroughly has consumers’ surplus been expunged from modern theoretical welfare economics, that a recent authoritative text does not even mention it”. On the other hand, in Samuelson and Nordhaus (1985), which surely must be considered an authoritative text, they say it is “...a concept relevant for many social decisions – such as deciding whether the community should incur the heavy expenses of a road bridge” (p. 421). However, in their example, they say that “To avoid difficult issues of interpersonal utility comparisons, assume that there are .. users .. identical in every respect” (p. 418), therefore brushing aside the main problem. Pigou, who first showed that individual surpluses could not be added, felt that the interrelationships of peoples’ utilities were stable and therefore when price changes were small, the interrelationship had little effect (Stigler 1950). It has generally been the view that, for various reasons, aggregating consumer’s surpluses is acceptable in practical contexts such as CBAs.

There were a number of reasons why the consumer surplus estimate from the TCMs applied in the KNP and Hluhluwe cases appeared both unreliable and invalid. First, in both cases, a large range in calculated consumer surpluses (an order of magnitude, from R1 bn to R24 bn in the case of KNP) was produced with different arrangements of the data (i.e. which countries were grouped into zones and which were excluded because of small sample sizes). The choice of functional form can strongly affect the results of a TCM, but the semi-log form that was used is regarded as appropriate and robust. The transformation of the dependent variable should help to deal with unequal variances in zones (Strong 1983). However, the within and between zone variances and differences in our cases seemed too large for this model, and including GNP per capita as a variable did not improve the variance explained. It is likely therefore that the consumer surplus estimate would be unreliable and possibly biased (i.e. total consumer surplus may be over- or underestimated). Second, the different numbers of tourists from different countries is still probably more historical than economic (e.g. South Africa has a strong historical link with England, but less so with the USA). A third reason is theoretical and reflects back on the history of economics given in Chapter 3. Aggregation is only really valid if the marginal utility of income is the same for rich and poor (Baumol 1977b in Ross 1999 p. 263, Dixon *et al.* 1986 p. 25, Staley 1989 p. 187). However, in the KNP and

Hluhluwe cases it would definitely have been the case that the relevant populations (i.e. people from all zones ranging from USA to Limpopo Province, South Africa) would have widely differing incomes (as well as tastes and other attributes) and therefore marginal utilities of money both within and between zones.

For these three reasons, therefore, we recommended that consumer surplus should not be used in contexts where source populations have widely different incomes, tastes and backgrounds within and between zones. However, in an economics based decision setting this would be problematic and the amenity in question would be consequently undervalued. We suggest a possible alternative in Section 9.5.8.

9.5.2 Measuring utility - or the ordinal/cardinal debate lives on

The nature of utility (subjective or objective), its measurement (ordinal or cardinal) and its relationship in EE to choice, preference, welfare and WTP are discussed here. With respect to EE, these discussions again reflect on the internal validity of the approach, given the implications of the use of WTP as a measure of utility.

One of the major debates in the first part of the twentieth century concerned the nature of utility. Was it, as Bentham, Jevons and Marshall felt, a measure of psychological satisfaction, by nature personal, or should/could it be 'objective' and concerned with measuring the "marginal rates of commodity substitution"? (Ross 1999 p. 169). The desire at this time was to eschew value-judgements and subjectivity, and pursue a more positive economics, and economists (generally) adopted the 'objective' interpretation of utility, at least in theory. In practice, the original or Jevonsian concept of utility (Chapter 3) still permeates through quite strongly, and practitioners still hope to measure how 'happy' people are with a project.

One of the important developments in economics at the beginning of the 20th century was the change from an assumption of the availability of a cardinal value function to the reliance on only ordinal information (Chapter 3) (apart from attempts to measure the cardinal utility of income by, for example, van Praag *et al.* (1982), which have been largely ignored by mainstream economics). This was because it was felt that cardinal value functions were too restrictive to use and too difficult to obtain (e.g. Henderson and Quandt 1980). The restrictiveness relates particularly to conclusions about the shape of resulting demand functions, and is not addressed here. The possibility to only use ordinal utility arose because of work at the turn of the 20th century which showed that ordinal rankings obtained for individual preferences are a sufficient basis for individual choice, i.e. for constructing indifference curves (Section 3.1) (indifference curves also addressing the problem of lack of independence between goods (Chapter 3)). Therefore, economics text books always mention that cardinal utility is no longer used, because ordinal utility is sufficient. But, WTP in the cost-benefit equation (where $WTP = \text{price paid} + \text{consumer surplus}$), is used as an *indicator* of utility and is certainly treated as cardinal in that WTP values for different commodities and individuals are considered comparable. But WTP is a biased measure of utility (Section 9.6.1), and may not necessarily indicate *consumer* preferences at all (Section 9.4.1).

If, at the individual level, only ordinal utilities are available we cannot aggregate them. Hanemann and Kanninen (1996) claim that referendum CVM and ranking CVM, which model behaviour in terms of ordinal preferences, avoid inter-personal comparisons of utility. Interpersonal comparisons of utility might be avoided in a sense by using ranks or yes/no responses, but this is somewhat irrelevant as interpersonal comparisons of money (the object of interest here) are not avoided (as money is included as an attribute of the scenarios ranked). In any case, the fact that the regressions are performed on the whole sample by its very nature assumes that the sample comes from a coherent 'population'. In other words the utility and WTP functions are assumed to come from the same population and therefore can be lumped – this is interpersonal comparability.

Lastly, although the aggregate value or B/C ratio is intended as ordinal, the monetary value makes an ordinal interpretation almost impossible to maintain as “symbol and reality become easily confused” (Self 1970, p.8 in Pearce 1983).

In summary, originally, utility theory was both descriptive (people make choices by maximising utility) and prescriptive (utility can be used to make social decisions; for example, Bentham felt it could be the basis of a new legal system). By the turn of the 20th century utility was used more descriptively and to predict demand, and attempts since then have concentrated on making it more strictly positive. Therefore utility was used as a “theoretical construct” to describe behaviour rather than to guide or prescribe decision-making; the “logical superstructure” of economics is not the utility function but the indifference curve (Bell *et al.* 1977). The descriptive basis was incorporated into the development of CBA within welfare (i.e. normative) economics, yet it was used as a prescriptive tool. How people make choices is *described* by maximising utility, but assumed to be *indicated* by WTP, and the option with the maximum WTP is *prescribed* as the best choice, given the *normative* basis of consumer sovereignty and ‘market rules’ (see e.g. Watson 1981). Thus, EE has mixed normative, prescriptive and descriptive elements, while trying to be ‘positivist’ and ‘objective’. Von Winterfeldt and Edwards (1986 p. 562) feel that “Economists have turned their backs on the intellectual consequences of cardinally measurable utility” and that “This is the intellectual partition that divides cost-benefit analysis from decision analysis.

In MCDA in general, it appears that the original concept of utility as a subjective measure of value is intended and maintained. Although when measuring the level of an attribute, for example different levels of air pollution, a measure may be objective, the translation to value may include a subjective component (e.g. marginal utility of income to different groups). Specifically, the overall aggregate evaluation of a project and trade-offs between criteria with SMARTx are subjective. These different intentions are fundamental to understanding the difference between the two paradigms, particularly as preferences or choices may change depending on if viewed from the point of view of a citizen or consumer.

Thus, SMARTx directly aggregates utility (where the term utility is used to correspond with the EE use), rather than a WTP indicator of utility. Additionally, this is not an aggregation of individual utility but rather an aggregate of ‘issue-based utility’, where the value of an alternative is measured by its ‘performance’ on a number of different

criteria. The group-value sharing model and the choice of criteria based on issues rather than stakeholders, helps to limit the problem of interpersonal comparability of utility - the utility measured by a criterion is not equivalent to the utility felt or measured by an individual. Also, MCDA is explicit about the nature of the scales used, SMARTx making use of interval scales while some MCDA methods build up a cardinal value function (or a range of consistent cardinal value functions) from purely ordinal information (e.g. conjoint scaling, Section 5.2.3). The linear or non-linear relationship between money or any other quantitative attribute is normally a focus of attention in SMARTx and non-linear functions were defined for several of the criteria in the three SMARTx case studies. Besides being a better reflection of values, non-linear relationships might reduce the potential pro-rich biases of EE (Section 9.6.1), and have better implications for dealing with sustainability (Section 9.6.2). The use of direct measures of value or performance might also be more transparent than the use of WTP (Section 9.7).

The final point in the cardinal/ordinal debate is that, in contrast to the claim that cardinal utility is too difficult to obtain, individuals seem quite happy to, and capable of, providing some kind of strength of preference information (which may be as comparable (or not) as WTP). This was seen in both the SMARTx 'scoring' process and in the EE applications (where CA and direct ratings (e.g. "what percentage importance..." questions) were used). This view stems from informal feedback from the enumerators in the EE studies, personal experience during the SMARTx workshops and the results of a questionnaire sent to SMARTx participants (Section 9.7.5). Whether these are strictly on an interval scale is debatable and probably unlikely, but perhaps 'quasi-interval' scales still provide more and relevant information than do rank orders. The relationship between these scales and true interval scales, and the implications of the use of 'quasi-interval' scales within SMARTx is a matter for research.

Although, in surveys, different parts of the given interval scale may be used by different people, or mean different things to different people, in particular cases this may not be an issue (as in the KNP study where scores were rather consistent, see Section 7.2.2). In the KNP case, we investigated the analysis of the CA data using only the ordinal information supplied, but found that despite the high consistency of responses, the ordinal model could not adequately represent the data. This study suggested that, for this type of survey, with good design, the use of 'scores' is efficient and appropriate, and that an LP formulation is more robust than an OLS formulation (see Section 7.2). The design would need to include a benchmark (e.g. the *status quo* as in these cases) and perhaps verbal cues for points along the scale.

For survey-based valuation, CA with interval-scale scoring has far larger informational and predictive power, and lower respondent fatigue than would a CA using ordinal information only. A conjoint scaling design, which would have greater power than the CA to construct interval scale scores, would not be appropriate in a survey. In some cases, conjoint scaling would be useful in a MCDA decision-group setting.

9.5.3 Mutual preferential independence and difference independence

The additive form of the cost-benefit equation and of SMARTx imply that the criteria or attributes (or individuals) being added are at least mutually preferentially independent. In the SMARTx applications we only informally tested whether people could give scores for one criterion without referring to the level of another criterion (essentially, waiting for warning signals, such as people saying ‘..well that depends on..’). However, we did not attempt to measure the extent to which independence held, and for the most part assumed mutual preference independence and that therefore the additive model was valid. This would mean that the aggregate values of alternatives $V(a)$ were only interpretable on an ordinal scale, but not making use of the cardinal properties of the numbers was a difficult feat to maintain (as with aggregate WTP values). Note that application of the group value-aggregation model would have imposed even stricter independence on the criteria (difference independence) (Section 4.1.3). Independence assumptions were not verified in the EE applications (other than trying to choose attributes in the CA models that we felt were preferentially independent).

From my experience, the types of questions that appear to be necessary to more carefully verify independence assumptions are likely to be considered tedious and off-putting by participants if included within the SMARTx process. However, Stewart has suggested that, where there are serious deviations from additive independence, results might be strongly affected (Stewart 1996) (although Stewart (1995) found that the additive *utility* model was a reasonable substitute for the multiplicative model). There is therefore a need to develop more ‘user-friendly’ questioning procedures for use within workshops, and also to examine the implications of interval rather than ordinal scale interpretation of aggregate value V where only mutual preferential independence held.

9.5.4 Survey, questioning and meeting biases

The various potential biases of CVM and CBV surveys were mentioned in Section 5.2.4. Although referendum CVM is theoretically less biased than the open-ended approach we used, the large sample sizes required for referendum CVM meant that it was not practical for our case studies. While many of the survey biases would have applied in our case studies, this was most evident in the open space study, which demonstrated at least vehicle bias, interviewer bias and embedding bias. Scoring in SMARTx may be subject to bias due to framing (e.g. framing performances in terms of losses or gains) and weights may be subject to biases for reasons discussed in the next section. In both these cases, the workshop process should limit the extent to which these biases happen. SMARTx may also be biased due to the deliberate manipulation of participants. For example, Wenstøp and Seip (2001) found that the environmental panel in a particular study wished to change their weights once they realised that the implication of their initial choice was that Norway would reduce its preparedness for oil spills. In our cases it was not evident that people tried to manipulate the process. There was perhaps one slight exception in the Cape Town water study where the evaluation of one alternative (Table Mountain Group Aquifer) according to one criterion (impact on groundwater) became contentious. After ‘robust’ debate, it became apparent that the ‘manipulation’ was in fact a result of differences in perceptions regarding the extent of the initial implementation phase of this project, and the basis on which the evaluation was being made (whether according to potential impacts based on current

information, or impacts as would be indicated after a required pilot test project). This was largely dealt with by more closely defining the criterion of relevance. While the number and extent of potential biases of the EE survey methods may be larger than those of SMARTx which can be mitigated by the group process, in either case, the analysts have the responsibility to ensure that all potential biases are minimised.

9.5.5 Weights and explicit trade-offs

It is assumed in EE that the importance of different impacts in an overall cost-benefit equation is captured by the relative monetary values and therefore weights are irrelevant (apart from the occasional use of income distributional weights). The SMARTx applications relied on the group-value sharing model to find the swing weights for the social, ecological and economic criteria groups. This examination was done directly between these criteria groups rather than by comparing one criterion from within each criteria group, which might have been more accurate. Another short-coming in the SMARTx applications was that we did not allot enough time during workshops to more carefully examine the trade-offs implied by the weights (for example, the indirect compensatory values (Section 9.6.2)). Perhaps surprisingly, there was little controversy over the weights, and certainly none over the rank order of the weights. There were three possible reasons for this. One reason may have been the assurance that the range of weights would be included in sensitivity analyses. Secondly, by this stage of the process, the group as a whole had a reasonably good appreciation of the range of impacts within the various groups, and thirdly, the swing weight process may have helped to focus this appreciation.

It has been shown that weights found by different methods do not correlate highly (e.g. von Winterfeldt and Edwards 1986, Bell *et al.* 2001), and that weights are affected by the ‘splitting bias’ where the number of levels in a value tree affects the weight given to a criterion (e.g. Bell *et al.* 2001). However, von Winterfeldt and Edwards (1986 p. 365) found that the additive model is reasonably robust to different weights, as we also found in the case studies here. Where economic or financial values are available from the particular study or from similar studies, these may sometimes be used as upper or lower bounds on weights. This was the approach applied in a study not reported here (Lamberth and Joubert, in review), where the financial value of the fishery for different line-fish, provided (together with the number of people involved) a guideline for the weight given to non-conservation criteria.

9.5.6 Uncertainty

Both SMARTx and EE-CBA rely on sensitivity analyses to deal with uncertainty or imprecision. This may be inadequate for dealing with uncertainty about long term social and environmental consequences (which may be irreversible), future political, social and economic conditions, and the dynamic relationships between these, as well as their impacts on preferences.

Normally, CBA sensitivity analyses would be limited to sensitivity to discount rates, and perhaps income distributional weights. With the expectation of rapidly changing income distribution in developing countries there is an implication of changing preferences over time even in the short term as people’s tastes (especially for environmental goods if tastes are measured by WTP) may change with their incomes (even if compensation does

occur (Scitovsky 1941)). Valuation of environmental impacts obtained *ex ante* may differ from that which the same respondents would provide given the distribution of income after the project. Uncertainty with regard to the value of environmental amenities or the consequences of decisions (external uncertainty), is considered in EE through ‘quasi-option values’ which represent the value of preserving future use or existence given some expectation of an increase in knowledge about the environment in question (Pearce and Turner 1990). Quasi-option value may also be seen as the value of foregone natural assets in the event of a shift in public tastes, or the cost society is willing to incur to acquire the scientific information needed for more accurate assessment of a project’s net benefits. Quasi-option values therefore are related to the rate of change of income expected in the near future: this itself is a source of uncertainty in developing countries, as is the distribution of that income.

In the SMARTx cases, besides sensitivity analyses, uncertainty was included through certain criteria defined as ‘the risk of ...’. For example in the Sand case we had “the risk of spread of exotic flora and fauna”, and in the Cape Town water study we had several criteria relating to risk such as “the risk of lack of public acceptance and buy-in” and “risk of lack of political support”. To deal with imprecision or disagreement on scores and weights the full range was included in sensitivity analyses. Thus, SMARTx has a relatively simple and transparent means of addressing uncertainty or risk and disagreements. Due to the many other complexities, the EE case studies did not include sensitivity analyses and did not consider uncertainty.

Other approaches to uncertainty in MCDA (Section 4.1.4) included multi-attribute utility theory (MAUT), semantic, ordinal and fuzzy approaches. We refer to that discussion and to our experiences in the case studies, in suggesting that the approach of Salo and Hämäläinen (1995), where ranges of scores are given (rather than points) and approaches where alternatives are given fuzzy membership in different categories, are also potentially useful in including uncertainty and imprecision. Currently, the lack of user-friendly software for these methods may limit their usefulness (particularly given workshop settings where participants may have very different backgrounds).

Whether sophisticated or simple approaches are used, in the end the decision taken will be a ‘gamble’ whatever the process used (Walters 1986), but with some exploration of the alternatives, the participants should be more comfortable that it is a reasonable gamble. Stakeholders may suggest that rather than taking a gamble with uncertain consequences, we should ‘wait and see’ or ‘do more research’ (Walters 1986); this is equivalent to the quasi-option approach from EE. However, in many cases decisions need to be made as a matter of urgency for other reasons. In developing countries this may be the need to ‘deliver’, combined with a real need for income growth amongst the poor. It must also be recognised that, in some cases, more research will not necessarily provide the answers, as ecological realities remain complex and consequences may be unknowable in any near term future.

9.5.7 External validity and reliability

A few additional remarks are necessary with regards to the external validity and reliability of EE and SMARTx. It is difficult to assess the reliability (replicability) and the external validity (extent to which results can be generalised) of

the SMARTx cases described. Firstly, there is no basis of comparison as another ‘decision-aid’ method was not used in parallel. Secondly, the SMARTx process is designed to construct values and develop the thinking around a project, but one cannot say ‘this was an SMARTx insight and that was not’. Thirdly, a good process can lead to a ‘wrong answer’ and lastly, one cannot always tell whether an answer is correct or not. If the group does not know that it prefers A to B at the start of a SMARTx exercise and at the end of the process the conclusion is that $A \succsim B$, the only real test is whether the instincts of the participants make them feel sufficiently uncomfortable to challenge the results (even just to say “that seems strange”). However, if the Sand case were used in other catchments as a basis for similar studies concerned with Reserve determination and designation of river management classes, the process would become more formalised (as is beginning to happen). For example, the types of criteria which must be considered may be pre-determined (with freedom to give zero weights perhaps), and scoring systems with careful definitions of end-points and a few intermediate points may be defined. In that case one would expect that the methods should become robust to personnel and other changes and therefore be reliable.

One would expect that one could more easily assume reliability of EE methods. Once one has developed a survey instrument, this should indeed produce similar values if reapplied, however, as with SMARTx there are many value judgement which precede this stage and of course, slightly differently worded surveys have been shown to produce very different results.

9.5.8 Summary discussion and proposals for Section 9.5

The total value of a natural amenity in economic terms is consumer surplus plus revenue, so ‘ignoring’ consumer surplus from the TCM, as suggested above, will undervalue the amenity. However, if a different framework were used (e.g. SMARTx), revenue could itself be a criterion (i.e. R136 mill in the KNP), while other values relating to the amenity (e.g. aesthetics, game viewing) could be measured by other criteria. This might account for the ‘surplus’ which the consumer enjoys that is not reflected in price. Therefore for the KNP case we might say that, for example:

$$\begin{aligned}\text{tourism value} &= \text{tourism revenue} + \text{tourism enjoyment, and} \\ \text{tourism enjoyment} &= \text{aesthetics} + \text{game viewing.}\end{aligned}\tag{9.5}$$

For example, in the Sand case, we included the operating margins from commercial conservation activities (aggregated with operating margins from all the various land-uses in the particular scenario). Other criteria included species and habitat diversity, as well as aesthetics. These latter reflect aspects of our SMARTx version of ‘consumer surplus’, or alternately, an explicit criterion, say, ‘tourism experience’ could be added to more directly measure how tourists feel. While not providing the economics consumer surplus concept, the new criteria reflect values that a consumer gains, which might not be reflected in revenue. This formulation, equation (9.5), might be reflected in a SMARTx value tree (e.g. Figure 9.4).

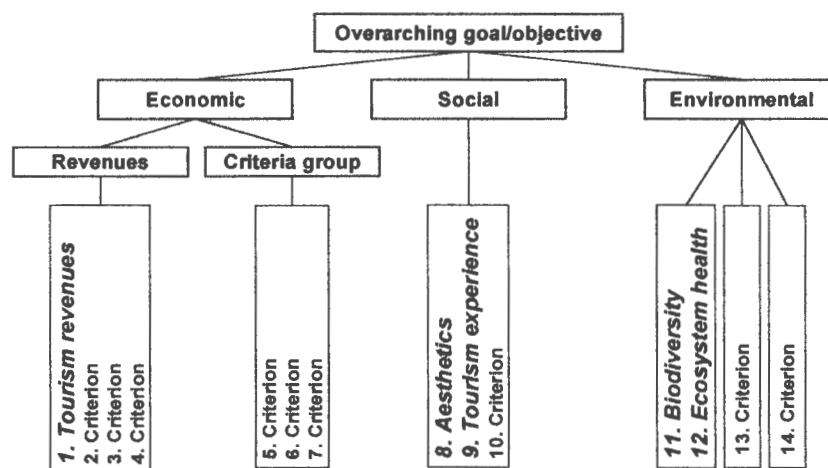


Figure 9.4. A SMARTx value tree including criteria that together might indicate tourism WTP (price+consumer surplus), where ‘price’ is included in revenue, and the criteria aesthetics, tourism experience, biodiversity and ecosystem health might contribute to ‘consumer surplus’. Several of these criteria were included in the Sand case (see Figure 6.2).

Viewed from the point of view of a comparison of EE and SMARTx it seems that ‘the cardinal-ordinal debate’ is far from over, even though there is a 100 year-old tradition behind the abandonment of cardinal utility in economics. Given the mixed signals potentially arising (evident from this debate and from Section 9.4.1), there is perhaps a need for a re-evaluation of the use of WTP as an indicator of utility, and of the assumption that only ordinal utility is available (particularly as the apparent implications of interval utility for demand functions may not be relevant in an EE valuation setting).

With regards to independence assumptions, the development of user-friendly questioning procedures for use within workshops would be invaluable. Lastly, a COMPAIRS approach (Salo and Hämäläinen 1995) for including ranges rather than point values would be a useful augmentation of SMARTx.

9.6 Efficiency, equity and sustainability criteria

Broadly speaking, social welfare decisions have three components: economic efficiency, social equity and environmental sustainability. This is reflected in the general literature and in national and international policies and programmes (e.g. the World Summit on Sustainable Development 2002, Mail and Guardian 2002). These three components are also a strong theme running through the National Water Act of South Africa (NWA, Act 36 1998) which formed the context for the Sand and KNP case studies. Based on these studies and others in the literature, these are often the three main groupings of criteria. Contrary to the view of Rauschmeyer (2001) these broad criteria clearly include ‘ethical’ concepts such as inter- and intra-generational equity.

CBA attempts to achieve efficiency by mimicking a perfectly competitive market. However, perfect *laissez-faire* cannot guarantee a general equilibrium that maximises aggregate social welfare, merely one that achieves efficiency given the current distribution. In addition, the second best theory (Lipsey and Lancaster 1956) implies that the allocation of resources given by a project under perfect competition need not result in the greatest net improvement in overall welfare if there is at least one sector in the real world that is not itself perfectly competitive.

Despite this, we will accept as a given that conventional economics or EE techniques are needed to supply economic efficiency criteria and that SMARTx cannot supply these. However, we will discuss equity and sustainability in more detail.

There are four aspects to effective inclusion of equity and sustainability considerations when evaluating projects

1. who chooses the items to value (EE) or criteria (MCDA),
2. how, once chosen, values are measured (via WTP or not),
3. the implications of additivity (substitutability) on equity and sustainability, and
4. the use of linear or non-linear value functions.

Who chooses the criteria has obvious implications for equity and for the acceptability of any conclusions drawn (in terms of the representativeness of the decision process). In the EE cases the analysts decided on the criteria of relevance (based on their expertise and discussions with the study team). However, it is common practice in the initial stages of an EE valuation to use focus groups and pilot surveys to determine which issues need addressing. In the SMARTx cases, the criteria were chosen by the team, once again relying on their relevant experience and the group-value sharing model. For reasons of equitable representation in the decision-process and ensuring that relevant issues are included (see also Section 9.4.2), it may be necessary to also use focus groups as in EE or other means to find the issues of concern of the affected population. However, there are no fundamental philosophical, methodological or practical reasons why the phase of choosing criteria should differ between EE and SMARTx and both are consistent with the Consultative Model of decision-making.

However, differences as to which issues are addressed, and therefore how well equity and sustainability are considered arise in practice due to how they are measured. This, together with points 3 and 4 is discussed below.

9.6.1 Equity

Equity as a concept in EE is limited to distributional issues (which are dealt with via distributional weights) and/or the Hicks-Kaldor compensation principle. This in itself may limit the extent to which the policy directive to consider equity in decision-making can be addressed. Additionally, the use of WTP may itself introduce biases.

Equity can be introduced as an issue into CBA with the use of income distributional weights, though these are rarely used in practice. Although standard texts on CBA mention weights (e.g. Pearce 1983) practical guides such as the South African Government handbook on CBA (CEAS 1989) and that of the American Economic Association (A Leiman, University of Cape Town, pers. comm.), explicitly avoid them. Without knowledge of the differences in marginal utility of income between rich and poor, distributional weights are relatively subjective. For this reason work such as Van Praag's attempts to measure the utility of income in a cardinal sense may be a prerequisite if weights are to help in social welfare optimisation (van Praag 1978 and van Praag *et al.* 1982). Although weights theoretically promote equity, implying some trade-off between this dimension and others, the trade-off is not explicit,

the procedure usually in the hands of the analyst, and the mechanism by which improved equity arises obscure. As mentioned, CBA should find an efficient solution *given* the current distribution. Without weights, this does not automatically lead to more equitable social welfare, the new efficient solution may be less equitable, and (together with the compensation system, discussed below) may entrench existing inequalities. The reasoning is usually that appropriate macro-level interventions will occur to adjust for equity.

An example of where weights might have been used in the EE cases was when the tourism values found in KNP were combined with other values (e.g. from harvesting natural resources) from the catchment. The former might be given a lower weight as they applied to a more affluent sector of society. In the open case study, when the value obtained for poorer areas were compared with those obtained from richer areas, they might also have been given a higher weight.

The operation of the Hick-Kaldor criterion may also have effects on equity. Confronted by the weakness of the Paretian test of optimality in decision-making, Hicks and Kaldor separately argued for a compensation test. They suggested that social welfare could be said to increase if the party that gains from a project could compensate the losers, while retaining some net benefit. The reasons for the theoretical rejection of the compensation approach are covered in the literature (e.g. Graaff 1971), but become especially compelling in a developing world setting.

Firstly, the compensation principle requires only that compensation could be paid, not that it actually is paid (though this was not the case in Hicks's 1939 paper). It may consequently accentuate any existing asymmetry in distributions of income and wealth. Secondly, the compensation principle has the potential for a pro-rich bias. If one accepts that the marginal utility attached to an extra unit of income by 'the poor' is greater than that attached to it by 'the rich', then the prices the poor are willing and able to pay for a specified change in utility will be less than those which the rich would offer. The Hicks/Kaldor approach is, therefore, biased in favour of projects that benefit the already affluent who are more able to 'compensate' those who have lost utility. Within certain contexts, the fact that a group is hypothetically in a position to 'purchase' utility may be an acceptable criterion. However, in South Africa, a move to more equitable income distribution is a national goal, particularly as the current distribution remains a legacy of the apartheid era. It is also a goal that cannot afford to rely on a hypothetical long-run balancing of wins and losses.

For example, in the open-space study (Section 8.4), what does it actually mean that the residents of the poorer area had a lower average WTP for open-space? Do they value their open space less than those in the wealthier area? If the total desired level of open space were fixed, could 'the rich' keep all of it in their area and compensate 'the poor' for having none? (In fact, for altruistic reasons, 'the rich' might be willing to pay in order that 'the poor' do have open space). If 'the poor' were to spend their compensation on education or health facilities, their welfare might well also improve. But what happens in the case where compensation does not actually occur?

Thirdly, the implicit bias may be accentuated if the environment is used differently by rich and poor. Where people are directly dependent on the environment for factors of production, TCM and hedonic pricing techniques are inappropriate. If, on the other hand, resources such as food were valued at market value (use value) they would be undervalued if life depends on them. This is especially problematic when, as in most southern African economies, the formal unemployment rate is over 40%. The cash value of the livelihood offered by the environment is then not the issue, rather it is the lack of any viable alternative.

For example, the environment may be a source of direct subsistence inputs to the rural poor (e.g. harvesting of secondary and natural resources in the Sand catchment, Chapter 6), but of recreational benefits to an urbanised affluent elite (e.g. tourism value in KNP, Chapter 7). Not only did harvesting supply value in terms of food and an informal income, but also in terms of providing relative self-sufficiency and therefore self-respect and a link to traditional lifestyles. All of these values were included in the SMARTx analysis as inputs to three separate criteria.

Where a project will directly affect the way of life of a community, any market approach will fail. This may be illustrated by the proposed damming of Epupa falls on the Cunene river in Namibia: the Himba people will lose much of their land, including grave sites and grazing land, and change from a pastoralist society to itinerant labourers (International Rivers Network 2001, Ezzell 2001). A market approach would have to attach a pecuniary value to something unknown by the people concerned.

The use of WTP as a measure may itself introduce biases which are particularly relevant in the developing world where income distribution is highly skewed, the environment is a direct source of subsistence to many, and goods are not traded in a formal market setting. In relatively homogenous populations comparability of WTP might be acceptable (in that case comparability of individual utility might also be acceptable). But the potential pro-rich biases of the monetary scale become relevant in heterogeneous populations. A rich man's WTP will outweigh a poor man's WTP because of ability to pay, especially if the relationship between money and utility is assumed to be linear. If one were to try to directly measure the utility of or preference for alternatives to a poor and rich man via direct scoring, the answer might well be different to that found with WTP. Even in a 'fully monetised' society, using WTP produced rank reversal in choices as compared to using other scales, including simple rating on a scale of 1 to 10 (Vatn and Bromley, 1994).

This type of reversal also occurred in the open-space study (Section 8.4), where respondents were asked to rank order preferred options, but in many instances their WTPs indicated different rank orders. These two orderings were both perfectly rational, but based on different issues: WTP was limited by income and protest at the payment vehicle, while the preferential rank order had no such constraints.

Where projects' impacts are borne by a non-monetised public, it may be significant that choices made in SMARTx are not based on WTP or willingness to accept, but purely on preferences and the existence of mutually agreeable

compromises. Obviously, financial implications, including those of implementation costs are relevant. In developing countries it might be that social and environmental gains are found to be large, but the government or other agent feels the costs of implementation are too high because of shorter term priorities. The inclusion of all types of value is therefore that much more important.

In contrast to EE-CBA, SMARTx can address equity issues directly by using improvement in income or non-income equity as project selection criteria. For example, 'land equity' and 'river access' (an equity issue) were included in the Sand case, 'personal well-being' in the Maclear case (which included consideration of the breadth of choices available), and impacts on the expenditure of poor households in the Cape Town water study (as well as health and employment). It can also improve 'procedural equity' by allowing stakeholders to participate in the process, permitting 'the poor' to address particular issues themselves. This means that the interests of stakeholders are not only reflected by a survey of their WTP for various goods and amenities, but also by their direct input in the decision-making process. The use of scores and weights obtained from stakeholders rather than income distributional weights takes the decision out of the hands of the analyst and places it into those of the communities involved. Thus, there may be 'content equity' in the direct inclusion of equity issues and "procedural equity" in the direct inclusion or representatives (McCarthy *et al.* 2001, Chapter 1). If growth and equity, or equity and the environment are seen as competing paradigms, then the introduction of equity as an explicit decision variable (as was the case in the Sand case) may be crucial to legitimise the decision-making problem.

The additive form of SMARTx, implies that losses in one criterion can be compensated for by gains in another criterion (as is the assumption with CBA). In the applications described here, applying the group-value sharing model meant that the decision problem was described in terms of impacts on criteria rather than impacts on stakeholders or specific groups (although effects on groups are still implicit). However, in some situations it may be necessary (for practical or political reasons) to know explicitly which groups receive the benefits and which bear the costs, and how large these are. These could perhaps be examined using the indirect compensatory values as demonstrated with the ecological criteria in the Sand case (Section 6.3.2). Extensive sensitivity analyses would be required to highlight to which criteria (and by implication which stakeholders) results were most sensitive. The sensitivity analyses, together with the selection and analysis of criteria (ranges, non-linearities, weights), should give the decision-maker and stakeholders an idea of the real impact of moving from a group's best option to their second best, increasing the likelihood of finding consensus or compromise. The ability to include non-linear relationships may also be important in including considerations of equity in SMARTx (e.g. the marginal utility of money may be different to different groups).

In general therefore, SMARTx can avoid the potential pro-rich biases inherent in CBA, provided that both rich and poor stakeholders are represented, the criteria, scores and weights reflect their values (in a non-monetary sense) and preferences are not governed by ability to pay.

9.6.2 Sustainability

At a practical level: weak sustainability - assuming substitutability between environmental, economic and social capital - can be handled in a utilitarian framework by asking: “Will the total utility decrease as a consequence of a project?”. Strong sustainability - rejecting substitutability for some elements - cannot effectively be dealt with in terms of utilitarianism or any form of weighted aggregation as it implies lexicographic preferences. The use of some concept of satisficing using lower bounds or reference points, like safe minimum standards, for some elements may instead be considered – however these are not then necessarily efficient. Both CBA and SMARTx are ‘compensatory’, because the summation of costs and benefits or scores means that increases in one criterion or attribute will compensate for decreases in other criteria. Both CBA and SMARTx therefore fall into the framework of weak sustainability. ELECTRE, a partially non-compensatory outranking MCDA approach (see Section 4.1.1) might allow for ‘stronger’ sustainability than SMARTx.

There are several influences on the extent to which sustainability can be addressed within the two approaches: the use of monetary measures or WTP, substitutability, discount rates, linear and non-linear value functions, and weights.

WTP and monetary scales

Sustainability is not dealt with directly in EE-CBA. Total economic value of the environment is made up of the sum of use value, and non-use or existence value, and different techniques have been developed to measure these different values (see Table 3.1 and e.g. Pearce and Turner 1990). The value of ecosystem goods and services is expressed by individuals’ WTP. For example, existence value is found with CVM, direct non-consumptive use like tourism with TCM, and direct-use like subsistence natural resource harvesting by finding the equivalent market value of the harvest. It is assumed that correct pricing of environmental effects and the use of the ‘correct’ discount rate in the cost-benefit equation will have sustainability consequences. However, the cost-benefit equation and its economic context are based on economic efficiency, not sustainability, and there is no reason to assume that efficient allocation of correctly priced environmental attributes will be sustainable (see the quote by Daly at the beginning of the next chapter).

CVM (and variations including CA) has been considered the only technique available in the EE paradigm for the measurement of the existence value of environment, including the value of biodiversity (although attempts have been made to value certain species based on their potential pharmaceutical value). A voluminous and vociferous literature has revealed a multitude of problems with this approach (e.g. Vatn and Bromley 1994, Angelsen and Sumaila 1995) which have been supported by the cases here. For example, the values obtained have been regarded both as overstating and understating the value of the environment, and the view is that CVMs tend to ascribe more value to the spectacular or charismatic than the functional (Vatn and Bromley 1994). Therefore, they do not adequately address sustainability, which relates to the healthy functioning of ecosystems, which may hinge on exceedingly uncharismatic species (e.g. macro-invertebrates in rivers), which might raise zero or even negative WTP bids. Existence value in fact says nothing about sustainability, but rather measures consumer or citizen preference.

Where indirect use values are concerned (e.g. the value of a close to natural flow regime) CVM is also sometimes used, but this may also be problematic (Angelsen and Sumaila 1995). The average person may not be aware of, for example, the health effects of altered flow regimes in rivers (e.g. increased levels of bilharzia), or be able to give a realistic WTP or WTA for resources upon which they may be directly dependent. In economic terms they have no existing demand curve for environmental functions so value cannot be attached to their loss or gain. But the change in flow regime and associated environmental effects may also have no clear effect on downstream human users (or there may be no human users). In this case one has to again rely on existence value.

If direct-use values (e.g. harvesting) are modelled to decrease with a change in ecosystem quality (e.g. soil erosion), the resulting decrease in the monetary value of the harvest will indicate changing economic efficiency, and, indirectly, environmental sustainability. It might be more straightforward to use the change in quality (e.g. tons of soil per year) directly as an ecological indicator and the change in monetary value of the harvest as an economic indicator.

Projects may have impacts on the environment (and society) which, although discernible are not easily monetised, and so may be omitted, or wrongly valued (Schulze 1994). Thus, especially in developing countries, it is likely that the monetary numeraire used in CBA will obscure the inclusion of environmental services in project assessment (van Pelt 1993). However, money is an available and well understood trade-off scale for many goods. The rationale has been therefore that we are used to money measuring the relative value of goods on the market, and if reliable hypothetical or surrogate markets can be established for non-market goods, money will similarly be a reliable common scale. However, the evidence here and in the literature casts doubt on the reliability of the scale for anything other than economic efficiency criteria. Also, the requirement of a monetary scale implies that 'the environment' and other factors need to be considered as tradable commodities. Although a perfectly rational requirement, it is difficult for many non-economists to see 'the environment' in these terms (Vatn and Bromley 1994), which is another reason that CVM has excited so much controversy, particularly in the light of the irreversibility of many environmental impacts. The relatively recent field of ecological economics attempts to redress the lack of appropriate recognition given to human dependence on functioning by placing economic activity within an ecological framework, and sometimes by using energy as a common measure. Again, lack of data on an appropriate scale may be a limiting factor, particularly in developing countries.

The problems illustrated seem to have arisen because the use of WTP as an indicator of preference (or utility) seems to have been confused with a belief that only WTP can 'count'. While this may be true for CBA it is not true for social choice in general. Although, in economics since Bentham, money was felt to be the only practical route to measuring utility, it was supposed to be a proxy rather than an end in itself. More direct measures of the value of the environment are apparently available using SMARTx. Pearce and Seccombe-Hett (2000) say that "...monetisation is simply a convenient means of expressing the relative values that society places on different uses of resources. Valuation is a means of measuring public preferences for environmental resources and is not a valuation of those

resources in themselves (so-called intrinsic value)". If this is the case, then there seems no reason to use money at all, if SMARTx can provide, as is its purpose, relative values of environmental effects in a way that will indicate sustainability rather than pricing efficiency. Ecological effects can be measured directly in SMARTx (as in the SMARTx cases described here) using whatever measure seems appropriate (natural or constructed) and including, where necessary, 'sustainability thresholds' (see next section). Nevertheless, all the diverse impacts need to be comparable in some way. Rather than using money to achieve commensurability, SMARTx relies on the construction of an abstract common trade-off scale between criteria, using the two stage process of scoring on interval scales and eliciting swing weights.

An advantage from the perspective of developing countries is that the use of a preference rather than monetary scale means that environmental choices are not limited by a group's ability to pay (and no adjustment need be made in the form of income weights). For example, using SMARTx, the ideal alternative in terms of environmental conservation, can be based simply on the fact that it is the ideal, not on whether people want to, or what they are able to pay (however much adjusted), for this ideal situation. Although the value functions derived may be translatable into monetary values (such as the indirect compensatory values in Chapter 6 and Section 8.1), these are not expressed *a priori* in an unrealistic market choice setting (Vatn and Bromley 1994). However, it is crucial to include the actual (financial) costs of project implementation in the MCDA exercise.

The evaluation criteria selected in the SMARTx cases described, were naturally grouped into broader criteria or objectives which reflected the potential achievements of the alternative in terms of the objectives of economic growth, equity and environmental sustainability (see also van Pelt 1993, Faucheux and Froger 1995). The criteria included a number of ecological indicators (e.g. aquatic and terrestrial habitat and species diversity, soil erosion, and the risk of invasive exotic species as in the Sand case) which contributed in some inexplicit way (reflected by weighted addition) to sustainability. The lower level measurable attributes or criteria were measured on natural scales (e.g. number of landtypes-Maclear), direct scoring (e.g. habitat diversity-Sand) and constructed scales (e.g. effect on estuaries-Cape Town water). However, in the SMARTx cases described we did not adequately deal with the secondary effects of environmental quality changes (e.g. the effects of soil erosion on commercial and subsistence farming activities). These could be informed by EE valuation.

Substitutability

The weak sustainability implied by both CBA and SMARTx means that environmental degradation may be compensated for by improvements in other criteria (e.g. income). Clearly, there are limits within which this would apply, and so safe minimum standards, thresholds or the precautionary principle are often invoked. SMARTx allowed us to develop non-linear relationships between quantitative criteria and their values, allowing for the specification of thresholds. These thresholds might allow for a slightly stronger sustainability criterion than would be implied by a linear additive model.

For example, the number of land-types preserved in the Maclear case was a non-linear criterion indicating the initial steeply increasing value as the number increased. This would mean that dropping below a certain number of land-types would be severely penalised in the summed scores. More stringent penalties might be introduced by setting non-negotiable minima.

In the Sand case, a logarithmic value function was chosen for operating margin as a realistic reflection of decreasing marginal utility of income. This in turn meant that depending on the current level of operating margin, one would be willing to 'give up' different amounts for an improvement in, say soil erosion. The non-linearity of operating margin had a profound effect on the interpretation of the trade-offs between scenarios in the Sand case, and the implied value of a sustainability criterion such as soil erosion. For example, the indirect compensatory value of a 100 point change in soil erosion ranged from R1.5 mill to R31.6 mill depending on the starting level of operating margin or in other words, the scenario from which one was moving. (While EE valuations usually assume linear relationships (e.g. between WTP and value), we included non-linear relationships between levels of the attributes in CA models 2, 3 and 4 in the KNP case, and in model 2 in the Hluhluwe case.)

Halvorsen *et al.* (1997) examined implied WTP derived from decision panels and from CVM and CA surveys. They found that these were generally higher than the CVM values and similar to the CA values. However, general conclusions cannot be drawn as the scenarios presented to participants were different in the three cases they describe. Using the implied WTP of the indirect compensatory values would mean one would need to address a new set of philosophical questions. For example, whose WTP is being reflected?. If this is a 'social' WTP, is it valid in economic terms or not (as it is relative not absolute)? Can or should this 'social' WTP be 'divided by' the population to find individual WTP? The latter calculation would be implied WTP in terms of OM and not individual WTP, and it would be implied WTP of those involved in the various commercial activities in the catchment, or rather, what society, as represented by specialists and stakeholders, think such people should be willing to pay.

Discount rates

The use of different discount rates in the cost-benefit equation as a means of changing the importance given to future effects involves an implicit trade-off. However, it is not always clear what effect changes in the discount rate have on intergenerational equity or on the promotion or rejection of environmentally damaging projects. By definition, applying discounting means that the value now of environmental benefits in the future will be small, although we cannot bank natural resources to earn interest as is implied by the inverse of discounting, and as applies to real money (even in the case of renewable resources, if the monetary interest rate is higher than the species' growth rate it pays to harvest to extinction). The use of high positive discount rates which is standard for CBA in developing countries, is likely to skew development towards projects that offer short term gains. On the other hand, high discount rates that penalise expensive projects may also promote intergenerational equity and long-term environmental quality. One option could be to apply different discount rates to different types of impacts. SMARTx may, of course, also lead to unsustainable results if long-term issues (e.g. ecological health) are not included in the criteria used to assess

alternatives. In theory, discounting could also be used in MCDA (e.g. Prelec and Loewenstein 1991, Stewart 1998), but its use is associated with more complex methods than SMARTx. The issue of different impacts occurring over time is usually dealt with in *ad hoc* ways. For example, in the Sand case, different impacts would be felt at different times and for different lengths of time. With some discussion during the workshops (and the fixing of a time horizon of 20 years), the view was taken that the weighted addition procedure reflected these sufficiently for our purposes. However, there is room for more research on use of discounting in MCDA.

9.6.3 Summary discussion and proposal for Section 9.6

In summary, as a practical measure for market-based goods and use value, money is an obvious and natural measure to use (particularly as a measure of economic efficiency), and where monetary values are easily determined, CBA may be an appropriate decision tool. However, the techniques for measuring both indirect (TCM) and non-use (CVM) values have been shown (here and in the literature) to produce unreliable results, and to require large resources of time, money and econometric expertise to produce even these. Some feel that the use of monetary units does not imply any bias towards goods sold in a market and that conversion to a monetary scale is not problematic (e.g. Pearce 1983, Angelsen and Sumaila 1995, Pearce and Seccombe-Hett 2000) and does not produce biased or incorrect results. That money is commonly used as a measure of value, however, does not indicate that it is a desirable, sensible, or possible measure in all cases (Vatn and Bromley 1994).

Therefore, in a broader social choice framework than CBA, it would seem appropriate to seek measures that do not require monetisation. The ability of SMARTx to directly include criteria measuring sustainability (ecological impacts) and equity (in terms of directly related criteria) seems a distinct advantage over the assumption that either weighting or monetisation will address these. However, as mentioned, the SMARTx model would need to be augmented in order to represent the broader population better (see Chapter 10 for discussions).

On the other hand, academic and decision-makers' interest in environmental valuation methods, has helped to demonstrate the importance of ecosystem services, although, unfortunately, this has also meant that some studies have estimated values quite out of any context, rendering them almost meaningless. Also of great importance is the contribution made by EE through the development of the typology of different types of value provided by the environment (Table 3.1).

For reasons of practicality and transparency, the additive model seems unavoidable. However, equity and sustainability can better be included through the careful use of non-linear value functions, thresholds and constraints than by assuming linear functions.

9.7 General practicality and provision of aid

The decision-aiding context of the case studies here, means that issues other than the theoretical purity and coherence of the methods are important. The practicality of the methods are compared in terms of their ease of use, flexibility,

time requirements, expense and transparency (or understandability), and in terms of their general contribution to facilitation of the decision-making process. Some of the comments about SMARTx arise from a questionnaire sent to participants, the results of which are given at the end of this section (Table 9.1 and Figure 9.5).

9.7.1 Costs in time and money

The relative costs of SMARTx and EE-CBA are of particular relevance to cash-strapped developing countries. EE surveys have been reported to sometimes cost more than the environmental benefit or cost they are intended to measure (Kontoleon *et al.* 2001), while adequate survey design and testing may take more than 30 months (e.g. Carson *et al.* 1994), require samples sizes of over 1000 (referendum CVMs generally need larger samples sizes than open-ended CVMs). Because of the need for large sample sizes, pilot testing and enumerator training, survey-based EE methods are expensive. For example, after two years, the broader study (Mander *et al.* 2001) of which the KNP study formed a part, was not able to produce monetary values in a form in which they could be aggregated with the KNP values (this was probably a result of both the complexity of the work and a lack of problem structuring). In the light of the expense of CVM surveys, and the apparently unreliable and biased results, running such surveys seems an inappropriate approach in general, and in developing countries in particular. (Of course, in other decision contexts, where a market is established, it may well be appropriate (e.g. Whittington *et al.* 1992)).

SMARTx as applied here was relatively inexpensive in terms of both time and money. However, in comparing costs we must be careful to compare like with like. The evaluation of alternatives includes whatever ecological, economic or other studies were necessary to be able to compare them, not simply the four or five workshops which comprised the process of comparison. Depending on the complexity of the study and whether many groups need to be accommodated in separate workshops, the costs could escalate dramatically. Aside from this, the SMARTx method imposes a high analyst and participant burden, to ensure that all stages are adequately addressed. In particular, participants need to be prepared to commit themselves to attending a series of workshops and ensuring that they apply their minds to scoring and weighting, and the analyst has to ensure that the process is as painless as possible, while remaining comprehensive and theoretically valid.

9.7.2 Simplicity, flexibility, ease of use, transparency and participation

It is especially important in new democracies that decision-making is transparent to those affected as well as participatory, requirements that may make EE-CBA an inappropriate tool. The issues of participation and transparency are also linked to that of equity, earlier referred to as procedural equity.

In terms of transparency, most of the concepts relevant to SMARTx can relatively easily be explained to participants, and the scoring process and addition of scores is easy to grasp. Transparency was one of the reasons for using SMARTx as other MCDA methods (e.g. 'pure' MAUT) are complex and off-putting to participants (e.g. Smith 1992), and may suffer from the same lack of transparency as EE-CBA approaches, although offering potential advantages in terms of including uncertainty. However, one respondent to the questionnaire found, for example, the

sensitivity analyses difficult to understand. This was a written sensitivity analysis, rather than examination during the workshop, and this highlighted the need to have a dedicated workshop session for evaluation of sensitivities and implied trade-offs (with the analyst preparing say indirect compensatory values etc. beforehand). The same respondent also felt that it would be difficult to translate the process to decision-makers (Table 9.1). We found that while participants found the scoring and swing weight concepts accessible, elicitation of weights in particular was quite hard work (although the fact that participants are 'forced' to apply their minds to scoring and weighting is a positive thing) (Table 9.1). However, there was general consensus about the weights in all the cases (and where not, the range of values was included in the sensitivity analyses). In the Sand case, we found that participants created strategies of their own for finding within criteria group weights. For example, the ecologists asked themselves which rehabilitation strategy they would undertake first if they only had the funds to undertake one, while the social group used the area of grazing land (together with discussion of particulars of scenarios) as a proxy for 'land equity'.

SMARTx is flexible in terms of the level of sophistication of value elicitation tools and analyses (see Section 5.3.4). For example, sophisticated software may be used to interactively give scores to alternatives, or paper thermometer scales may be used or beans distributed amongst alternatives in a matrix. Also, an analysis may relatively easily be done at a 'back of the envelope' level or with a high degree of information and detail. For example, in the Maclear case, the scenarios and criteria became progressively more defined throughout the process, starting with fairly broadbrush scenarios, to ones in which the exact location and extent of potential forestry expansion was known. The participants may well progress from the former, and stop before arriving at the latter, but at a point where they are comfortable that there is sufficient information and clarity to make their decision (Watson 1981). In contrast, one cannot really do a 'rough' EE valuation or a rough CBA, without it effectively becoming MCDA. However, although value and weight elicitation techniques for use with innumerate or less-numerate participants are available (e.g. from the participatory rural appraisal methods), these were not tested in these applications, and there is a need for more extensive use and development of such approaches. SMARTx is also flexible in its ability to integrate different types of value (quantitative, qualitative, tangible and intangible. This was regarded as one of its main strengths by the respondents to the questionnaire (Table 9.1).

The involvement of the public in SMARTx is more limited than EE surveys in terms of numbers, but higher in content, as different stakeholders can participate directly in the process of choosing criteria, and providing scores and weights. Trade-offs between the issues of concern (e.g. equity, environmental sustainability and economic efficiency) are the focus of workshop discussions and analysis and are not hidden within an aggregation of monetary values of several different types and from several different sources. Also, the range of alternatives considered may be expanded by suggestions from stakeholders and not bounded by the options originally tabled (Gregory *et al.* 1992, Munasinghe 1993, Keeney 1996), thus encouraging creativity in problem solving. The use of EE valuation surveys may help diminish the role played by 'special interests' (lobby groups, big business) who have concentrated costs or benefits or have the time and staff to attend meetings (Loomis 2000), whereas the already disempowered might not. A process of ensuring representativeness at SMARTx workshops is needed, and whether this is achieved may remain controversial (approaches are discussed in Chapter 10).

Participation in EE-CBA is broader than that of SMARTx in that people's WTP is taken from a larger sample. However, although money is a familiar (and transparent measure), the EE process is not transparent to the public or decision-makers, and participation of the public is essentially passive.

First, the purpose of EE valuation surveys (in terms of the use of the values provided) is not transparent to the respondent, leading to the suggestion that participants should be asked to give "informed consent" to the use of their values, or if not, that they are regarded simply as "gist" values. Fischhoff (2000) suggests that respondents often give "gist" responses: "we generally support better river health" rather than "contract" values: "we are willing to pay a 24% entrance fee surcharge to fund river health improvements" (the KNP study), or "we are willing to pay 20% more to see more lions and rhino" (the Hluhluwe study). Clarification of these differences is important so that where values are more realistically of the "gist" type, decision-makers do not interpret them as "contract" values that can be realised (e.g. by using results of a TCM to increase or initiate entrance fees). This might tend to increase any environmental inequities present as the affluent generally have better access to recreational facilities, and increased entrance fees would further limit access of the poor.

Second, the effects of discount rates and income weightings, if used in cost-benefit equations, are not obvious to the public or decision-makers (Munda *et al.* 1994a) and are determined by ethical criteria in texts, by analyst, or by government (Dubourg and Pearce 1994).

Third, in general, although less so in the Hluhluwe case, the overall design and analyses required were rather complicated (even though the particular econometric and statistical approaches were straightforward compared to those in the literature). The complexity of the KNP and open space studies tended to focus attention on the resulting Rand values (and high standard deviations tended to be ignored), rather than on clarifying values and differences between scenarios. The potential lack of transparency in the KNP study is particularly problematic, as this was measuring just one of the values to be included in evaluation of river health scenarios alongside valuation of other impacts (some from more and some from less complex methods). The analyses are difficult even for economists to do correctly, and normative assumptions are hidden in the apparent objectivity of the monetary measure (Watson 1981). Decision-makers are generally not in a position to understand the subtleties of normative assumptions, and of survey, sampling and econometric complexities and the consequences of these on results (Weston 1994, Pearce and Seccombe-Hett 2000). This "considerable analytical effort" required by EE-CBA cannot be limited by doing a 'rough' analysis, although this is possible with SMARTx (Watson 1981 p. 245).

While it can be argued that MCDA can also suffer from lack of transparency, it is our contention that the participatory nature of SMARTx, and the simplicity of SMARTx compared to other MCDA approaches, greatly increases transparency, as compared to EE.

Finally, a conclusion of all of the case studies reported here (both EE and SMARTx), was that people found it relatively easy to give ratings and ‘percentage importance’ (rather than rankings) and thus provide some form of strength of preference information (see Section 9.5.2).

9.7.3 Facilitation

The problem structuring focus and facilitation associated with the SMARTx process, appears to be an important part of the ‘value added’ and helps clarify relevant issues and it is relatively inexpensive in terms of time and money. In particular, the ‘value tree’ has proved to be an invaluable visual aid. The lack of a clear problem-structuring process in the EE cases seemed to be detrimental in terms of overall project management and cohesion.

Application of the SMARTx process helped to clarify the problem, facilitated communication between different groups, and allowed the integration of different types of value. This was apparent from the responses to a questionnaire sent to participants (end of this section). Generally there were three main areas in which the use of SMARTx was valuable. Firstly, the analyst provided general facilitation, and together with the process itself, this assisted in problem structuring, in a way which ensured at the very least that the problem was manageable. Second, SMARTx facilitated communication and the development of a shared understanding of the problem between different disciplines or groups. This does not mean that there was meek agreement across disciplines on all issues, rather that ‘robust debate’ happened within a shared understanding of broader goals. This also meant for example, that while early on in the process, participants might express worries about the weight elicitation stage, by the time this stage actually arrived, they found themselves largely in agreement. By this stage they had more insight into impacts on criteria within different disciplines or groups, and better appreciated the swing weight concept. The workshop context and emphasis on debate means that biases such as ‘framing bias’ to which SMARTx (as a stated preference technique) is vulnerable, should be reduced. Third, the process focused attention on the understanding and evaluation of trade-offs between criteria. Note that these three facets are the particular form which Krzysztofowicz’s ‘rule *d*’ took in these cases. Within this process, the value tree (particularly the graphic representation) was found to be particularly useful and illuminating by participants. However, it was clear that there was a need for a dedicated ‘debriefing’ workshop or workshop session during which results and trade-offs could be more fully explored.

9.7.4 Perceptions

Finally, an important issue is that of perceptions. SMARTx is regarded with some scepticism by some decision-makers and stakeholders in South Africa and is viewed as being “subjective”, “arbitrary” (Cornelius pers. comm.) and “complex” (Mander pers. comm.). It is important to note that these were *not* the views of participants in any of the cases, but rather of people who knew that SMARTx involved weighting different impacts, and felt that weights would be too subjective or arbitrary. Wenstøp and Seip also mention that the inclusion of subjective elements in MCDA was a reason why it was perceived as “lacking legitimacy” in Norway (Wenstøp and Seip 2001 p. 53). Ironically, one of the questionnaire respondents (see below) felt that the SMARTx approach had not enabled him fully to capture subjective elements (Table 9.1, respondent 9). However, until these perceptions are overcome, SMARTx (or MCDA) will continue to be considered as a poor substitute for EE, rather than as a potentially

complementary or even superior approach, and where SMARTx is used the risk is that insufficient time will be allotted to the process, and its use will tend to be limited to that of a 'scoring and weighting' tool. The subjective elements of SMARTx need to be acknowledged and the process needs to ensure that participants are clear when they may be attempting to objectively (e.g. ecological impacts) or subjectively (e.g. aesthetic impacts) quantify impacts on a value scale, or subjectively evaluate the relative importance of these impacts. In contrast to these perceptions, there appears to be a high level of interest in EE valuation techniques in South Africa and an increasing emphasis on their use. Also, the monetary values produced either by EE or the SMARTx indirect compensatory values were always of interest to the participants, and people seemed to relate to the values more than to scores.

9.7.5 SMARTx questionnaire

To gauge the usefulness of SMARTx, a questionnaire was sent to participants in the Sand case, the Maclear case and another case not reported here (see Figure 9.5 and Table 9.1). Nine people responded although not to all questions. In general (8/9) found the process useful, particularly in terms of the holistic integration of different views or factors, and (in so doing) gaining an insight into the relative importance of these. The majority also found the various concepts relatively easy to understand (thermometer scales and scoring (7/9), swing weights (5/9), sensitivity analyses (7/9)). However, some found it difficult to actually give scores and weights (which perhaps means that they 'applied their minds'). A few comments were made regarding lack of understanding (e.g. of sensitivity analyses), the difficulty of transferring the understanding gained through the process to the decision-maker, or the practicality of the methods.

Table 9.1. Comments arising from the questionnaire. R = respondent code.

Question	R	Comments
<i>Were the results of MCDA used/implemented?</i>	1	Yes, in the Sand River project Phase 2.
	2	Yes, the results were used to propose potential (best) land and water use practices in the Sand sub-catchment. This is now being implemented.
	3	Yes, integration of different aspects of the study. Also to calibrate (normalise) different issues.
	4	No, full project not completed.
	7	Yes (some), provided support for the reduction in forestry in the catchment, particularly in terms of a 'quantitative' analysis of the situation. General recommendations accepted.
<i>If result were not used, why?</i>	3	Lack of understanding of concepts.
	5	Didn't specifically use results in report but they may have provided insights that influenced it.
	7	In terms of specifics, it would be difficult to get buy-in to decisions from government departments. Not sure we could explain sufficiently to get them to 'understand'.
	8	Method not yet fully practical.
	9	Am involved with usual impact assessment which is a relatively subjective issue. I feel that the MCA results did not fit with my own analysis.
<i>What general insights were gained?</i>	2	Helped to prioritise the relative importance of a vast range of factors (in relation to each other)
	3	Relative importance of issues
	4	Go / no go decisions on projects are still made on the basis of single key issues which override all others. So MCDA works well on creating a hierarchy of the minor solvable problems
	6	An additional tool
	7	Insights in terms of decision-making support systems
	8	Too early to say
<i>Other comments</i>	2	A good way to incorporate different data from different disciplines to provide a holistic picture.
	3	Publish, present and communicate the approach.
	5	Useful in exposure to a different way of looking at things
	7	The sensitivity analysis was difficult to understand
	8	Helped in understanding of diverse views
	8	Feel that method may still be too theoretical / academic.
	9	If we spent some more time refining the system, 'emotional' and subjective issues will be brought into the MCDA in a meaningful manner

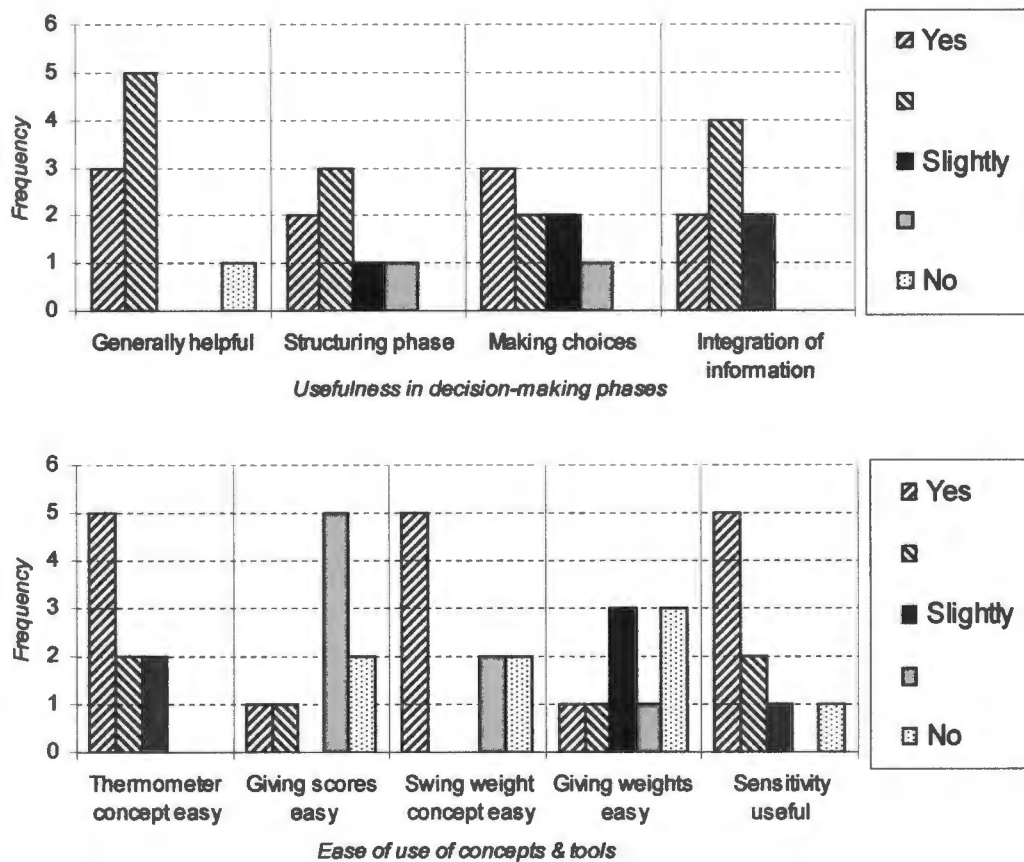


Figure 9.5 Frequency of responses to the questionnaire.

9.7.6 Summary discussion and proposals for Section 9.7

There are practical reasons why using SMARTx has advantages over EE, particularly in the context of environmental decision-making. In brief, SMARTx is relatively cheap, simple, transparent and flexible, and participants find that it clarifies 'the problem' and their values (Gregory *et al.* 1992). However, research is needed on elicitation of swing weight and of non-linear value functions, in particular in situations where stakeholders with wide ranges of backgrounds are participating. In addition, environmental (or any other) criteria can readily be included, whether they have a natural quantitative scale, or are qualitative, tangible or intangible, without the need to rely on potentially biased monetary proxies that may not reflect environmental quality or functioning. The integrative role of SMARTx was mentioned by questionnaire respondents as one of the most valuable aspects of using SMARTx. Besides the burden on participants in SMARTx, there is also a high analyst burden, and where possible it would be preferable to have both a facilitator and analyst to run the workshops. If CA or other surveys are used, a recommendation would be to exploit the ease of elicitation of scores, but to provide verbal cues to try to encourage similar use of the scale by different individuals in surveys, and to use differences in scores from the *status quo* (e.g. Roe *et al.* 1996) in the analysis.

9.8 Summary of performance and proposals

The concern is with practical environmental decision-making, and many would agree that a useful decision-making method (the implementation of a theory or philosophy) should provide useful information, elicit reliable and

appropriate values (which reflect what they need to reflect in order to 'fit' the philosophy), should help to clarify people's values and be relatively easy to apply, particularly so as to be transparent. The clarification of values is perhaps the more important, especially as it can be difficult to separate the method from the philosophy (i.e. if a philosophy is incorrect, can the method still be worth using (Hubin 1994), in terms of clarifying values or providing useful information). This chapter has discussed several theoretical and procedural issues raised in the case studies, which for the most part were in the context of land- and water-use decision-making. Problems with SMARTx related primarily to the representativeness of results. Problems with EE were of a more theoretical and methodological nature arising primarily out of the use of WTP as a measure of value (Table 9.2).

Table 9.2. A 'score-card' for the SMARTx and EE derived from the discussion in this chapter. A ✓ means that the approach generally performed well in this area, a ✗ that the approach generally performed poorly. Where either ✓ or ✗ is followed by the opposite, this indicates that there are specific areas in which the opposite applies, or ? indicates arguments either way.

	SMARTx	EE
1. RESONANCE WITH and VALIDITY WITHIN PREVAILING POLITICAL MILIEU and POLICY CONTEXT		
	✓✗ but group values may not be societal values	✗✓ but can supply efficiency measures and CVM may provide citizen values
Theoretical validity of EE	-	-
EE: What does economic theory require?	-	? but generally consumer values
EE: What values do stated preference methods produce?	-	? but CVM probably citizen values
EE: Interpretation – mean or median	-	?
EE: Interpretation – economic behaviour and/or preferences	-	?
EE: Interpretation - values aggregated from different methods	-	✗ - money equal value to all, consumer and citizen values equal, different levels of accuracy
SMARTx: Are group-values also citizen values?	✗ group values may not be societal values	-
2. GENERAL VALIDITY and RELIABILITY		
	-	-
Consumer surplus and the travel cost method	-	✗ TCM not appropriate in context of case studies
Measuring utility - ordinal/cardinal	✓	?
Preferential independence	✓ ...?only informal checks	?
Biases	✓ few	✗ many survey biases
Weights and additive model	✓ weights a focus of attention ✗ different methods produce different values But ✓ additive value model is robust to these	?weights usually only for income distribution
Uncertainty	✓	✗✓ sensitivity analyses
EFFICIENCY, EQUITY and SUSTAINABILITY CRITERIA		
Efficiency	✗	✓
Equity	✓ directly included	✗ indirectly included, biases against equity
Sustainability	✓ directly included	✗ indirectly included
GENERAL PRACTICALITY and PROVISION of AID		
Costs	✓ relatively cheap ✗ analyst burden ✗ participant burden and commitment	✗ relatively expensive ✗ analyst burden ✓ Low burden, if short questionnaire
Simplicity, flexibility, ease of use, transparency and participation	✓ generally simple ✓ deep involvement ✗ few people participate	✗ generally complex, difficult ✗ passive involvement ✓ large sample
Facilitation	✓	✗
Perceptions	✓ from participants ..✗ from non-participants	✓

- We felt that the EE philosophy, relying as it does on consumer values, was inappropriate within the prevailing political Consultative Model, and suggested that SMARTx could be used, with modifications, instead. We found that the values produced by CVMs were not internally consistent with the EE philosophy, but could be consistent with the Consultative Model. However, SMARTx appeared to provide similar 'citizen values' more directly

than via a WTP measure. Potential modifications to SMARTx to allow for its valid use within the Consultative Model are discussed in Chapter 10.

- Interpretation of values from EE methods is generally difficult. One does not know whether economic behaviour is indicated, whether to use the mean or median and how to interpret aggregated consumer and citizen monetary values.
- We found that in contexts where source populations have highly skewed and variable incomes, tastes and preference, the TCM produced unreliable measures of consumer surplus. Given that consumer surplus measures of this type would not be required within the Consultative Model, we therefore proposed that SMARTx could be used to indicate non-economic ‘surplus’, while actual revenues would remain as economic efficiency criteria.
- We proposed that surveys could realistically include rating rather than ranking, as the former is more efficient and powerful in terms of data requirements, and that with associated verbal cues and *status quo* benchmarks, these scores appeared to be reasonably consistently used.
- Given the potential biases introduced by monetary measures, and the indirect nature of the inclusion of equity and sustainability within EE, we proposed that SMARTx could more directly and reliably include these criteria.
- We found that, in general, SMARTx was simpler to apply and understand, and therefore more appropriate as an integrative framework. However, there was therefore a need to investigate the adaptation and practical use of participatory rural appraisal approaches to scoring and weighting within SMARTx.

Building on the issues raised here and potential solutions offered in this chapter, the concluding chapter proposes an approach to environmental decision-making (within similar contexts) that combines aspects of both paradigms but attempts to address their shortcomings. The proposed approach considers, in particular, the three main shortcomings of SMARTx that are shaded in Table 9.2.

10. Synthesis and recommendations

"The purely economic man is indeed close to being a social moron" Amartya Sen 1977, p. 336

"It is absurd to expect that market equilibria will automatically coincide with ecological or demographic equilibria, or with a reasonably just distribution of wealth and income"

Daly 1976, p. 251

The overall aim of this thesis was to examine two different approaches to support and inform environmental decision-making, namely, EE and value measurement MCDA. These two approaches were characterised as paradigms, the reasons for this classification becoming apparent through the thesis. The purpose of this final chapter is to synthesise the discussion of shortcomings and strengths of the methods in Chapter 9 and, based on this, to propose an approach (or family of approaches) for environmental decision-making.

Part I introduced the two paradigms by giving a brief historical background and a description of some of the specific methods of each approach. Chapter 2 provided a basic statement of measurement theory as applicable to both value measurement MCDA and EE. Chapter 3 described the development of the concept of utility and its use in economics and introduced the theme of the reliance on ordinal or cardinal utility, and the use of individual and aggregate WTP as a basis for social welfare choices. Chapter 4 briefly outlined the origins of MCDA after the time of von Neumann and Morgenstern, described some of the approaches within MCDA, concentrating on those based in value measurement theory, and including a brief description of different approaches to group decision-making in MCDA. Figure 10.1 summarises the intellectual contributors and theoretical developments by combining Figure 3.1 and Figure 4.1 from these earlier chapters. Chapter 5 provided a common general formulation for EE and SMARTx in terms of the search for a utility maximising alternative. Then, together with brief mention of the practical tasks undertaken within either paradigm, the EE valuation and SMARTx evaluation methods used in this thesis were detailed. This also clarified that SMARTx is a framework, a decision-making process and an evaluation method, while EE provides valuation tools and consequent information which may be used within a CBA (or other) decision-making framework (in this format then referred to as EE-CBA). Part II illustrated the application of the two approaches to various case studies. Chapter 6 described a study in which SMARTx was used to develop and evaluate alternative land- and water-use scenarios in the Sand catchment. Chapter 7, in a similar policy context, described the use of various EE techniques to find the effects of changes in river quality on tourism value in the Kruger National Park. Both of these studies were ultimately aimed at assisting in the choice of a particular river quality. Chapter 8 described two SMARTx studies, one developing and evaluating land-use scenarios in Maclear, and one evaluating alternative water supply augmentation and demand management options of Cape Town. Two EE case studies followed. One study valued the contribution of different attributes of a game park, and within this, different species of the Big Five. The second study valued different types of open space within Cape Town with a view to developing an urban open space system.

In Chapter 9 of Part III, the applications were discussed in the light of the background in Part I, problems of reliability, validity, practicality and resonance with the prevailing political model. Chapter 9 ended with a summary of these problems (Table 9.2) and some suggestions as to how these could be overcome.

We begin this chapter with a description of different decision contexts (Section 10.1) and a summary of the shortcomings of the methods and proposals for improvements outlined in Chapter 9 (Section 10.2). These are supported by additional ideas in the literature regarding future directions for investigations into social choice (Section 10.3). Then, in Section 10.4, a more detailed proposal for environmental decision-making is described, and compared with other new hybrid techniques (Section 10.5). The thesis closes with some suggestions for research necessary to investigate aspects of this proposal (Section 10.6), and final conclusions.

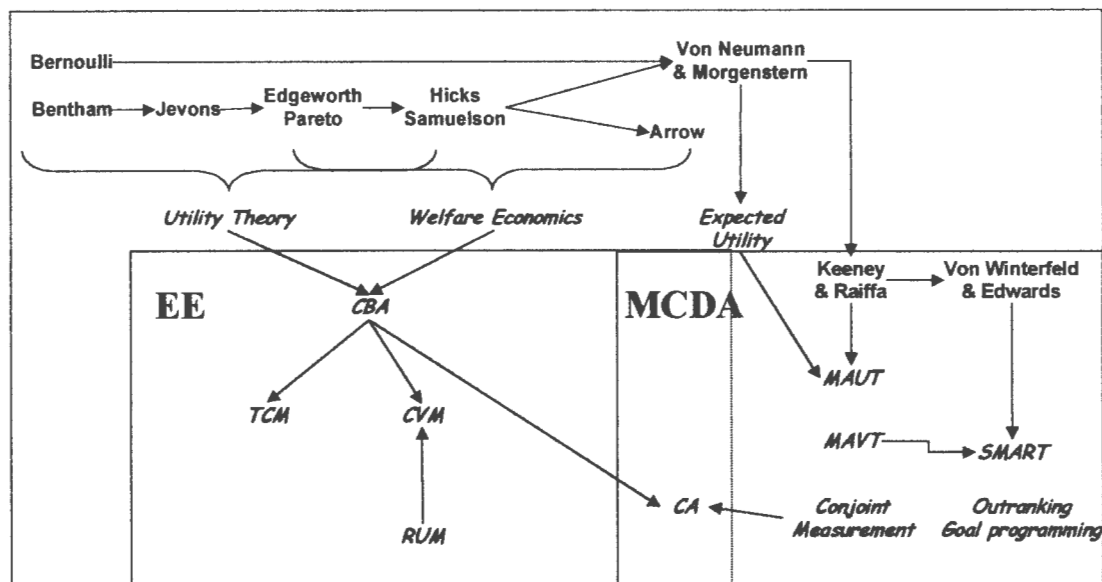


Figure 10.1. Intellectual and methodological links between EE and MCDA.

10.1 Decision and evaluation contexts

Before expanding on the proposed approach, we clarify with which decision contexts we have been concerned both in terms of the types of decisions and the type of political or social environment in which the decisions are being made. There are five distinct areas in which EE valuation has been previously applied (Figure 10.2): (1) the comparative evaluation of alternatives (projects or policies), (2) the evaluation of a single project or policy, (3) the valuation of a particular habitat or species or recreational site, (4) the valuation of environmental damages with a view to establishing compensation, and (5) the determination of ‘efficient’ levels of pollution.

In cases where there are no environmental and social impacts the usual CBA or cost-effectiveness analysis is appropriate. SMARTx is not normally applied in contexts (2) to (5), although there are situations in which it might be appropriate. In the case of single project evaluation (2), SMARTx might be still useful: aggregate monetisable benefits or cost may be compared against aggregate non-monetised benefits or costs leaving a final trade-off question: “are the aggregate benefits worth the costs?”. In the case of the valuation of a particular single habitat or species in

isolation, SMARTx is inappropriate. However, as noted by Pearce and Seccombe-Hett (2000), this type of EE application is often so context-free as to make the results fairly meaningless. Therefore, the context discussed below is that of (1) - evaluation of alternatives – resulting in the process outlined in the double edged box in Figure 10.2, and described in more detail in Section 10.4. Thus, the main area of interest in this thesis and in this closing chapter is the first: the comparative evaluation of alternatives. In other situations, especially (4), EE valuation is the only feasible approach, and the use of CVMs have become fairly well established, particularly in the USA, where results are accepted legally as a starting point in the judicial process (e.g. Arrow *et al.* 1993). In situation (5) establishing actual WTP may also be the only route, as this provides a means of implementation. In developing world situations, WTP surveys have also been very usefully applied in determining appropriate pricing for municipal water provision, where the current water provision is largely via private vendors (e.g. Whittington *et al.* 1992). CVM and TCM have also been used to examine price structures for game parks. In the latter case, while these approaches indicate the revenue maximising fee, which is probably what developing economies need to do, they do not, for example, address equitable access to the good in question.

The political, policy and social context is that of a developing country with a strong leaning towards consultative democracy. Thus the recommendations are proscribed by the contexts in which we have applied and thought about the methods. Occasional reference is made in the remainder of the chapter to the applicability of the proposed approach in other contexts.

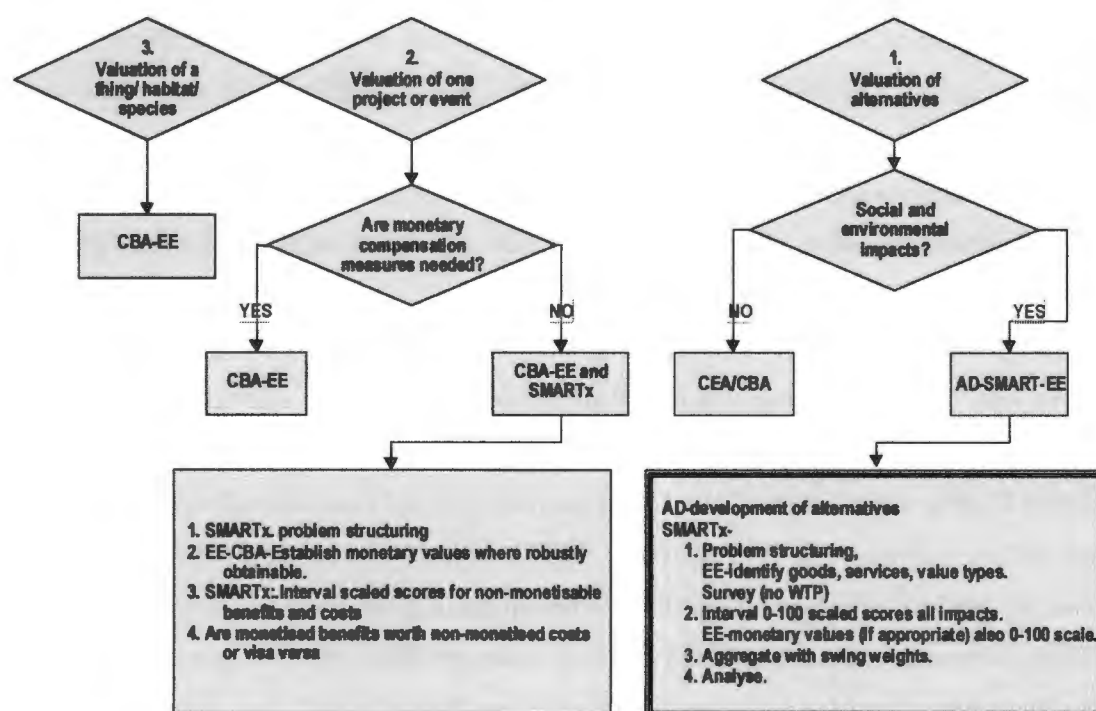


Figure 10.2. Flow chart for choosing an appropriate framework and methodologies. The process in the double bordered block is described below.

10.2 Summary of strengths, shortcomings and proposals

The approaches applied in the case studies were compared, together with their broader philosophies, according to four criteria, viz. resonance with decision and political context, validity and reliability, ability to include equity and sustainability criteria, and practicality. For the most part the problems in EE arose through the use of WTP as an indicator of preference, utility, or welfare, and the use of stated preference surveys that elicited WTP values that were not internally consistent with the EE philosophy. The main problem with SMARTx would arise if the group-value sharing model (applied here in Phase I projects) was assumed to also produce social values. We illustrate this summary of the shortcomings of the various approaches from the broader point of view of collective or social choice (Figure 10.3). Our argument was that (in the decision context described above), where there seems to be no external reason why the EE or Market Model should predominate in social choice applications, SMARTx would provide a better *framework*. Then also, scoring using value measurement theory was preferable to monetary valuation for non-economic or non-monetary criteria. Within this framework, EE methods would be used to provide values for certain impacts. These monetary values would be translated in to dimensionless value as for all other criteria.

We have suggested SMARTx as the overarching framework. Some additional points need to be made about this suggestion. First, we could perhaps substitute ‘MCDA’ as a framework rather than specify SMARTx, but a full exploration of the potential use of all MCDA methods was beyond the scope of this thesis. Specific MCDA methods may provide an alternative framework or may be appropriate at different stages within an investigation. For example, depending on the information available, MODM methods could be used to create a set of efficient alternatives for more in depth investigation. Outranking methods, may also be useful in filtering such a set, or fuzzy methods may help to classify alternatives into groups (e.g. Perny 1998). These may well be used in a ‘backroom’ technical evaluation context, before proceeding to evaluation with broader representation. However, at the stage of evaluation with which we have been concerned, where people with different specialities, cultures and levels of numeracy are to participate, our contention has been that SMARTx appears to be one of the simplest and most transparent of the MCDA methods. Second, before adopting such a framework, it is perhaps incumbent on the analyst to make explicit that the process rests on the assumption that individual and societal values are constructed through dialogue and interaction, and therefore that group processes are a useful way of finding them. Lastly, our justification for proposing SMARTx as the framework was based on its resonance with the Consultative Model of decision-making in South Africa, and on problems with EE methods as well as with the underlying philosophy. In other political systems where the emphasis is more firmly on the Market Model (e.g. the USA) this justification may not apply. However, further support for proposing SMARTx outside of any particular system is provided in the next section. Finally, alternative ways of dealing with the representativeness problem of SMARTx are given in Section 10.4.1.

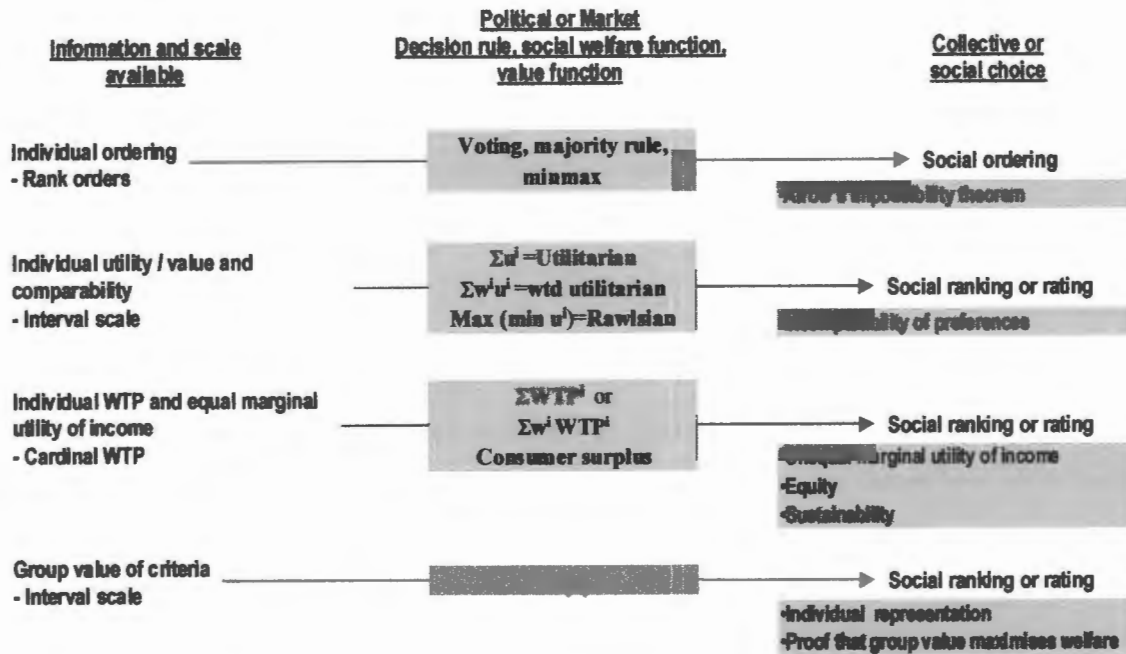


Figure 10.3. Approaches to social choice, showing the critical weaknesses or shortcomings of the approach in the bulleted lists on the right.

10.3 Social choice theory and environmental choice

Additional arguments for our proposal of SMARTx as a framework are suggested in this section. These arise out of suggestions in the social choice, economics and multicriteria literature about new directions for social choice theory, and suggestions in the environmental ethics and economics literature about ways of including the environment in the 'social welfare function'.

Most of our discussions have been concerned with 'instrumental rationality' (Hargreaves Heap *et al.* 1992) or 'substantive rationality' (Faucheux and Froger 1995), and has been 'consequentialist' (Sen 1997). In other words the emphasis has been on the rational analysis of consequences. This sort of analysis must obviously must play a pivotal role, but there have been suggestions from various quarters that there needs to be more emphasis on right procedures, 'procedural rationality' and the process of construction of preferences (e.g. Sen 1997, Rauschmayer 2001). Sen (1997) also contends that the emphasis on individual values may evolve to allow more explicit consideration of whether the relevant values in a particular context are preferences, choices, satisfactions, selfish or altruistic. These two issues are summarised well by Hargreaves Heap *et al.* (1992 p.215): "An alternative model of collective choice would be most likely to present it not as a process of preference aggregation, in which there is a mapping from a set of individual orderings to a social ordering, but as a process of dialogue in which reasons are exchanged between participants in a process that is perceived to be a joint search for a consensus... Such a dialogic concept of collective choice would necessarily work not with fixed preference to be amalgamated, but with preferences that were altered or modified as competing reasons were advanced in the course of discussion...". SMARTx, which emphasises dialogue and the development of a shared understanding between participants, is a process which could provide such an alternative model of collective choice. A process for developing preferences and choices would seem to be preferable to a model which either takes revealed market choices at face value, or in a

survey, creates a hypothetical market in which people have to make a rapid choice in an unfamiliar market for a good which they are not used to considering as a marketable good. Decisions which have long-term environmental and consequently social and economic implications, need to include ethical considerations, and allow for the evolution and revision of values (Sen 1995, Sen 1997, Wenstøp and Seip 2001). Anyone who has been harassed into buying something in a market in a foreign country where they have not quite found their feet or adjusted to the currency can understand that foolish choices may result!

Alternative forms of social welfare functions have been suggested to accommodate the nature of environmental goods and services as public goods on which we depend and which we value in their own right. In the normal social welfare function, the environment enters as a function of the individual i 's preferences along with other market goods $W=G\{u^1(b,x), u^2(b,x), \dots, u^n(b,x)\}$, where b is a basket of other goods and x a vector of environmental goods. The alternative proposed is that x also enters as an argument on its own: $W=G\{u^1(b,x), u^2(b,x), \dots, u^n(b,x), x\}$ (Johansson-Stenman 2002). In the SMARTx approach, the ecologists evaluated environmental effects that would not necessarily enter as part of any individual's preference function (particularly if this were represented by WTP). The weighted addition of scores in SMARTx may therefore correspond more closely to this social welfare function. While EE values are all anthropocentric, the use of SMARTx, implying alternative social welfare functions of this type, may allow varying degrees of anthropocentricity. The degree may be decided by the team as a metaproblem and informed by the survey proposed in Section 10.4.1. Extending this argument, one might want to state more explicitly the dimensions of a decision problem in terms of various ethical norms (Rauschmayer 2001).

10.4 Recommendations for practical aid for environmental decision-making

In the light of the discussions in this and previous chapters, an approach to environmental decision-making is proposed below, which is particularly appropriate in a developing world setting. The initial intention had been to combine aspects of both paradigms to have "methodological pluralism" (Norgaard 1989) and a resulting "multi-methodology" (Mingers and Brocklesby 1997). However, in the end, the process recommended rests firmly within the MCDA paradigm.

Abandoning the market-based use of WTP means that we have abandoned the economics basis of EE valuations. Instead, we have maintained the apparently less intrinsically value-laden theoretical basis of measurement and value theory. While SMARTx relies on subjective values, these are explicit rather than hidden behind the apparently objective use of WTP. However, it is important to include conventional economic values, monetisable impacts and the financial costs of implementation.

In terms of broader philosophies, the proposed approach sits reasonably well within Popper's refutation of positivism. Although there seems to be debate about whether Kuhn's view on paradigms and their shifts (e.g. Dow 1994, Argyrous 1994), supported the possibility of methodological pluralism or not, perhaps to the pragmatists amongst us, this does not really matter.

In brief, within the context described (Figure 10.2), it is proposed that the overall SMARTx decision-making framework, process and scoring/weighting 'toolbox' is used, together with alternative development as described in the cases (AD) (together therefore AD-SMARTx) and augmented as described in Section 10.4.1.

The process is therefore that described in Section 5.3, Chapter 6 and Section 8.1, augmented with option (3) as described in Section 10.4.1 below (or (2) in some contexts) (Figure 10.4). In addition to the descriptions in this thesis, detailed and accessible recommendations on all stages of SMART-type analyses are given in, for example, von Winterfeldt and Edwards (1986), Goodwin and Wright (1998) and Belton and Stewart (2002), and references therein. These are not repeated here. The process would also be augmented and informed by the EE paradigm in a number of ways. Firstly EE thinking can assist with the problem structuring stages (Section 10.4.2) by helping to identify the types of values for assessment, and secondly, EE valuation methods may be needed for the monetary valuation of certain impacts (Section 10.4.3). Any other adjustments to the 'business as usual' SMARTx approach are also mentioned below. The augmented SMARTx process might need to be renamed, but I will resist the temptation to call it SMARTy.

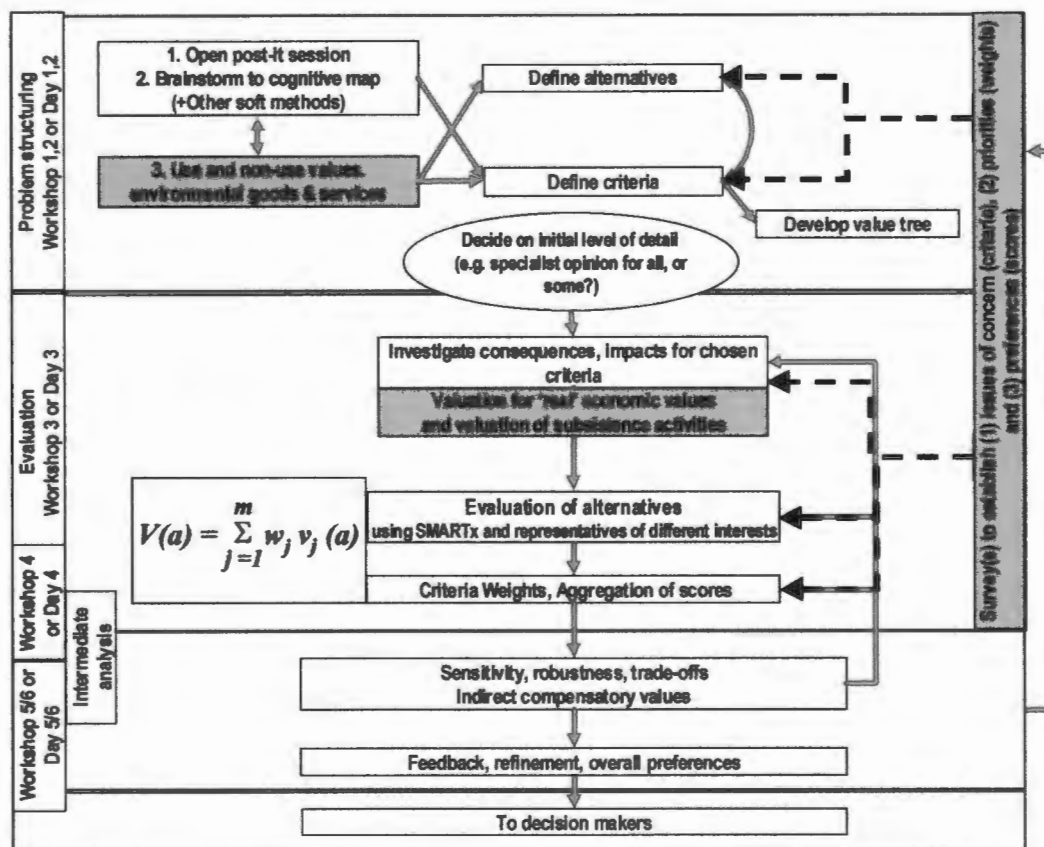


Figure 10.4. Proposed decision-making framework. EE contributions are shaded.

10.4.1 Adjustments to SMARTx and the group-value sharing model

The group-value sharing model applied in SMARTx develops group values through discussion and debate. Depending on the make-up of the group, certain people may value the alternatives from the point of view of their

discipline, while others may, in addition, have to represent a constituency (e.g. a sociologist might have to represent the views of the rural poor on issues of land equity). The analyst needs to ensure that the various roles people may have to play are clear.

The problems with the original SMARTx were largely related to representativeness. The number of people who can be involved is limited and this may be of concern to those who prefer the economic notion of aggregate individual values or fear that those involved are not truly representative, or might bias the process through self-interest. In any case, the process needs adjusting in order to conform to the Consultative Model. There are essentially three ways (which are not necessarily mutually exclusive) in which the representativeness problem of SMARTx as applied in this thesis could be addressed:

- (1) SMARTx workshops could be undertaken with different stakeholder groups separately, with the separate results being presented to the decision-maker for evaluation. While this approach might be useful in some contexts and possible where there were not 'too many' groups to consider, the final decision-maker would be faced with the task of establishing comparability between the different evaluations.
- (2) SMARTx workshops could be run separately with different stakeholder groups, but with a common value tree being established at an early stage. The evaluations of the different groups then makes reference to this common value tree, although weights of zero may apply to different criteria for different stakeholder groups. This has two benefits, firstly that groups become aware of concerns other than their own, and secondly that the problem is not presented as a conflict between groups, but if anything, as conflicts between criteria. Sensitivity analysis on the effects of different weights would still highlight for the groups and the decision-maker the specific consequences to different groups. An additional stage might search for a consensus set of weights using the Delphi or some other technique, or the decision-maker would decide. A variation of this might begin with separate groups, which combine at a subsequent stage. A process where stakeholders worked together (with breakaway sessions for particular stages) would be preferable, as this would allow for more justifiable comparison and aggregation of utilities (via the group-value sharing model), and increase the likelihood of consensus being reached. But, separate initial smaller workshops might also have advantages, because, if all stakeholders are combined at the start, some might be disadvantaged because of differences in status, power, eloquence, education, or familiarity with 'workshopping'. Separate initial workshops may also be necessary due to numbers or animosities.
- (3) SMARTx workshops could be undertaken with one group made up of relevant specialists and key stakeholders. A survey or series of surveys could be held at key stages in this process (Figure 10.5). The surveys could ask questions to raise 'issues of concern' (to be translated into criteria by the workshop participants and in some cases to become parts of alternatives), to prioritise these issues (to provide guidance for weighting), to evaluate the alternatives using simple rating questions or perhaps CA (to provide guidance for scoring and weighting). For cost reasons, it would be preferable to limit the number of surveys needed, but it might also be useful to run separate smaller surveys at different stages. This would mean that the public would be more involved in the

process, but also that the surveys would be better informed by utilising the information gleaned from earlier surveys. In a situation where a number of surveys are undertaken at successive stages providing inputs to those taking part in workshops, the group-value sharing model starts to have elements of the Iterative Open Planning Process mentioned in Section 4.1.3. Figure 10.5 suggests how information from the survey(s) (preferences of individuals) might be taken into account within workshops (social process of value formation), thus ensuring that workshop participants are directly informed as to the views of the general public and include these in the process.

Adjustments (2) and (3) might be appropriate in different contexts. There are practical advantages to (3). While surveys are expensive, they may be cheaper when seen as an alternative to a number of workshops being run for a number of different groups. Moreover, if surveys do not need to find WTP, they should be (a) simpler and (b) require smaller sample sizes. Option (2) also has potential representativeness problems because ensuring that all groups are included may be problematic, particularly as groups who feel they have been left out may well derail the process, and the workshops may still be dominated by the more eloquent, experienced or numerate. Also, process (3) may limit the undue influence of more powerful groups who have the time and money to attend a series of workshops. However, in preferring (3) to (2), we are in danger of doing the inverse of what we earlier did in preferring SMARTx to EE. Process (3) does substitute depth of involvement with breadth of representation. However, we believe that simple questions regarding preferences will add more value to the process than equivalent EE surveys assessing WTP. Lastly, some might prefer (2) as explicit separate assessments of benefits and costs to different groups are made. While we believe either (2) or (3) could be appropriate, the proposal below includes only (3).

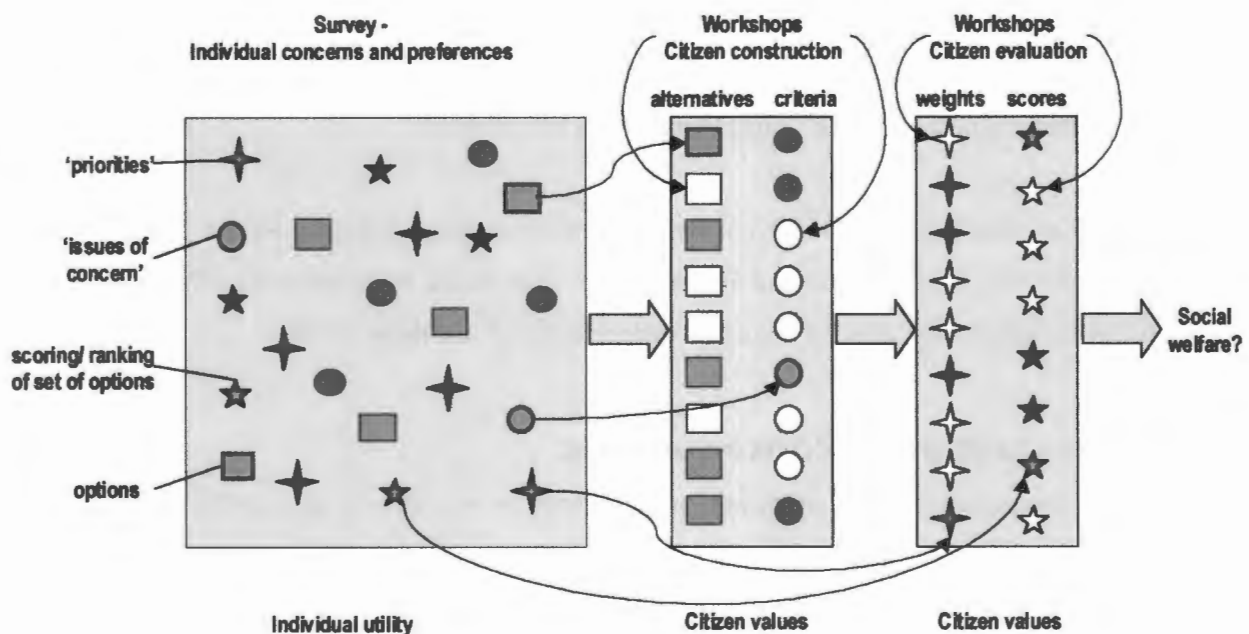


Figure 10.5. Inputs from SMARTx and a survey combined to inform social welfare decisions.

10.4.2 Problem structuring

Initial problem structuring including alternative development would follow the process outlined in Sections 5.3, 6.1 and 8.1 (Figure 10.4) and the references mentioned above. Within the context of projects with potential wide social,

economic and environmental impacts, there is a need for augmentation of the problem structuring stage, for broader input, and which makes use of some of the thinking which has evolved within the EE paradigm over the last few decades. The three sets of EE concepts that are relevant here are those of (a) environmental goods and services, (b) consumptive and non-consumptive use values and non-use values and (c) costs and benefits (including costs of implementation).

The explicit consideration of environmental goods and services is probably the most useful in this context, as this will highlight benefits resulting from healthy ecosystem functioning. The intention is not to require that all identified goods and services become criteria and/or are monetised (the normal ‘pruning’ of the value tree should apply (von Winterfeldt and Edwards 1986)), but rather to ensure that all consequences of ecosystem quality changes are explicitly considered in the scores and weights given. For example, while soil erosion was included as a criterion in the Sand study, the long-term downstream effects on agriculture were not included. It is assumed that the many values which might contribute to, for example, ‘existence value’ will already have been identified in the initial post-it session (e.g. species diversity, aesthetics). The role of examining use and non-use values would be to determine explicitly whether all aspects have been considered. For example, as discussed in the previous chapter, tourism has direct economic benefits in terms of park revenues and multipliers, which could be measured by criteria such as ‘gross margin’ and ‘regional economic benefit’. However, tourists’ experiences in terms of aesthetics, game viewing, perceived naturalness etc. might also be relevant, although not translated into economic behaviour: consideration needs to be given as to whether the criteria have captured these values (or whether they need to). Inclusion of these criteria may occur through an additional post-it session (e.g. “what are the important riverine goods and services in this catchment?”). Responses could be checked against a checklist of potential goods and services, to ensure that nothing vital was omitted. Using a post-it session rather than checklist format should limit the tendency to check every possible thing rather than only relevant issues. Finally, direct questioning as to “what will be the costs of this alternative” and “what will be the benefits”, might bring to light further issues.

The initial value tree constructed after the initial post-it/cognitive mapping stage (see Section 5.3), can then be updated. Because this stage offers the potential for many issues to be raised, more attention will have to be paid to pruning: i.e. the inclusions of only ‘relevant’ issues, so that there are not ‘too many’ criteria.

10.4.3 Evaluation of consequences

Information on the consequences of the alternatives might already be available or new studies (e.g. hydrological modelling) required. An important contribution of EE would be in the assessment of the relevant and relative market values of subsistence harvesting (e.g. of fish from a river) (Figure 10.4). Though not discussed in this thesis, this sort of approach was used in the measurement of the value of harvesting of natural resources in the Sand river case. Other innovative applications have been quite revolutionary in ensuring the inclusion of values usually ignored by conventional economics (i.e. because subsistence use is not reflected in the market). An important financial criterion to include is the cost of implementation, which may sometimes be unavailable in any detail especially in a Phase I project (e.g. in the Sand case). In that case, costs might need to be estimated (and scored) using direct specialist

opinion. Based on the previous discussions the conclusion was that CVM should not be used in the evaluation of land- and water-use options in developing world settings as it is inappropriate where skewed income distributions exist, it is expensive and the results are unreliable. TCM was felt to be inappropriate for anything other than local amenities for precisely the same reasons. In addition, neither explicitly address environmental sustainability issues.

In the absence of user friendly software which allows the inclusion of ranges of scores and weights where people are unsure of exact values (e.g. as in COMPAIRS, Salo and Hämäläinen 1992), ranges could be recorded and included in all sensitivity analyses.

10.4.4 Analysis and feedback

Unfortunately, it is difficult to tell in advance how many workshops will be required. In each of the SMARTx cases described here an additional workshop (or workshop session) would have been preferable, bringing the total number to five or six, depending on the case study. This additional workshop (or workshop session) would have been the dedicated 'debriefing' and feedback workshop or session that we felt had not been given sufficient time in our applications. Some of the implications of the scoring and weighting stages may be immediately investigated during the workshop, particularly if software such as VISA is being used (or if the problem has been set up in an Excel spreadsheet). However, the analyst will often need to perform more extensive analyses outside of the workshop setting. While this incidentally allows the participants to digest the process, scores and weights, it also means that they will benefit from a more detailed feedback from the analyst, and be able to refine and update scores and weights based on the feedback and the digestion process. The analyses (described in detail in the case studies in Chapters 6 and 8) include systematic assessment of sensitivity to scores and weights (and ranges of both where given), the elaboration of the implied trade-offs between criteria and the calculation of indirect compensatory values. The results need to be reported back to the participants with appropriate visual aids, and time allowed for participants to further explore sensitivities and trade-offs interactively, allowing them to verify that the values are acceptable to them.

10.5 Comparison with other combined approaches

One of the criticisms of CVM is that respondents are not familiar with the environmental good in question, and are therefore not in a position to give informed answers. The National Oceanic and Atmosphere Administration panel investigating the use of the CVM (Arrow *et al.* 1993) recommended that the analyst decide on the desired level of knowledge of respondents and suggested that this should be similar to that of the average voter in an election and that respondents should understand exactly what they are valuing. It is generally acknowledged that it is difficult to convey sufficient information in the time available for in-person or telephone interviews, especially as respondents may have very different levels of knowledge and cognitive abilities. With insufficient information, respondents tend to rely on cues (from the interviewer, for example) and heuristics and may not give sufficient thought to their statements (see for e.g. the biases in Sections 5.2.4 and 9.5.4).

Recently there have been attempts to combine information from individual preferences with that from expert or stakeholder groups (Kontoleon *et al.* 2001). The purpose of these methods is to improve the level of information and time available to respondents. These two studies, as well as others (e.g. Whittington *et al.* 1992) demonstrate that when respondents are given more time to think about their responses, their resultant WTP bids can be significantly affected. Two 'hybrid' methods reported in recent literature, which combine aspects of MCDA with EE valuation are briefly described below. These are the Market Stall method (Hanley *et al.* 2001) and the Valuation Workshop method (Kenyon and Hanley 2000). The two methods are still firmly based in the EE paradigm as they search for more 'informed', 'accurate' or 'reliable' estimates of WTP.

The Market Stall approach discussed by Hanley *et al.* (2001) attempts to deal with the issues of shortages of information and of time to digest in CVM and CA applications by setting up two group meetings. The first meeting presented information, gave time for discussion and asked for individual WTP of those present (for wild goose conservation). At a follow up meeting a week later, further points could be clarified and discussed and the WTP question was repeated. The average WTP from the market stall increased over the week, but was still about one third of that from a CVM carried out for the same issue, and had lower variance (sample sizes for the CVM and the Market Stall were approximately the same). The view was that the Market Stall WTP estimates were more reliable than the traditional CVM estimates.

Kenyon and Hanley (2000) discuss the use of Valuation Workshops that combine citizen juries and CVMs. The citizen jury (e.g. Brown *et al.* 1995) is an approach based on the idea of juries within the criminal law system. A group of about a dozen randomly selected citizens is provided with information by expert 'witnesses', whom they may question, and after deliberation, the jury provides recommendations on policy matters, or chooses between several options. Thus the output is not a 'value', economic or otherwise, but a choice (from which value might be inferred by regression). The authors therefore combined the citizen jury (with its qualities of participation, deliberation and preference construction) with CVM (with its quantitative outputs). The jurors completed a CVM survey, and were then asked to list the positive and negative aspects of the projects under consideration (floodplain restoration) and to rank order each list. They were given an opportunity to change their WTP after further questioning of the 'witnesses'. The average WTP increased slightly with the second offer, but was similar to that from a separate CVM run for the same issue, while the variance was much smaller for the Valuation Workshop (as was found in the Market Stall approach). The sample size for the Valuation Workshop was very small ($n=28$) compared to that for the CVM ($n=233$). Once again the view was that the Valuation Workshop WTP estimates may be more informed and therefore 'accurate' than those from the CVM. The citizen jury part of the process only adds a list of ranked 'pros and cons'. There are no real functional or process links between the two tasks, except to refine the WTP estimates, however the qualitative output of the citizen jury part provided useful information for decision-makers.

Both approaches therefore aim to (a) improve the reliability of WTP estimates with increased information and time for deliberation, and (b) augment the decision-making process with qualitative information. However, it is somewhat

questionable to use average WTP from a small meeting if the basis of the decision-making process is still 'economic'. That is, if the basis of a decision will be whether the monetised benefits are greater than the costs, this is fundamentally supposed to be based on aggregate individual utilities (as indexed by WTP). If this is the basis, then a survey of a relatively large sample seems to be the only justifiable approach. Another question is how to combine different types of information in the Market Stall and Valuation Workshop approaches. Citizen juries (as part of the Valuation Workshop) presumably are intended to provide societal or citizen views congruent with maximising social (rather than economic) welfare (Kontoleon *et al.* 2001). Also, the implicit aim seems to be to allow respondents' views to converge with those of 'experts' (Kontoleon *et al.* 2001), by increasing information and time. One wonders then why not use the specialists directly.

If there has been a move to the use of citizen values as a basis for decision-making, it seems that either approach falls short of the augmented SMARTx approach. SMARTx can integrate qualitative inputs which both of the hybrid methods obtain but cannot integrate. Furthermore, through the deliberative process SMARTx establishes the trade-offs between these and monetary criteria, which neither of the other method achieves. Lastly, the augmented SMARTx obtains individual-based qualitative inputs which may inform and guide specialists' deliberations.

10.6 Research issues

As mentioned, all the case studies were action research, which meant that we could not set up experimental situations to compare or combine different approaches. This was because of severe time and budget constraints, respondent or participant burden and the explicit requirements of the client. Therefore, many interesting questions could not be examined and remain open research questions. With a willing client, who has accepted the potential benefits of the main approach, it might well be possible to pursue this examination within an action research setting. Some of the many areas of research are mentioned below.

The proposed combination of SMARTx with a simple (non WTP) survey should be examined for use in practical environmental decision-making contexts and exposed to further practical and philosophical examination. Many of the research needs identified in the course of this thesis were not connected to the augmented model, and are probably repeated in many theses and papers. These included the need for user-friendly guides to swing weight and value function elicitation, independence checking, and value tree structuring. The latter is of particular importance in environmental decision-making, where various attributes and criteria may functionally depend on each other. The SMARTx approach lacks any formal way of dealing with uncertainty, imprecision and time, and finding reasonably simple ways of addressing these would also be a direction of research. A slightly newer research question arose from the suggestion that the indirect compensatory values found with SMARTx might be related in some way to WTP. This relationship could be explored by carrying out separate SMARTx evaluations and EE valuations on the same problem, so as to directly compare WTP results and indirect compensatory values. The study by Halvorsen *et al.* (1997) seems to be the only such study to date (see Section 9.6.2).

The augmented SMARTx approach raises some new questions. The exact mechanism by which the survey information is to inform the workshop process (development of alternatives, scoring and weighting) needs to be investigated, both practically and in the light of social choice theory. On the practical side, we need to know which information it would be most useful to gather (open-ended issues of concern, prioritisations of specific sets of attributes or alternatives) and at which stages of the workshop process this would be most useful. The proposed use of rating in these surveys also needs to be investigated more deeply to establish the relationship between these scores and true interval scales and the implications of using less-than-interval scales in weighted additions, for example. A slightly more remotely connected research question would be the relationship between the V_i of SMART and the $EXP[V_i^{*i}]$ (i.e. expected or average value of repeated estimates of V_i^{*i}) of the random utility model (Section 5.1).

10.7 Conclusion

This thesis has tried to consolidate various debates about the use of SMARTx-MCDA and EE valuation approaches to environmental decision-making and to augment these with new insights from application of the methods in real world decisions. The basis of the EE paradigm in notions of consumer sovereignty and the market make it unacceptable to many who feel these are inappropriate for environmental decisions as the environment is a public good not (or not fully) traded in the market, and that environmental decisions should be based on social or citizen values rather than individual and/or consumer values. In addition, the unreliability and bias of some of the valuation methods used in EE make their contribution as decision-aids dubious. In contrast, the basis of SMARTx is in measurement theory and therefore rests on a perhaps less controversial set of axioms. The problems with using SMARTx are mostly connected with process (although potential for bias also exists). The augmented SMARTx method proposed above does not solve all the process issues, but suggests some ways in which they could be limited. One of the strongest arguments in favour of use of SMARTx is that it *informs* the decision process and assists in the separation of facts from values, in contrast to EE which is complex and seems to confound positive and normative issues.

We have shown that the use of WTP may introduce several biases in evaluation, these range from its theoretical underpinnings in consumer sovereignty and the efficiency of the market (whereas environment quality is a citizen issue and aspects of it are not traded on the market), through to its potential pro-rich biases. Monetary values or prices used to be considered a proxy for utility, but have now become an end in themselves. However, the techniques used mean that the values obtained will often have variances so large as to make interpretation meaningless. Many have agreed that it is difficult to find monetary values for certain project impacts (e.g. Pearce 1983, Gregory *et al.* 1992, Munasinghe 1993, Vatn and Bromley 1994, van Pelt 1993, Munda *et al.* 1994a, Angelsen and Sumaila 1995) and it may seem unnecessary to belabour the point. Some have suggested that these impacts may be considered using MCDA (e.g. Gregory *et al.* 1992, van Pelt 1993, Munasinghe 1993, Munda *et al.* 1994a,b), but many remain sceptical. The discussions above and in Chapter 9 (summarised in Table 9.2) suggest that projects with environmental and social impacts may be better handled within a SMARTx rather than EE *framework*, and possibly that some derived monetary values (i.e. the indirect compensatory values) may better reflect trade-off values than their EE counterparts.

The theoretical and practical consideration of equity and sustainability within EE has proved problematic. While not capable of solving all the problems associated with operationalising these concepts, SMARTx offers some advantages (e.g. Gregory *et al.* 1992, van Pelt 1993, Munasinghe 1993, Munda *et al.* 1994a, Faucheux and Froger 1995), in part through the use of a hierarchical system of objectives, goals and criteria selected and developed by specialists and stakeholders, where lower level criteria contribute to the broader goals of equity, sustainability and efficiency. Only the lower level criteria need to be measurable, whether quantitatively or qualitatively. While EE may be appropriately used to determine certain growth and efficiency considerations, intangible and qualitative data can more easily be included in SMARTx and effects on particular groups more explicitly assessed. While SMARTx does not directly incorporate uncertainty, its flexibility allows the use of a variety of approaches which can augment its handling of uncertainty. Thus a narrow and exclusive approach to project appraisal which necessarily limits creativity without ensuring rigour is avoided.

Rather than continuing to apply the same techniques with increasing analytical sophistication, but the same fundamental flaws, these arguments suggest a new approach is needed which combines the strengths of both EE and MCDA and minimises the weaknesses, and therefore initial recommendations for an augmented SMARTx-based approach which is informed by EE concepts were made above.

While the discussions and recommendations apply to any project appraisal, they are more persuasive in developing countries, where impacts are often felt by those to whom any small change in income or quality of life is highly significant, due either to a low income base or to direct dependence on the environment. They are especially relevant to projects with environmental impacts as these will not be directly revealed in the market, and cannot appropriately be measured in hypothetical markets.

"In response to your question 'what is worth doing and what is worth having?' I would like to say simply this. It is worth doing nothing and having a rest." Michael Leunig 2001, p.26.

References

- Abu-Taleb MF and Mareschal B, 1995. Water resources planning in the Middle East: Application of the PROMETHEE V multicriteria method. *European Journal of Operational Research*, 81:500-511.
- Adelman L, 1991. Experiments, quasi-experiments, and case studies: A review of empirical methods for evaluating decision support systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(1):293-301
- Angelsen A, and Sumaila UR, 1995. Hard methods for soft policies: Environmental and social cost-benefit analysis. Michelsen Institute for Development Studies and Human Rights, Bergen. WP 1, 1995.
- Argyrous G, 1994. Kuhn's paradigms and neoclassical economics: Reply to Dow. *Economics and Philosophy*, 10:123-126.
- Arrow KJ, 1950. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58:328-346.
- Arrow KJ, 1963. *Social Choice and Individual Values*. Yale University Press, New Haven.
- Arrow KJ and Reynaud H, 1986. Social choice and multicriterion decision-making. Cambridge, MA: MIT Press.
- Arrow KJ, Solow R, Portney PR, Learner EE, Radnar R and Schuman, 1993. Report of the National Oceanic and Atmosphere Administration (NOAA) panel on contingent valuation. Washington DC, *Federal Register*, 58:4601-4614.
- Aumann RJ, 1964. Subjective Programming. In: (eds) Shelby MW and Bryan GL, *Human Judgements and Optimality*, Chapter 12. Wiley and Sons, New York.
- Bakus GJ, Stilwell WG, Latter SM and Wallerstein MC, 1982. Decision making: With applications for environmental management. *Environmental Management*, 6(6):493-504.
- Bana e Costa CA, 1990. (ed) *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin. 660 p.
- Bana e Costa CA and Pirlot M, 1997. Thought on the future of the multicriteria field: Basic convictions and outline for general methodology. In (ed) Climaco J, *Multicriteria Analysis*. Springer-Verlag, Berlin.
- Bana e Costa CA and Vansnick J-C, 1994. A theoretical framework for Measuring Attractiveness by a Categorical Based Evaluation Technique (MACBETH). *Proceedings of the XIth International Conference on MCDA*, Coimbra, Portugal, August, 1994.
- Bana e Costa CA, Stewart TJ and Vansnick J, 1995. Multicriteria decision analysis: Some thoughts based on the tutorial and discussion sessions of the ESIGMA meetings. In: *Proceedings of the 14th European Conference on Operational Research*, Jerusalem, Israel. July 3-6.
- Barzilai J and Lootsma FA, 1997. Power relations and group aggregation in the multiplicative AHP and SMART. *Journal of Multi-Criteria Decision Analysis*, 6(3):155-165.
- Baumol WJ, 1977a. Neumann-Morgenstern utility theory. In: *Economic Theory and Operations Analysis*. 4th edition. Prentice Hall, New Jersey.
- Baumol WJ, 1977b. General equilibrium and welfare economics. In: *Economic Theory and Operations Analysis*, Prentice-Hall, New Jersey.
- Beck MP and Lin BW, 1983. Some heuristics for the consensus ranking problem. *Computers and Operations Research*, 10(1):1-7.
- Becker GS, 1974. A theory of social interactions. *Journal of Political Economy*, 82:1063-1094.
- Beinat E, 1995. *Multiattribute Value Functions for Environmental Management*. PhD thesis, Free University of Amsterdam. No. 103 of Tinbergen Institute Research Series.
- Beinat E and Nijkamp P, 1998. (eds) *Multicriteria Analysis for Land-use Management*. Kluwer Academic Publishers, Dordrecht.
- Bell DE, Keeney RL and Raiffa H (eds), 1977. *Conflicting Objectives in Decisions*. International Series on Applied Systems Analysis, 1. International Institute for Applied Systems Analysis- John Wiley and Sons, Chichester.
- Bell ML, Hobbs BF, Elliott EM, Ellis H, 2001. An evaluation of multi-criteria methods in integrated assessment of climate policy. *Journal of Multi-Criteria Decision Analysis*, 10:229-256.
- Belton V and Pictet J, 1997. A framework for group decision using a MCDA model: Sharing, aggregating or comparing individual information?. *Revue des systèmes de décisions*, 6(3):283-303.
- Belton V and Stewart TJ, 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Belton V and Vickers S, 1990. Use of a simple multi-attribute value function incorporating visual interactive sensitivity analysis for multiple criteria decision making. In (ed) Bana e Costa C. *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin.
- Bentham J. 1789. *An Introduction to the Principles of Morals and Legislation*. In: (eds) Burns JH and Hart HLA. (Intro by Rosen F). 1996. *The collected works of Jeremy Bentham*. Clarendon Press, Oxford.
- Black RDC, 1970. (ed) Introduction to: Jevons WS, 1871. *The Theory of Political Economy*. Penguin Books Ltd.
- Blamey RK, Common M and Quiggin J, 1995. Respondents to contingent valuation surveys: Consumers or citizens? *Australian Journal of Agricultural Economics*, 39(3):263-288.

- Bockstael NE, Freeman AM iii, Portney PR, and Smith VK, 2000. On measuring economic values for nature. *Environmental Science and Technology*, 34:1384-1389.
- Bodily SE, 1979. A delegation process for combining individual utility functions. *Management Science*, 25(10).
- Bogetoft P and Pruzan P, 1991. *Planning with Multiple Criteria*. North Holland Elsevier Science Amsterdam.
- Boxall PC, Adamowicz WL, Swait J, Williams M and Louviere J, 1996. A comparison of stated preference methods for environmental evaluation. *Ecological Economics*, 18:243-253.
- Braden JB, 2000. Value of valuation: Introduction. *Journal of Water Resources Planning and Management*, 126(6): 336-338.
- Brans JP and Mareschal B, 1990. The PROMETHEE methods for MCDM; the PROMCALC, GAIA and BankAdviser software. In: (ed) Bana e Costa C, *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin. 660 p.
- Brock HW, 1980. The problem of utility weights in group preference aggregation. *Operations Research*, 28(1): 176-187.
- Brown RS, unpubl. Stated preference valuation methods: Surrogate markets or substitutes for markets? Dept. of Economics, University of Queensland, Australia.
- Brown S, Schreier H, Thompson W A, and Vertinsky I, 1994. Linking Multiple accounts with GIS as decision support system to resolve forestry / wildlife conflicts. *Journal of Environmental Management*, 42(4):349-364.
- Brown TC, Peterson GL and Tonn BE, 1995. The values jury to aid natural resource decisions (speculations). *Land Economics*, 71(2):250-260.
- Brundtland GH, 1987. 'The Brundtland Report' or 'Our Common Future'. Report to the United Nations by the World Commission on Environment and Development. Oxford University Press, 1987, p.43 (Chair – Gro Harlem Brundtland).
- Cameron TA, 1988. A new paradigm for valuing non-market goods using referendum data. Maximum likelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management*, 15(3):355-379.
- Cameron T, 1992. Combining contingent valuation and travel cost data for the valuation of nonmarket goods, *Land Economics*, 68(3):302-317.
- Carson RT, 2000. Contingent valuation: A user's guide. *Environmental Science and Technology*, 34(8):1413-1418.
- Carson RT, Hanemann WM, Kopp RJ, Krosnick JA, Mitchell RC, Presser S, Ruud PA and Smith VK, 1994. Prospective interim lost use value due to DDT and PCB contamination in the southern Californian bight. Report of the Natural Resource Damage Assessment Inc. and Industrial Economics Inc. to the National Oceanic and Atmospheric Administration, NOAA contract No. 50-DGNC-!-00007.
- Central Economic Advisory Services, 1989. *Cost-Benefit Analysis in South Africa*.
- Chang Ni-Bin, Wen CG, Wu SL, 1995. Optimal management of environmental and land resources in a reservoir watershed by multiobjective programming. *Journal of Environmental Management*, 44(2): 145-161.
- Charnes A and Cooper WW, 1961. *Management models and industrial applications of linear programming*, Vol 1, Wiley, New York.
- Churchman CW, 1972. Morality as a value criterion. Paper read at 1st Multiple Criteria Decision Making meeting at University of South Carolina.
- Clark M. and Herington J, 1988. (eds.) *The Role of ELA in the Planning Process*. Mansell Publishing Ltd.
- Cocks D and Ive J, 1996. Mediation support for forest land allocation: The SIRO-MED system. *Environmental Management*, 20(1):41-52.
- Cohen D, 2001. Analysis of conjoint data - a review of a case study "Estimating potential impacts of a change in river quality on the tourism value of Kruger National Park: An application of travel cost, contingent and conjoint valuation methods". Honours Project (supervisors TJ Stewart and AR Joubert), Dept. Statistical Sciences, University of Cape Town.
- Cohon JL, 1978. *Multiobjective Programming and Planning*. Academic Press, New York.
- Cook WD and Kress M, 1985. Ordinal rankings with intensity of preference. *Management Science*, 31(1): 191-198.
- Cook WD and Kress M, 1991. A multiple criteria decision model with ordinal preference data. *European Journal of Operational Research*, 54:191-198.
- Cornelissen AMG, van den Berg J, Koops WJ, Grossman M, Udo HMJ, 2001. Assessment of the contribution of sustainability indicators to sustainable development: A novel approach using fuzzy set theory. *Agriculture, Ecosystems and Environment* 86:173-185.
- Cowen R, 1999. Notes for Geology 115 course at University of California Davis. <http://www-geology.ucdavis.edu/~GEL115/115CH17oldirrigation.html>.
- Cowen T, 1993. The scope and limits of preference sovereignty. *Economics and Philosophy*, 9: 253-269.
- Daily GC and Ehrlich PR, 1995. Population extinction and the biodiversity crisis. In: (eds) Perrings CA, Maler KG, Folke C, Holling CS, Jansson BO, *Biodiversity conservation: Problems and policies*.
- Daly HE, 1976. The steady-state economy. In: (ed) Wilson KD, *Prospects for growth*. Praeger, New York.
- Dantzig GB, 1951. Maximization of a linear function of variables subject to linear inequalities. In: (ed) Koopmans TC, *Activity Analysis of Production and Allocation*. Wiley, New York.

- Davies P, Williams AT and Curr RHF, 1995. Decision making in dune management: Theory and practice. *Journal of Coastal Conservation*, 1(1):87-96.
- Dee N, Baker J, Drober N and Duke K, 1973. An environmental evaluation system for water resources planning. *Water Resources Research* 9(3):523-535.
- de Graaff J, and Kuyvenhoven A, 1996. Economic evaluation of watershed development. Inaugural conference of the European branch of the International Society for Ecological Economics, 23-25 May, 1996. Paris.
- de Montis A, de Toro P, Droste-Franke B, Omann I and Stagl S. Undated. MCDA and sustainable development – A comparison of methods. www.kfunigraz.ac.at/vwlwww/omann/finalversion.pdf.
- Decision Explorer V3.04. 1996. © University of Strathclyde and Mathew Jones, (93-96), Published by Banxia Software Ltd. Glasgow, UK.
- Dept. of Environmental Affairs, South Africa, 1992. Guidelines for report requirements. Guidelines Document 3 of Integrated Environmental Management Guideline Series.
- Diamond P and Hausman J, 1994. Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives*, 8(4):45-64.
- Dixon JA, Scura LF, Carpenter RA and Sherman PB, 1986. *Economic Analysis of Environmental Impacts*. Earthscan Publications Ltd. London.
- Dodgson J, Spackmann M, Pearman A and Phillips L, 2001. *Multi-criteria analysis: A manual*. Department for Environment, Food and Rural Affairs, United Kingdom. A report prepared by National Economic Research Associates.
- Dow SC, 1994. Kuhn's paradigms and neoclassical economics: A comment. *Economics and Philosophy*, 10:119-122.
- Dubourg WR and Pearce DW, 1994. Paradigms for environmental choice: Sustainability versus optimality. In: *Proceedings of the International symposium: Models of sustainable development. Exclusive or complementary approaches of sustainability?*
- Dupuit J, 1844. In *Annale des Ponts et Chaussées*, translated: *International Economic Papers*, 1952, No.2.
- Eberhard R and Joubert AR, 2000. An evaluation of alternative water supply augmentation and water demand management options for the City of Cape Town using multi-criteria decision analysis. A report to the City of Cape Town.
- Eberhard R and Joubert AR, 2001. An evaluation of alternative water supply augmentation and water demand management options for the City of Cape Town using multi-criteria decision analysis. Phase 2: Assessment of three additional supply augmentation options. A report to the City of Cape Town.
- Edwards W, 1954. The theory of decision making. *Psychological Bulletin*, 51(4):380-417.
- Edwards W, 1971. Social utilities. *Engineering Economist*, Summer Symposium Series, 6.
- Edwards W, 1977. How to use multiattribute utility measurement for social decision making. *IEEE Transactions on Systems, Man and Cybernetics*, SCM-7: 326-340.
- Edwards W, 1992. (ed) *Utility Theories: Measurements and Applications*. Studies in Risk and Uncertainty. Kluwer Academic Publishers, Netherlands.
- Edwards W and Barron FH, 1994. SMARTS and SMARTER: improved simple methods for multiattribute utility measurement. *Organizational Behavior and Human Decision Processes*, 60:306-325.
- Ellsberg D, 1954. Classic and current notions of 'measurable utility'. *The Economic Journal*, 64: 528-556
- Ezzell C, 2001. The Himba and the Dam. *Scientific American*, June 2001.
- Faith DP and Walker PA, 1996. Integrating conservation and development: Effective trade-offs between biodiversity and cost in the selection of protected areas. *Biodiversity and Conservation* 5(4):431-446.
- FAO, Food and Agriculture Organisation, 2000. Applications of the contingent valuation method in developing countries - A survey. FAO Economic and Social Development Paper No. 146.
- Farber S and Griner B, 2000. Using conjoint analysis to value ecosystem change. *Environmental Science and Technology*, 34(8):1407-1412
- Faucheux S and Froger G, 1995. Decision-making under environmental uncertainty. *Ecological Economics*, 15(1): 29-42.
- Fischhoff B, 1977. Cost benefit analysis and the art of motorcycle maintenance. *Policy Sciences*, 8:177-202.
- Fischhoff B, 2000. Informed consent for eliciting environmental values. *Environmental Science and Technology*, 34:1439-1444.
- Fischhoff B, Lichtenstein S, Slovic P, Derby SL and Keeney RL, 1981. *Acceptable Risk*. Cambridge University Press, Cambridge.
- Fishburn PC, 1976. Cardinal utility: An interpretive essay. *Revista di Scienze Economiche e commerciali*, 12:1102-1114.
- Freeman AM III, 1993. *The Measurement of Environmental and Resource Values: Theory and Methods*. Resources for the Future, Washington DC.
- French S, 1988. *Decision Theory: An Introduction to the Mathematics of Rationality*. Ellis Horwood Ltd. Chichester.

- French S 1989. *Readings in Decision Analysis*. Chapman and Hall, London.
- Fuggle RF and Rabie MA, 1992. *Environmental Management in South Africa*. Juta.
- Garson GD, 2002. *Guide to Writing Empirical Papers, Theses, and Dissertations*. Marcel Dekker, Inc. New York.
- Gibb 1997. Impact assessment phase of the environmental impact assessment on a proposed steel plant and transport options. Report by Gibb Africa, project No J19795A.
- Goicoechea A, Hansen DR and Duckstein L, 1982. *Multiobjective Decision Analysis with Engineering and Business Applications*. John Wiley and Sons, New York.
- Goicoechea A, Stakhiv EZ and Li F, 1992. Experiment evaluation of multiple criteria decision models for application to water resource planning. *Water Resources Bulletin*, 28(1):1-14.
- Goodwin P and Wright G, 1998. *Decision analysis for management judgement*. John Wiley & Sons, Chichester, England.
- Graaff J de van., 1971. *Theoretical Welfare Economics*. Cambridge University Press, Cambridge.
- Green PE and Krieger AM, 1993. Conjoint analysis with product-positioning applications. In: (eds) Eliashberg J and Lilien GL. *Handbooks in Operations Research & Management Science: Marketing*, 5:463-515.
- Gregory R, Keeney R and von Winterfeldt D, 1992. Adapting the Environmental Impact Statement process to inform decisionmakers. *Journal of Policy Analysis and Management*, 11(1): 58-75.
- Haeussler EF and Paul RS, 1996. *Introductory mathematical analysis for business, economics and the life and social sciences*. 8th edition. Prentice Hall International Editions. New Jersey. 941 pp.
- Halvorsen B, Strand J, Sælensminde K, Wenstøp F, 1997. Comparing contingent valuation, conjoint analysis and decision panels: An application to the valuation of reduced damages from air pollution in Norway. In: (eds) Stewart TJ and van den Honert RC. *Lecture Notes in Economics and Mathematical Systems*, 465; *Proceedings of the 13th International Conference on Multiple Criteria Decision Making*, South Africa, January, 1997.
- Hanemann WM, 1989. Welfare evaluations in contingent valuation experiments with discrete response data: Reply. *American Journal of Agricultural Economics*, 71(4):1057-1061.
- Hanemann WM and Kanninen B, 1996. *The Statistical Analysis of Discrete-response CV data*. Dept. of Agricultural and Resource Economics, University of California at Berkeley. Working Paper No. 798.
- Hanley N, 1987. Valuing non-market goods using contingent valuation: A survey and a synthesis. University of Sterling discussion papers in economics, finance and investment, No.138.
- Hanley N and Spash C, 1993. *Cost-Benefit Analysis and the Environment*. Edward Elgar, Aldershot. 385 p.
- Hanley N, Wright R, Macmillan D and Philip L, 2001. *Willingness to pay for the conservation and management of wild geese in Scotland*. Technical Report B. Scottish Executive, Central Research Unit.
- Hargreaves Heap S, Hollis M, Lyons B, Sugden R and Weale A, 1992. *The Theory of Choice: A Critical Guide*. Blackwell Publishers, Oxford.
- Harrison NE, 2000. *Constructing Sustainable Development*. State University of New York Press.
- Harsanyi JC, 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63:309-321.
- Hausman DM, 2000. Revealed preference, belief, and game theory. *Economics and Philosophy*. 16:99-115.
- Henderson J and Quandt R, 1980. *Microeconomic theory: A mathematical approach*. 3rd edn. McGraw-Hill.
- HET- history of economic thought website. New School University, Department of Economics. <http://cepa.newschool.edu/het/essays/paretian/paretosocial.htm>). The Paretian System – IV – Social Welfare.
- Hicks JR, 1939. The foundations of welfare economics. *The Economic Journal*, 49:696-712.
- Hicks JR, 1943. The four consumer's surpluses. *Review of Economic Studies*, 11(1): 31-41.
- Hicks JR and Allen RGD, 1934. A reconsideration of the theory of value. *Economica*, 14:52-76, 196-219.
- Hobbs BF, 1979. Analytical multiobjective decision methods for power plant siting: A review of theory and applications. Division of Regional Studies, Brookhaven National Laboratory, New York.
- Hobbs BF, 1985. Experiments in multicriteria decision making and what we can learn from them: An example. In: (eds) Haimes YY and Chankong V. *Decision Making with Multiple Objectives*, Springer-Verlag, Berlin.
- Hobbs BF, Chankong V, Hamadeh W, and Stakhiv EZ, 1992. Does choice of multicriteria method matter? An experiment in water resources planning. *Water Resources Research*, 28(7):1767-1779.
- Howard RA, 1966. Decision analysis: Applied decision theory. In: (eds) Hertz DB and Melese J. *Proceedings of the fourth International Conference on Operational Research*. Wiley-Interscience, New York.
- Hubin DC, 1994. The moral justification of benefit/cost analysis. *Economics and Philosophy*, 10: 169-194.
- Hwang CL and Yoon K, 1981. *Multi-Attribute Decision Making*. Springer-Verlag, New York.
- International Rivers Network, 2001. *Epupa Currents*. <http://www.irn.org/programs/epupa/index.shtml>
- Jaquet-Lagrèze E, 1990. Interactive assessment of preferences using holistic judgements: the PREFCALC system. In: (ed) Bana e Costa C. *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin.
- Jakobsson KM and Dragun AK, 1996. *Contingent Valuation and Endangered Species: Methodological Issues and Applications*. Edward Elgar, Cheltenham UK.

- Jayne TS, Rubey L, Lupi D Tschirely D and Weber M, 1996. Estimating consumer response to food market reform using stated preference data: Evidence from eastern and southern Africa. *American Journal of Agricultural Economics*, 78:820-824.
- Jevons WS, 1871. *The Theory of Political Economy*. (ed) Black RDC. 1970. Penguin Books Ltd.
- Johansson-Stenman O, 2002. Environmental policy when people's preferences are inconsistent, non-welfaristic, or simply not developed. Forthcoming in: (eds) Bromley D and Paavola J *Economics, Ethics and Environmental Policy: Contested Choices*. Blackwell.
- Joubert AR, 1997. Application of multi-criteria decision analysis to the environmental impact assessment for the Maputo iron and steel project transprot options. Report to Gibb Africa for Maputo Iron and Steel Impact Assessment.
- Joubert AR and Pollard S, 2000. Using multi-criteria decision analysis in catchment management: A case study of the Sand River, Mpumalanga, South Africa. *African Journal of Aquatic Science*, 25: 238-242.
- Joubert A R, Leiman A, de Klerk H, Katua S, Aggenbach J-C, 1997. Fynbos (fine bush) vegetation and the supply of water: A comparison of multi-criteria decision analysis and cost-benefit analysis. *Ecological Economics*, 22: 123-140.
- Kahneman D and Knetsch JL, 1992. Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22(1): 57-70.
- Kaldor N, 1939. Welfare propositions and inter-personal comparisons of utility. *The Economic Journal*, 49: 549-552.
- Keeney RL, 1987. An analysis of the portfolio of sites to characterize for selecting a nuclear repository. *Risk Analysis*, 7: 195--218.
- Keeney RL, 1992. On the foundations of prescriptive decision analysis. In: (ed) Edwards W, *Utility Theories: Measurement and Applications. Studies in Risk and Uncertainty*. Kluwer Academic Publishers.
- Keeney RL, 1996. Value focused thinking: Identifying decision opportunities and creating alternatives. *European Journal of Operational Research*, 92:537-549.
- Keeney RL and Nair K, 1976. Evaluating potential nuclear power plant sites in the Pacific Northwest using decision analysis. Reference in: (eds) Bell DE, Keeney RL and Raiffa H. *Conflicting Objectives in Decisions. International Series on Applied Systems Analysis, 1*. International Institute for Applied Systems Analysis- John Wiley and Sons, Chichester.
- Keeney RL and Raiffa H, 1976. *Decisions with multiple objectives: Preferences and value trade-offs*. John Wiley and Sons, New York.
- Kenyon W and Hanley N, 2000. Economic and participatory approaches to environmental evaluation. Unpublished manuscript. Scottish Agricultural College, Edinburgh.
- Kersten G, undated. Support for group decisions and negotiations: An overview. http://www.iiasa.ac.at/Research/DAS/interneg/research/mise/intro_gdn.htm
- Kontoleon A, Macrory R and Swanson T, 2001. Individual preferences, expert opinion and environmental decision making: An overview of the issues. Paper prepared for the Symposium on Law and Economics of Environmental Policy. [www://cserge.ucl.ac.uk/kontoleon_et_al.pdf](http://www.cserge.ucl.ac.uk/kontoleon_et_al.pdf).
- Krzsztowicz R, 1979. Group utility assessment through a nominal-interacting process. *Unpubl. Working paper*. Dept. of civil Engineering, MIT, Cambridge, Massachusetts, 1979.
- Kuhn TS, 1962. *The Structure of Scientific Revolutions*, University of Chicago Press.
- Lamberth S and Joubert AR, *in review*. Prioritizing species for research, conservation and management: A case study of exploited fish species. *Conservation Biology*.
- Larichev OI and Moshkovich HM, 1995. ZAPROS-LM - A method and system for ordering multiattribute alternatives. *European Journal of Operational Research*, 82:503-521.
- Lawes Agricultural Trust, 1998. GENSTAT 5 Release 4.1. IACR Rothamstead.
- Leunig M, 2001. *The Curly Pyjama letters*. Viking, Penguin Books Australia.
- Lewis HS and Butler TW, 1993. An interactive framework for multi-person, multi-objective decisions. *Decision Sciences*, 24:1-22.
- LINDO Systems Inc., © 2000. HyperLindo/PC, Release 6.1 LINDO Systems Inc. 1415 North Dayton St. Chicago, IL, 60622. <http://www.lindo.com>.
- Lipsey RG and Lancaster K, 1956. The general theory of second best. *Review of Economic Studies*, 24:11-32.
- Little IMD, 1957. *A Critique of Welfare Economics*. Oxford University Press, Oxford.
- Loomes G and Sugden R, 1982. Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92:805-824.
- Loomis JB, 1995. Four models for determining environmental quality effects on recreational demand and regional economics. *Ecological Economics*, 12:55-65.
- Loomis JB, 2000. Environmental valuation techniques in water resource decision making. *Journal of Water Resources Planning and Management*, 126(6): 339-344.

- Lootsma FA, 1989. Conflict resolution via pairwise comparison of concessions. *European Journal of Operational Research*, 40:109-116.
- Lootsma FA, 1993. Scale sensitivity in the multiplicative AHP and SMART. *Journal of Multi-criteria Decision Analysis*, 2:87-110.
- Lootsma RA, 1996. Comment on: The European school of MCDA: Emergence, basic features and current works. (by Roy B and Venderpooten D). *Journal of Multi-Criteria Decision Analysis*, 5:22-38
- Lootsma FA, 1997. *Fuzzy logic for Planning and Decision Making*. Applied Optimization, Volume 8. Kluwer Academic Publishers, Dordrecht.
- Losa F, van den Honert R and Joubert AR, 2001. The multivariate analysis biplot as a tool for conflict analysis in MCDA. *Journal of Multi Criteria Decision Analysis*, 10(5):273-284.
- Maass A, Hufschmidt MM, Dorfman R, Thomas Jr HA, Marglin SA and Fair GM, 1962. *Design of Water-Resource Systems*. Harvard University Press, Cambridge, USA.
- Mackenzie J, 1993. A comparison of contingent preference models. *American Journal of Agricultural Economics*, 75:593-603.
- Maguire LA and Boiney LG, 1994. Resolving environmental disputes: A framework incorporating decision analysis and dispute resolution techniques. *Journal of Environmental Management*, 42(1):31-48.
- Mail and Guardian 2002. What is sustainable development? *Mail and Guardian* 18(11), March 22-27, 2002.
- Malczewski J, 1999. *GIS and Multicriteria Decision Analysis*. John Wiley & Sons, Inc, New York.
- Mander M, Cox D, Breen C and Turpie J, 2001. Incorporating economic considerations into the allocation, quantification and management of the environmental reserve: Synthesis report: executive summary, lessons learned and recommendations. Working paper No. 7 of Water Research Commission Project No. K5/978.
- Martunnen M and Hämäläinen RP, 1995. Decision analysis in environmental impact assessment. *European Journal of Operational Research*, 87:551-563.
- Massam BH, 1999. The classification of quality of life using multi-criteria analysis. *Journal of Geographic Information and Decision Analysis*, 3(2):1-8.
- McCarthy JJ, Canziani OF, Leary NA, Dokken DJ and White KS, 2001. Climate change 2001: Impacts, adaptation and vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change. <http://www.ipcc.ch/pub/tar/wg2/112.htm>
- McConnell KE, 1986. The damages to recreational activities from PCBs in New Bedford harbour. Report prepared by Industrial Economics Inc. Mass. For the Ocean Assessment Division, National Oceanic and Atmospheric Administration.
- McDaniels TL, 1995. Using judgement in resource management: A multiple objective analysis of a fisheries management decision. *Operations Research*, 43(3):415-426.
- McFadden D, 1976. Quantal choice analysis: A survey. *Annals of Economic and Social Measurement*, 5: 363-390.
- McFadden D, 1999. Rationality for economists? *Journal of Risk and Uncertainty*, 19, 73-105.
- Merkhofer MW and Keeney RL, 1987. A multiattribute utility analysis of alternative sites for the disposal of nuclear waste. *Risk Analysis*, 7: 173--194.
- Microsoft 1995. Microsoft Corporation © (1985-1997).
- Mill JS, 1848, Principles of Political Economy, Book 3, ch1. in Black, 1970 p. 11.
- Miller GA, 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63: 81-97.
- Mingers J and Brocklesby J, 1997. Multimethodology: Towards a framework for mixing methodologies. *Omega, International Journal of Management Science*, 25(5): 489-509.
- Mishan EJ. 1972. *Cost-Benefit Analysis*. Praeger, New York.
- Mitchell R and Carson R, 1989. *Using surveys to value public goods*. Resources for the future, Washington DC.
- Morgan MG and Henrion M, 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, New York.
- Munasinghe M, 1993. Environmental economics and sustainable development. World Bank Environmental paper number 3. The World Bank, Washington DC.
- Munda G, Nijkamp P and Rietveld P, 1994a. Monetary and non-monetary evaluation methods in sustainable development planning. In: *Proceedings of the International symposium: Models of sustainable development. Exclusive or complementary approaches of sustainability?*, 2:887-898.
- Munda G, Nijkamp P and Rietveld P, 1994b. Qualitative multicriteria evaluation for environmental management. *Ecological Economics*, 10: 97-112.
- Niewijk RK, undated. Cato Review of Business & Government (<http://www.cato.org/pubs/regulation/reg17n1-niewijk.html>)
- Olson DL, Fliedner G and Currie K, 1995. Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research*, 82:522-539.

- Ortolano L, 1974. A process for federal water planning. *Water Resources Bulletin*, 10(4).
- Ozernoy VM, 1997. Some fundamental problems in the selection and justification of discrete alternative MCDM methods. In: (ed) Clímaco J. *Multicriteria Analysis. Proceedings of the XIth International Conference on MCDM*, 1994, Portugal.
- Parsons W, 1995. *Public Policy: An introduction to the theory and practice of policy analysis*. Edward Elgar Aldershot, UK.
- Pearce DW, 1983. *Cost-Benefit analysis* (2nd edition). Macmillan, London.
- Pearce DW and Turner RK, 1990. *Economics of Natural Resources and the Environment*. Harvester Wheatsheaf, London.
- Pearce DW and Seccombe-Hett T, 2000. Economic valuation and environmental decision-making in Europe. *Environmental Science and Technology*, 34:1419-1425.
- Perman R, Ma Y and McGilvray, 1996. *Natural resource and environmental economics*. Longman. London.
- Perny P, 1998. Multicriteria filtering methods based on concordance and non-discordance principles. *Annals of Operations Research*, 80:137-165.
- Pollard S, Perez de Mendiguren J-C, Joubert A, Shackleton C, Walker P, Poulter T, White M, 1998. Save the Sand Phase I Feasibility Study: The development of a proposal for a catchment plan for the Sand River Catchment. Submitted by AWARD to the Department of Water Affairs and Forestry. 280 pp.
- Prato T, 2000. Multiple attribute evaluation of landscape management. *Journal of Environmental Management* 60: 325-337.
- Prelec D and Loewenstein G, 1991. Decision making over time and under uncertainty: A common approach. *Management Science*, 37(7):770-786.
- Raiffa H, 1968. *Decision Analysis*. Addison-Wesley, Reading, USA.
- Raiffa H and Schlaifer R, 1961. *Applied Statistical Decision Theory*. Belknap Press/Harvard University Press, Cambridge.
- Ralls K and Starfield A, 1995. Choosing a management strategy: Two structured decision-making models for evaluating the predictions of stochastic simulation models. *Conservation Biology*, 9(1):175-181.
- Ramsey FP, 1931. Truth and probability. In: *The Foundations of Mathematics and Other Logical Essays*. Harcourt Brace, New York.
- Rauschmayer F, 2001. Reflections on ethics and MCA in environmental decisions. *Journal of Multi-Criteria Decision Analysis*, 10(2):65-74.
- Republic of South Africa, 1998. National Water Act, Act No. 36 of 1998.
- Republic of South Africa, 1998. National Forests Act, Act No. 84 of 1998.
- Republic of South Africa, 1997. Water Services Act, Act No. 108 of 1997.
- Republic of South Africa, 1998. National Environmental Management Act, Act No. 107, 1998.
- Ridgley MA, 1995. A multicriterion approach to setting instream flow standards. *Journal of Environmental Systems*, 24(1):69-86.
- Ridgley MA and Rijsberman F, 1992. Multicriteria evaluation in a policy analysis of a Rhine estuary. *Water Resources Bulletin*, 28(6):1095-1110.
- Ridgley MA and Rijsberman F, 1994. Multicriterion analysis and the evaluation of restoration policies for a Rhine estuary. *Socio-Economic Planning Sciences*, 28(1):19-32.
- Roberts FS, 1979. *Measurement Theory with Applications to Decisionmaking, Utility and the Social Sciences*. Encyclopedia of Mathematics and its Applications, Vol 7. Addison-Wesley Publishing Company, Reading.
- Roe B, Boyle K and Teisl M, 1996. Using conjoint analysis to derive estimates of compensating variation. *Journal of Environmental and Economic Management*, 31:145-159.
- Rogers P, 1999. Environmental indicators. *Journal of Water Resources Planning and Management*, 125(2): 74-75.
- Rosenhead J, 1989. *Rational analysis in a problematic world*. John Wiley and Sons, New York, USA. 386 pp.
- Ross D, 1999. *What people want. The concept of utility: From Bentham to game theory*. UCT Press.
- Rouse WG and Sheridan TB, 1974. Computer-aided group decision making: Theory and practice. *Proceedings of the International Conferences on Systems, Man and Cybernetics*, IEEE Systems, Man and Cybernetics Society, Dallas. 505-513.
- Rowe MD and Pierce BL, 1982. Sensitivity of the weighting summation decision method to incorrect application. *Socio-Economic Planning Sciences*, 16(2):173-177.
- Roy B, 1990. The outranking approach and the foundation of ELECTRE methods. In: (ed) Bana e Costa, *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin. 660 p.
- Saaty TL, 1980. *The analytic hierarchy process*. McGraw-Hill, New York.
- Saaty TL and Vargas LG, 1987. Uncertainty and rank order in the analytic hierarchy process. *European Journal of Operational Research*, 32: 107-117.

- Salo AA and Hämäläinen RP, 1992. Preference assessment by imprecise ratio statements. *Operations Research*, 40(6): 1053-1061.
- Salo AA and Hämäläinen RP, 1995. Preference programming through approximate ratio comparisons. *European Journal of Operational Research*, 82:458-475.
- Samuelson PA, 1938. The numerical representations of ordered classifications and the concept of utility. *Revue of Economic Studies*, 6:65-70.
- Samuelson PA, 1956. Social indifference curves. *Quarterly Journal of Economics*, 90:1-22.
- Samuelson PA and Nordhaus WD, 1985. *Economics*. McGraw-Hill Book Company, Singapore.
- Sarin RK, 1990. Conjoint measurement: Theory and methods. In: (ed) Bana e Costa C. *Readings in Multiple Criteria Decision Aid*, Springer Verlag, Berlin.
- Savage LJ, 1954. *The Foundations of Statistics*. Wiley, New York.
- Schlaifer R, 1959. *Probability and Statistics for Business Decisions: An Introduction to Managerial Economics under Uncertainty*. McGraw-Hill, New York.
- Schulze PC, 1994. Cost-benefit analyses and environmental policies. *Ecological Economics*, 9: 197-199.
- Scitovsky T, 1941. A note on welfare propositions in Economics. *Review of Economic Studies*, 9(1): 77-88.
- Scott D and Suppes P, 1958. Foundational aspects of theories of measurement. *Journal of Symbolic Logic*, 23:113-128.
- Self P, 1970. Nonsense on stilts: The futility of Roskill. *New Society*, 2 July.
- Sen AK, 1970. The impossibility of a Paretian liberal. *Journal of Political Economy*, 78 (1):152-157.
- Sen AK, 1974. Choice, orderings and morality. In (ed) Körner S. *Practical Reason. Proceedings of the First Bristol Conference on Critical Philosophy*. Blackwell, Oxford.
- Sen AK, 1977. Rational fools: A critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs*, 6:371-344.
- Sen AK, 1979. Personal utilities and public judgements: Or whats wrong with welfare economics?. *The Economic Journal*, 89(September): 537-558.
- Sen AK, 1997. Individual preference as the basis of social choice. In: (eds) Arrow KJ, Sen AK and Suzumura K: *Social Choice Re-examined*. Volume 1 of *Proceedings of the IEA Conference held at Schloss Hernstein, Berndorf, Austria*.
- Shabman L and Stephenson K, 2000. Environmental valuation and its economic critics. *Journal of Water Resources Planning and Management*, 126(6):382-388.
- Simon HA, 1945. *Administrative behaviour*. Collier-Macmillan, Toronto, Canada.
- Smith LG, 1992. *Impact Assessment and Sustainable Resource Management*. Longman Scientific & Technical, John Wiley & Sons, Inc., New York.
- Smith PGR and Theberge J B, 1987. Evaluating natural areas using multiple criteria: Theory and practice. *Environmental Management*, 11(4):447-460.
- Smith PN, 1994. Applications of fuzzy sets in the environmental evaluation of projects. *Journal of Environmental Management*, 42(4):365-388.
- Sondheim MW, 1977. A comprehensive methodology for assessing environmental impact. *Journal of Environmental Management*, 6: 27-42.
- Staley CS, 1989. *A History of Economic Thought: From Aristotle to Arrow*. Blackwell, Cambridge.
- Stansbury J, Woldt W, Bogardi I and Bleed A, 1991. Decision support system for water transfer evaluation. *Water Resources Research*, 27(4):443-451.
- Stevens SS, 1946. On the theory of scales of measurement. *Science*, 103: 677-680.
- Stevens TH, Belkner R, Dennis D, Kittredge D and Willis C, 2000. Comparison of contingent valuation and conjoint analysis in ecosystem management. *Ecological Economics*, 32(1):63-74.
- Stewart TJ, 1992. A critical survey on the status of multiple criteria decision making theory and practice. *OMEGA, International Journal of Management Science*, 20(5/6):569-586.
- Stewart TJ, 1995. Simplified approaches for multicriteria decision making under uncertainty. *Journal of Multi-Criteria Decision Analysis*, 4: 246-258.
- Stewart TJ, 1996. Robustness of additive value function methods in MCDM. *Journal of Multi-Criteria Decision Analysis*, 5: 301-309.
- Stewart TJ, 1998. Measurement of risk in fisheries management. *ORiON*, 14(1/2):1-15
- Stewart T J, Joubert A, Scott L and Low T, 1997. Multiple Criteria Decision Analysis: Procedures for Consensus Seeking in Natural Resources Management. Report to the Water Research Commission. WRC Report No 512/1/97.
- Stewart TJ and Joubert AR, 1998. Conflicts between conservation goals and land use for exotic forest plantations in South Africa. In: (eds) Beinat E and Nijkamp P, *Multicriteria Analysis for Land-use Management*. Kluwer Academic Publishers, Dordrecht.

- Stewart TJ, Joubert AR, Liu DF, 2001. Group decision support methods to facilitate participative water resource management. Report to the Water Research Commission. WRC Report No. 863/1/01.
- Stewart TJ and Scott L, 1995. A scenario-based framework for multicriteria decision analysis in water resources planning. *Water Resources Research*, 31(11): 2835-2843
- Stewart TJ, Scott L and Iloni K, 1993. Scenario based multicriteria policy planning for water management in South Africa. Report to Water Research Commission: WRC Rep No: 296/1/93.
- Stigler GJ, 1950. The development of utility theory I & II. *The Journal of Political Economy*, 58:307-327, 373-396.
- Strong EJ, 1983. A note on the functional form of travel cost models with zones of unequal populations. *Land Economics*, 59(3): 342-349.
- Tecle A, 1992. Selecting a multicriterion decision making technique for watershed resources management. *Water Resources Bulletin*, 28(1):129-140.
- Thayer MA, 1981. Contingent valuation techniques for assessing environmental impacts. *Journal of Environmental Economics and Management*, 8(1):27-44.
- Thurstone L, 1927. A law of comparative judgement. *Psychological Review*, 34:273-286.
- Tochner KD, 1977. Planning systems. *Philosophical Transactions of the Royal Society of London*, A287:425-441.
- Turpie JK, 1995. Prioritizing South African estuaries for conservation: A practical example using waterbirds. *Biological Conservation*, 4(3): 35-37.
- Turpie JK, Adams JB, Joubert AR, Harrison TD, Colloty BM, Maree RC, Whitfield AK, Wooldridge TH, Lamberth SJ, Taljaard S and Van Niekerk L, 2002. Assessment of the conservation priority status of South African estuaries for use in management and water allocation. *WaterSA*, 28(3):191-206.
- Turpie JK and Joubert AR, 2001. Estimating potential impacts of a change in river quality on the tourism value of Kruger National Park: An application of travel cost, contingent and conjoint valuation methods. *WaterSA*, 27(3):387-398.
- Turpie JK, Joubert AR and Mander M, 2000. Incorporating economic considerations into the determination of the environmental reserve: Proposed framework for incorporating economic criteria into the quantification, allocation and management of the environmental reserve for rivers. Working paper No. 4 of Water Research Commission Project No. K5/978.
- Turpie J, Joubert A, van Zyl H, Harding B and Leiman T, 2001. Valuation of open space in the Cape Metropolitan area: A pilot study to demonstrate the application of environmental and resource economics methods for the assessment of open space values in two case study areas: Metro South and Metro South-east. *A report to the City of Cape Town*.
- UN website, http://www.johannesburgsummit.org/html/whats_new/feature_story39.htm
- UNEP, 1995. Middle East Peace Process: Multilateral working group on water resources. Training course on "Project design and appraisal tools for integrated environmental and economic assessment". 4-8 December 1995. United Nations Environment Programme (UNEP), Kenya.
- van den Honert RC and Lootsma FA, 1996. Group preference aggregation in the multiplicative AHP: The model of the group decision process and Pareto optimality. *European Journal of Operational Research*, 96:363-370.
- van Huylenbroeck G, 1995. The conflict analysis method: Bridging the gap between ELECTRE, PROMETHEE and ORESTE. *European Journal of Operational Research*, 82:490-502.
- van Pelt M. JF, 1993. Ecologically sustainable development project appraisal in developing countries. *Ecological Economics*, 7: 19-42.
- van Praag BMS, 1978. The perception of income inequality. In: (eds) A.F. Shorrocks and W. Krelle, *Personal Income Distribution*. North Holland, Amsterdam.
- van Praag BMS, Hagenaaers AJM and van Weered H, 1982. Poverty in Europe. *Review of Income and Wealth*, 28(3): 345-359.
- Varis O, Kuikka S and Taskinen A, 1994. Modelling for water quality decisions: Uncertainty and subjectivity in information, in objectives and in model structure. *Ecological Modelling*, 74:91-101.
- Vatn A and Bromley DW, 1994. Choices without prices without apologies. *Journal of Environmental Economics and Management*, 26: 129-148.
- VISA, Visual Interactive Sensitivity Analysis. 1995. © V. Belton, University of Strathclyde. Visual Thinking international, Ltd. info@VisualT.com.
- von Neumann J and Morgenstern O, 1947. *Theory of Games and Economic Behaviour*. 2nd edition. Princeton University Press, Princeton.
- von Winterfeldt D and Edwards W, 1986. *Decision Analysis and Behavioural Research*. Cambridge University Press, Cambridge. 604 p.
- Walters C, 1986. *Adaptive management of renewable resources*. Macmillan Publishing Company, New York.
- Watson SR, 1981. Decision analysis as a replacement for cost/benefit analysis. *European Journal of Operational Research*, 7:242-248.
- Wenstøp F, undated. Rationality, emotion and MCDM. Unpublished manuscript. Norwegian School of Management BI.

-
- Wenstøp F, Carlsen A, Bergland O and Magnus P, 1997. Valuation of environmental goods with expert panels. In (ed) Climaco J, *Multicriteria Analysis*. Springer-Verlag, Berlin.
- Wenstøp F and Seip K, 2001. Legitimacy and quality of multi-criteria environmental policy analysis: A meta analysis of five MCE studies in Norway. *Journal of Multi-Criteria Decision Analysis*, 10(2):53-64.
- Weston SC, 1994. Toward a better understanding of the positive/normative distinction in economics. *Economics and Philosophy*, 10:1-17.
- Whittington D, Kerry Smith V, Okorafor A, Okore A, Liu JL and McPhail A, 1992. Giving respondents time to think in contingent valuation studies: A developing country application. *Journal of Environmental Economics and Management*, 22:205-225.
- Xiang W-N, Gross M, Gy-Fabos J and MacDougall EB, 1992. A fuzzy-group multicriteria decisionmaking model and its application to land-use planning. *Environment and Planning B: Planning and Design*, 19:61-84.