

PROJECT EVALUATION, SUSTAINABILITY AND ACCOUNTABILITY

Combining cost-benefit analysis (CBA)
and multi-criteria analysis (MCA)

Frans J. Sijtsma



REG-publication 27

ISBN 90-73709-24-5
ISBN (new) 978-90-73709-24-5

RIJKSUNIVERSITEIT GRONINGEN

PROJECT EVALUATION, SUSTAINABILITY AND ACCOUNTABILITY

Combining cost-benefit analysis (CBA)
and multi-criteria analysis (MCA)

Proefschrift

ter verkrijging van het doctoraat in de
Economische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
donderdag 19 oktober 2006
om 13:15 uur

door
Frans Jan Sijtsma
geboren op 26 mei 1964
te Veenwouden

Promotores:

prof. dr. J. Oosterhaven
prof. dr. A.J.M. Schoot Uiterkamp
prof. dr. D. Strijker

Beoordelingscommissie:

prof. dr. A. van der Veen
prof. dr. J.H. Voogd
prof. dr. mr. C.J. Jepma

Wêr't wy op de wrâld ek binne,
oeral skynt deselde sinne.

Wherever in the world that we may be,
it is the same sun everywhere that we see.
(Frisian proverb)

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Sijtsma, F.J.

Project evaluation, sustainability and accountability - Combining cost-benefit analysis (CBA) and multi-criteria analysis (MCA) / Frans J. Sijtsma - Groningen : Stichting Ruimtelijke Economie Groningen (REG-publicatie; 27). - Met lit. opg.
ISBN 90-73709-24-5
ISBN (new) 978-90-73709-24-5

Copyright 2006 Stichting Ruimtelijke Economie Groningen. Groningen, The Netherlands

Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced in any form, by print or photo print, microfilm or any other means, without written permission by the publishers.

Printing: Facilitair Bedrijf - University of Groningen

PREFACE

This book discusses project evaluation in a context where sustainability and accountability are crucial; therefore the scope is specific and narrow. However, three somewhat broader threads in my professional career as an economist converge to formulate the main messages of this work. In order to fully appreciate these main messages it seems useful to mention these threads.

The first thread in my career as an economist concerns how to reconcile the limited perspectives of an economist with myriad other valuable perspectives. This concern has largely occupied my thinking throughout my studies, and the late Taco Kastelein has played an essential role by helping me ‘frame’ my thinking, through introducing me to the history of economic thought and the more philosophically-oriented methodology of economics. Later on, especially during my career when performing country credit ratings and conducting business cycle analyses at the international research department of the AMRO Bank, I learned much more about both the power and the limitation of mainstream economic analysis within practical settings.

The second thread is my interest in spatial economics, and the importance of different spatial levels in economic analysis, whether labour markets, infrastructure and transport, regional development, or investment and innovation. While working with the very stimulating groups of spatial economists and economic geographers at the University of Groningen, the importance of distance and space in economic analysis became part of my professional ‘backbone.’

The final thread worth mentioning here is a fascination with how to draw science and society closer together. As a coordinator of the Science Shop for Economics, Management and Organisation, I have been involved in numerous economic and other research to assist organisations with scientific input. Working towards fruitful interaction between science and society, involving stakeholders in research, reconciling widely disparate viewpoints, and bridging communication gaps in practical ways has become ‘daily business.’

Considering the many people who deserve gratitude for contributing to this thesis, one may wonder why there is only one author on the title page. I cannot thank everybody.

Be that as it may, I would like to thank the co-authors of the case studies and background studies in part three: Dirk Strijker, Gert Jan Rotmensen, Agnes van den Berg, the late Katrin Bettels, Jella Hoefsloot, Ulrich Futh, Isabelle Maignan, Charles Vlek, Linda Steg, Dick Feenstra, Winnie Gerbens, Sigi Lindenberg, Henk Moll, Ton Schoot Uiterkamp, Martijn Broekhof, Michiel Nijboer, Arjen van Witteloostuijn, Anna Kusk, Krista Koeleman and Sybren de Jong.

I would like to thank the organisations that have provided financial and other support to the case studies: Vereniging Natuurmonumenten, Waddenvereniging and the Ahold Company.

I am grateful to the Faculty of Economics for graciously allowing me the opportunity to write the thesis, which has – for one day a week over a number of years – taken me away from my regular task at the Science Shop.

I would also like to thank Elise Kamphuis, my close colleague at the Science Shop, for her continuous support despite the fact that it was she who ‘suffered’ the most from my aspirations.

I am grateful to Henk Janssen at the Science Shop for assisting me with the layout and many other small ‘services’ throughout the years.

I thank Susan Davis (ssndavis@hotmail.com) for correcting my English and providing greater flow and richness to the text.

Johan Zwart has single-handedly designed a cover for the book which has markedly improved on the original idea.

I would also like to thank my three promoters: Jan Oosterhaven for his unrelenting scrutiny of the manuscript throughout its various stages, as well as his fierce defence of welfare economics, which has prevented me from arguing my case too easily. I thank Ton Schoot Uiterkamp for his subtle steering with many thought-provoking metaphors. I thank Dirk Strijker for his incisive and crystal clear common sense approach to both pure science and applied research, which has been a great help in many parts of the work.

I thank the examiners of my thesis Catrinus Jepma, Anne van der Veen and Henk Voogd for reading the manuscript.

I would also like to thank my family for their support in countless ways. Several ‘family threads’ come together too, in important elements of the thesis: the love of accounting and making balance sheets of my father Ouwe Sijtsma, the ease of reconsidering old alternatives and finding new ones of my mother Annie Sijtsma-Venema, a strong commitment to everything fragile and the importance of up-to-date information of my father-in-law Rients Tijsma, and the careful observation of reality without judgement of my mother-in-law Tjitske Tijsma-Bergema.

Finally, because home is where the heart is, I am grateful to my wife and soul sister Saepke and our amazing and sweet children Nyckle ‘car-designing’ Owe and Emke ‘horse-feeding’ Tjitske.

Veenwouden/Groningen, 9th of August 2006.

FJS

SUMMARY

Introduction

Sustainability has become a high profile objective. Decision-makers in governments and businesses must choose among different project alternatives which, in varying degrees, contribute to sustainability. Decision-makers also have to account for their choices to a large audience or a broad range of stakeholders.

This thesis is about the positive and negative aspects of using the main judgement-oriented evaluation tools of Cost-Benefit Analysis (CBA) and Multi Criteria Analysis (MCA) together in the context of sustainability and accountability. These evaluation tools are used for what might be called “accountable project evaluation in a sustainability context” or, put more simply, using them in the ‘sustainability context’.

Accountability presents a major challenge to the development of evaluation tools. If the demands of accountability are neglected in ‘open’ decision-making processes, the risk looms large that the evaluation will either be judged as irrelevant or unwarranted interpretation of results will occur. The main challenge of this thesis with regard to accountability is to improve the communicative quality of the judgement-oriented evaluation tools.

Sustainability also requires unique qualities from evaluation tools. Sustainability has been defined in various – sometimes confusing and vague - ways. This thesis gives an operational content to sustainability by identifying key elements of the concept. Sustainability evaluation – which enhances accountability - requires:

- Triple E assessment (that is, evaluation of impacts on Economic development, Extreme poverty and Environmental degradation);
- evaluation at both global and decision-making levels;
- evaluation over a (long period of) time;
- evaluation which increases everyone’s understanding (including awareness about limitations).

After careful examination of CBA and MCA performance within the ‘sustainability context’, the thesis provides both a theoretical basis for a combination of CBA and MCA, and insight into the practical applicability of this combination.

MCCBA

An integration of CBA and MCA as a fruitful addition to the evaluator’s toolbox within a sustainability context is presented in this thesis. A new tool, ‘MCCBA’, stands for ‘multi criteria cost benefit analysis’. The acronym reveals that the approach is a new combination of *existing* elements of CBA and MCA, rather than something entirely new.

The common ground between the two methods, which allows for the integration of CBA and MCA, is that full judgement is often impossible in a sustainability context, and that ‘decision aid’ is therefore the realistic aim for any evaluation tool; structuring and framing decision-making is central to the MCCBA approach.

Three major building blocks are presented to further specify the style of integration. The first is that stakeholder involvement is useful (mainly) for incorporating a spectrum of views and for assessing consensus. The second is that standardisation along a global long-term Triple E criteria structure is needed. The third is the recognition that judgement and measurement in evaluation are closely related.

We will discuss the major elements of the MCCBA approach that offer *direction* to the

evaluation at each stage; MCCBA has eight stages indicated below in Table ES-1.

Table ES-1 The eight stages of the MCCBA approach

MCCBA	
Stage one:	Identify function, project alternatives and scale of the evaluation
Stage two:	Involve a broad group of stakeholders
Stage three:	Organise judgement criteria on Triple E impacts
Stage four:	Quantify impacts physically
Stage five:	Aggregate monetary scores, consensus based
Stage six:	Aggregate non-monetary scores, consensus based
Stage seven:	Interpret trade-offs
Stage eight:	Perform sensitivity analysis and reconsider project alternatives

Stage one: Identify function, project alternatives and scale of the evaluation

As to the function of the evaluation, the most the analyst-evaluator can do is to *clear as much ground as possible*. To undertake this, the analyst works with judgement concepts that are necessarily partial, but are acceptable and clear to stakeholders of the evaluation. As to the definition of project alternatives, the specification of a null-alternative against which project alternatives are assessed would be very useful were it to be standardised. Furthermore, given its *functional limitations*, MCCBA should by itself be seen as important in finding new alternatives.

The MCCBA approach is *strict in the evaluation at different spatial levels*, as it will help clarify impacts in the sustainability context and prevent unnecessary and unwarranted perceptions of complexity among stakeholders. The *global level is standard* in the MCCBA approach; it acts as insurance against any tendency to overlook impacts. However, in most evaluations there will be at least one other spatial level closer to the decision-making, such as the regional, national or organisation level. In order to remain practical, the second scale of the analysis should stick closely to that which the decision-makers think is the natural scale for the analysis.

Stage two: Involve a broad group of stakeholders

In the MCCBA approach involvement of stakeholders should focus primarily on achieving a broad range of perspectives; doing this *may act as a mechanism for exploring consensus among stakeholders*. Seeking consensus has to start from lower-order measurement and judgement issues and work upwards. The highest-order judgement issues will usually be beyond consensus.

Stage three: Organise judgement criteria on Triple E impacts

MCCBA in the sustainability context has two standardised requirements that give direction to the organisation of criteria: The first is to *define separate Triple E criteria without redundancy and double counting*. As judgement and measurement are closely related, one should try to use or develop standardised measurement of the criteria; this standardisation preferably holds on different spatial scales. The second is to *analyse and present short and medium-term impacts separately from long-term impacts*. For the short and medium-term impacts on economic development, the CBA practice of discounting can be applied, but the other Triple E impacts are shown in their own right. All long-term Triple E impacts (including long-term impacts on economic development) are also shown in their own right. Because of their long-term relevance, these impacts are conceptually on a more equal footing without discounting.

Stage four: Quantify impacts physically

In this stage the measurement of impacts in their natural dimensions occurs. This ‘fact finding’ phase is probably the most important phase in any evaluation. Stage four ends with the *performance matrix*, which shows the scores of the alternatives on different criteria. In the sustainability context this performance matrix will usually be too large and thus too complicated to warrant clear interpretations. Stages five and six are therefore crucial, as they reduce the information to a smaller number of dimensions.

Stage five: Aggregate monetary scores, consensus based

A CBA is performed in this stage of MCCBA for the impacts that fit well within it. The limitation to impacts that ‘fit well’ is the reason why it is ‘consensus based’.

What should be included in the CBA part? The legitimacy of the MCCBA approach rests on the notion that CBA is a powerful evaluation tool that relates to everyday economic notions such as efficiency and monetary measurement, thus staying close to market-related valuation. Only those impacts that broad groups of stakeholders regard as well captured within the monetary measurement possibilities of CBA should be fit into CBA.

In order to clarify how MCCBA differs from mainstream CBA, the MCCBA approach uses the phrase ‘limited CBA’, which underlines the limited scope of this type of CBA. Nevertheless, this type of CBA has a robust quality. A pure MCA approach often lacks the crucial economic realism that is well captured in CBA.

MCCBA’s consensus based valuation with CBA can to some extent use the straightforward but labour-intensive procedure of asking for willingness to pay or willingness to accept monetary values. The further the decision-maker is removed from common everyday economic valuations (e.g., consumer preferences), the less reliable WTP or WTA valuations will be, and the more that MCA-like measurement techniques should be preferred.

Stage six: Aggregate non-monetary scores, consensus based

In stage six a consensus based MCA is performed on the remaining impacts. The results of the CBA can be seen as a separate MCA branch to be combined with the other branches of remaining impacts. The MCCBA challenge here is twofold: 1) *to reduce the number of criteria to a minimum – in a CBA style* and 2) *to use consensus based judgement criteria and measurement*.

Ad 1) In the sustainability context the clarity of the judgement concept will be easily blurred when stakeholder involvement leads to huge value trees. To reduce the number of criteria, the MCCBA approach above all derives inspiration from the CBA practice of avoiding double counting and incorporating causality. As with *CBA practice*, the MCCBA approach uses indicators situated at the end (or at one point) of a causal chain and avoids using intermediate indicators, which would easily lead to double counting.¹ As a final method of reducing criteria, MCCBA captures the most important criteria or criterion and ignores criteria with low weight or unclear or small impacts. Naturally completeness suffers from this procedure but the gain is an increased likelihood of understanding for many stakeholders – including the grasp of the limitations. At this stage the selection of the most important criteria should certainly be contingent on the increased problem understanding brought forward by the evaluation.

Ad 2) The second challenge of MCCBA at this stage is to perform an MCA which incorporates use of judgement concepts and measurements that have achieved consensus among broad groups of stakeholders, or about which consensus can be developed. Compared

¹ In fairness, however, one may observe that inspiration can emerge from the MCA side too, as a thorough analysis of causality to avoid redundancy of criteria is well-established in *MCA theory*.

to the performance matrix, what happens here is that several sub-criteria can be aggregated to comprise an overall criterion. Because the higher order valuations are most often disputed, further aggregation is avoided as much as possible in the MCCBA approach. The end best result of the MCCBA is an aggregation of the performance matrix, based on a broad consensus among stakeholders.

Stage seven: Interpret trade-offs

The major difference between the end result of stages five and six of MCCBA and that of CBA and MCA, is that insufficient judgement is provided. This thesis argues that although this is regrettable, in many instances it will be the optimum and realistically feasible result, because accountability is best served with MCCBA, and the provision of more judgement is likely to become uncertain, thus subjective and contestable.

However, although ultimately one may be unable to pass full judgement on the various project alternatives, there are more possibilities for judgement, while one also remains within the confines of an objective and accountable approach. Proceeding in small steps is recommended, as non-involved stakeholders need to understand the process, and the limitations of the analysis should remain clear to all. The goal of stage seven is to *interpret* the consensus based aggregation of the performance matrix with the aim to analyse *trade-offs*.

Generally, an important method for analysing trade-offs is by means of *ratio-analysis* on the basis of the aggregated performance matrix. Thus one examines, for instance, net-CBA outcomes per outcome on another criterion (Stewart, 2003). This type of analysis may closely resemble cost-effectiveness analysis.

Ratio-analysis may gain in strength if *comparisons across projects* can be made. Perhaps ratio-analysis standardisation of measurement in project evaluation enhances the possibility for comparison across projects. A clear noteworthy example of this standardised measurement is the Quality Adjusted Life Years (QALY) from health economics.

A related type of analysis is *stakeholder perspectives analysis*, which is to explicitly adopt the perspectives of specific stakeholders in order to assess the importance of different criteria to them – (what if priority were given to this criterion rather than that one or to those?) – and to present a preferred option from the perspectives. Several elements from the MCA toolbox are useful in this analysis; one will sometimes realise that despite choice of perspective, – from a logical set – the result remains the same.

Stage eight: Perform sensitivity analysis and reconsider project alternatives

The overall aim of stage eight of MCCBA is to share the increased problem understanding developed during the evaluation through an application of sensitivity analysis. The sensitivity analysis may have an important function for accountability; stakeholders can assess how several individual assumptions and estimates work out on the overall outcomes.

However, practical experience with sensitivity analysis shows that it can only have a limited role in sharing problem understanding. Therefore, MCCBA uses another powerful tool for sharing the increased problem understanding: it explicitly requests a reconsideration of the project alternatives. The following (types of) questions should be posed: *Does the increased problem understanding show that new alternatives should be considered or that existing alternatives should be combined?*

Questions about the consistency and feasibility of the selected goals in relation to the different project alternatives are relevant: increased problem understanding may also shed light on this.

Three MCCBA case studies

After presenting CBA and MCA and their integration into MCCBA, the thesis discusses three MCCBA case studies. The main aim of the case studies is to provide insight into the

applicability of MCCBA. The selection of the cases was made because, taken altogether they highlight the crucial strengths of MCCBA, which are not normally tackled by either CBA or MCA.

Evaluation of the Dutch Ecologische Hoofdstructuur (EHS): The value of combining monetary and non-monetary measurement.

This study, conducted in 1995, evaluated the implementation of the EHS (literally: Ecological Main Structure) in the Netherlands within the time period 1990-2020. The EHS implied that 244,000 hectares of new nature were to be developed on what was until then mainly conventional farmland.

Evaluation of the German Emssperrwerk: The value of more scale levels

The evaluation of this case study was carried out in 1998, a few months before the German government decided to build the Emssperrwerk. The Emssperrwerk had two official purposes, first as a (movable) dam in the event of extremely high water, thus preventing water from the sea threatening the hinterland in the EmsDollard region. Its second function was to facilitate large cruise ships built at the Meyer shipyard in Papenburg (approximately 30 kilometers upstream the river Ems) in order to reach the deep waters of the North Sea.²

Design and evaluation of Sustainable Corporate Performance (SCP) policy: The value of causal analysis.

The final case study of this thesis performed in 2001-2002 discusses Corporate Social Responsibility and sustainability, the combination of which is called Sustainable Corporate Performance (SCP). The Ahold Company, a leading Dutch multinational, was then in the process of formulating its SCP policy. Two closely connected concerns were particularly challenging to this group of quality supermarket and foodservice companies: first, the confusion about what SCP was or should be, and second, the lack of standardisation in measurement of SCP.

Lessons from the case studies (per stage)

The case studies revealed that in *stage one* the most important key word is acceptance; that is, acceptance of a change in function and acceptance of the different regional or organisational scales of the evaluation. This acceptance need not be ‘automatic’; stakeholders may have to be persuaded to accept certain points and success is not guaranteed.

As for the involvement of stakeholders in *stage two*, despite a ‘solution’ proposed in chapter 5 to focus merely on a broad group of stakeholders, case studies illustrate that even this type of involvement may remain difficult for various reasons.

The Triple E structure, which is proposed as a standardised element of MCCBA in *stage three*, is only preliminary. If no significant impacts can be found, the structure should be changed and criteria skipped. The global level assessment needed in MCCBA in many cases requires existing expertise and data; otherwise, too much time may be required to estimate impacts.

The case studies clearly and concretely acknowledge that *stage four* is indeed critical in many respects. One especially noteworthy aspect is the role of stakeholders in facilitating access to information and data.

Much has been learnt from the case studies about the aggregation of impacts and how it can be made ‘consensus based’. They clearly show that potential consensus only concerns the

² The huge Meyer shipyard in Papenburg specialises in building cruise ships. These ships have become ever larger and deeper, outgrowing the depth of the river Ems. With the Emssperrwerk the water level in the Ems can be raised temporarily when a cruise ship sails to sea once or twice a year.

minimum importance of criteria and ease of understanding measurements. This conclusion relates to *stages five and six* – to both monetary and non-monetary aggregation. From *stage six* alone it has become clear that simplicity of the analysis should/may overrule the demand of completeness.

Stage seven proved to be very rewarding in the case studies and its value is difficult to overestimate. It may turn out that the the main results of the evaluation are generated during this stage, as it could strongly increase the judgement value of the results from stages five and six.

The main conclusion from *stage eight* is that although the sensitivity analysis can show increased problem understanding – a stumbling block may be the attention it receives when presenting main results. The explicit demand to reconsider project alternatives can potentially remove this stumbling block, but was not yet practiced in the case studies.

The main building blocks of the MCCBA approach as identified above are that: 1) stakeholder involvement is useful for obtaining broad views and for checking on consensus; 2) standardisation of a global and long-term Triple E criteria structure is necessary; 3) judgement and measurement are closely related, and in an MCCBA both should be understandable to broad groups of stakeholders. Having discussed the case studies, a general conclusion is that mainly the aspect of the stakeholders is in need of ‘qualification’: stakeholder involvement may also be important for acquiring relevant data, but difficulties remain around fruitful involvement. The other building blocks have proven to be valid in practice.

In conclusion, we return to the title of this thesis: ‘Project evaluation, sustainability and accountability – Combining cost-benefit analysis (CBA) and multi-criteria analysis (MCA)’. The MCCBA combination developed in this thesis has proven to be a useful project evaluation tool in the sustainability context. Through its consensus based aggregation, MCCBA improves the communicative quality of the evaluation, thereby enhancing accountability. Through its decision aid style and its emphasis on standardisation, tracking causality, and allowance for simplicity to (here and there) overrule completeness, MCCBA acts as a safeguard against over-ambition, and is a facilitator for clarity, utility and practicality in a sustainability evaluation.

To end this summary we will give an overview of the main practical essentials of MCCBA.

Main practical essentials of MCCBA

How does the MCCBA approach evaluate whether a project is sustainable? How does it assure that accountability is well served?

The main practical essentials of the MCCBA approach can be encapsulated as advice to two target groups, experts/practitioners and laymen/users.

The first target group comprises the analysts/evaluators, and it is to them that this thesis is primarily aimed. The main points of advice to this expert group are:

1. Do not give ‘best’ solutions to problems and do not accept preferences too easily as ‘given.’ Be keen towards finding ‘possibly better’ solutions; be content to only frame and structure decision-making and remember that in sustainability matters most people will be ‘learning how to think.’
2. Sustainability should not be about everything. Surely sustainability entails the economic development of the planet and its people in different places, but

above that, only issues about which there is some consensus that global level outcomes are greatly worrying should matter: environmental degradation and extreme poverty are definitely cases in point.

3. Work with two evaluation scale levels: a decision-making and a global level. Evaluate the project at both levels (of thinking). In practice you may work with a lower than global level; the global level is merely an indication – to be on the safe side – of that highest level where economic, environmental, or extreme poverty impacts can be found.
4. Involve stakeholders, but focus on involving stakeholders with strongly different views, rather than on involving as many stakeholders as possible.
5. Starting from the impact matrix, use both CBA and MCA techniques for aggregation. Do not decide alone how to aggregate: try to achieve a consensus among your stakeholders on which aggregation is easiest to understand.
Furthermore, when aggregating:
 - Use the consumer-citizen divide of preferences, the first are usually well-captured in CBA, the second best delineated in MCA.
 - Very reluctantly discount long-term impacts, even if they are economic. Long-term economic impacts are matters of deliberation and learning how to think; they are not matters of mechanical discounting. Preferably use discounting only for short and medium-term economic impacts.
 - Use MCA scaling and weighing techniques to your benefit, but only in as much as they can be assured of wide consensus. This will often mean not scaling and weighing the highest order criteria.
 - Unless ‘not-all-that-important impacts’ can be easily aggregated, in the final stages of the evaluation focus only on the most important impacts.
6. Take pride in interpretation of sub-aggregated outcomes. You can:
 - Calculate various ratios (e.g., cost-effectiveness ratios),
 - Compare standardized impact scores with other projects (e.g., QualityAdjusted Life Year [QALY]; or GlobalWarmingPotential [GWP] or Naturescores [as in this thesis])
 - Analyze outcomes from different stakeholder perspectives (e.g., winners/losers, environmentalists-entrepreneurs).
7. In the final stage of your evaluation, address explicitly the question of whether increased problem understanding stipulates that new project alternatives should be considered or old alternatives should be combined.

The second target group consists of decision-makers and stakeholders. To this laymen ‘user’ group, the main points of advice arising out of the MCCBA approach are:

1. If you think that expert evaluators can judge easily whether a project is sustainable or not, that is a wrong assumption. You will have to think together.
You may be certain what to do afterwards, but it is likely that the evaluator will not. However, you should all definitively possess greater understanding of the problem situation.
2. You might think that judging sustainability is about looking into the distant future. Perhaps. But be aware; sustainability surely requires that you scrutinize your project from a high international or even global level – next to examining it from your ‘own’ level, certainly. Good courses of action

should preferably be good at both levels.

3. It is normal, even essential, that you understand how the evaluation results developed – although you need not understand every single detail. If you do not comprehend the process which has yielded the results, something has gone wrong.
4. You may think that it is best to involve as many stakeholders as possible in an evaluation. However, achieving a consensus on the evaluation outcomes with all stakeholders is illusory, as it is impossible to structure the evaluation by involving everyone. A consensus about easily understood ways to measure the most significant impacts is, however, a realistic goal. To that purpose, the involvement of a small group of stakeholders with widely differing views on a project seems more fruitful.

TABLE OF CONTENTS

PREFACE.....	V
SUMMARY	VII
PART ONE - INTRODUCTION	
SUSTAINABILITY AND PROJECT EVALUATION	1
CHAPTER 1: CBA AND ACCOUNTABLE EVALUATION IN A SUSTAINABILITY CONTEXT.....	3
1.1 Introduction.....	3
1.2 Central research question	4
1.3 The structure of this thesis.....	4
1.4 A note on the style of the thesis	6
CHAPTER 2: THE CHALLENGES FOR PROJECT EVALUATION TOOLS IN AN ACCOUNTABLE SUSTAINABILITY CONTEXT	9
2.1 Introduction.....	9
2.2 The judgement perspective in evaluation	9
2.3 Type of projects	11
2.4 Synthesis of the essential characteristics.....	23
2.5 Short summary of evaluation demands in a sustainability context.....	28
PART TWO – THEORY	
CBA, MCA AND THEIR INTEGRATION IN A SUSTAINABILITY CONTEXT	29
CHAPTER 3: CBA IN A SUSTAINABILITY CONTEXT.....	31
3.1 Introduction.....	31
3.2 The stages of CBA	37
3.3 CBA and the incorporation of broad welfare notions	52
3.4 CBA and valuing gains for some, losses to others.....	61
3.5 CBA and valuing environmental impacts	69
3.6 Conclusions: CBA's performance in a sustainability context	82
Appendix to chapter 3 Illustration of pitfalls of environmental valuation - The Costanza attempt.....	84
CHAPTER 4: MCA IN A SUSTAINABILITY CONTEXT.....	89
4.1 Introduction.....	89
4.2 The stages of an MCA.....	92
4.3 MCA and organising stakeholder involvement.....	107
4.4 MCA and interdependence of criteria and double counting	110
4.5 MCA and its judgement orientation.....	115
4.6 Conclusions: MCA's performance in the sustainability context.....	120
CHAPTER 5: MCCBA: INTEGRATING CBA AND MCA IN THE SUSTAINABILITY CONTEXT	123
5.1 Introduction.....	123
5.2 Common ground for integrating MCA and CBA in MCCBA	123
5.3 Major building blocks for integrating MCA and CBA in MCCBA.....	126
5.4 Stages of the MCCBA approach.....	128
PART THREE – USING MCCBA	
THREE CASE STUDIES	139
CHAPTER 6: INTRODUCTION TO PART THREE.....	141
CHAPTER 7: EVALUATION OF THE DUTCH ECOLOGISCHE HOOFDSTRUCTUUR (EHS): THE VALUE OF NON-MONETARY MEASUREMENT	143
7.1 Introduction.....	143
7.2 Original results.....	144
7.3 The Ecological Main Structure evaluation in light of the MCCBA approach.....	158
CHAPTER 8: EVALUATION OF THE GERMAN EMSSPERRWERK: THE VALUE OF MORE SCALE LEVELS	169
8.1 Introduction.....	169
8.2 Original results.....	170
8.3 The Emssperrwerk in the light of the MCCBA approach	183

TABLE OF CONTENTS

<i>Appendix 1 to chapter 8: Sensitivity analysis for four crucial parameters</i>	<i>191</i>
<i>Appendix 2 to chapter 8: The results of the total nature score approximation for Niedersachsen and Germany</i>	<i>193</i>
CHAPTER 9: DESIGN AND EVALUATION OF SUSTAINABLE CORPORATE PERFORMANCE (SCP) POLICY: THE VALUE OF CAUSAL ANALYSIS	195
9.1 <i>Introduction</i>	195
9.2 <i>Designing SCP policies following the MCCBA approach.....</i>	196
9.3 <i>Exploring a concrete example: Evaluating EurepGAP</i>	213
9.4 <i>Conclusions</i>	221
<i>Appendix to chapter 9: Some examples of sector-level SCP projects</i>	223
 PART FOUR – CONCLUSIONS AND FINAL REFLECTIONS	
THE MERITS OF THE MCCBA APPROACH	225
CHAPTER 10: CONCLUSIONS AND FINAL REFLECTIONS	227
10.1 <i>Introduction</i>	227
10.2 <i>Additional conclusions (additional to conclusions of chapter 5)</i>	227
10.3 <i>Final reflections</i>	229
 LITERATURE CITED.....	
SAMENVATTING.....	263
APPENDIX: ELABORATE TABLE OF CONTENTS.....	273
LIST OF TABLES.....	281
LIST OF FIGURES	282
 PUBLICATION SERIES OF THE STICHTING RUIMTELIJKE ECONOMIE GRONINGEN	
	283

PART ONE - INTRODUCTION

SUSTAINABILITY AND PROJECT EVALUATION

CHAPTER 1: CBA AND ACCOUNTABLE EVALUATION IN A SUSTAINABILITY CONTEXT

1.1 Introduction

Sustainability has become a high profile objective. Decision-makers in governments and businesses have to choose between different project alternatives that contribute in varying degrees to sustainability. However, to pass sound judgement as to whether or not a specific project contributes to sustainability is not always a straightforward matter. Furthermore, decision-makers have to account for their choices to a large audience or a broad range of stakeholders. Project evaluation tools should therefore be useful, not only for helping decision-makers to choose between alternatives but also to compel them to account for their choices.

Since at least the 1950s, Cost-Benefit Analysis¹ (CBA) as a project evaluation tool has been used across a wide spectrum for a large number of projects. CBA is often associated with objective decision-making and increasing accountability of government decisions (Farrow and Toman, 1999). CBA has strong links with economic science, which is one reason why it has a strong “economic flavour” about it. This economic flavour, however, gives rise to two problems. First, the fact that many public policy decisions and business decisions have become accessible to stakeholder involvement, has meant that, increasingly, the judgement criterion and technical calculations of CBA are not easily understood by stakeholders without expertise in economics. Second, discussions about sustainability require that environmental and social justice aspects be assessed within CBA. Such an assessment is problematic, as CBA tends to give market-related aspects, or easily monetised aspects, a predominant position.

Multi-Criteria Analysis (MCA), is probably the most important alternative available to the analyst-evaluator who may be somewhat discontent with CBA. Confronted with the demands posed by the sustainability context, the analyst-evaluator may appreciate MCA’s apparent simplicity as well as its flexibility in integrating a wide range of concerns. MCA practice and theory rest heavily on the fruitful interaction between the analyst and a single decision-maker or a small group of decision-makers or stakeholders. This dependency may facilitate the introduction of ‘subjective’ elements in the evaluation, which contrasts sharply with CBA’s ‘objective’ stance. Moreover, although stakeholder involvement may be more commonly practiced within MCA than within CBA, accountability towards broader groups of stakeholders may suffer from the introduction of these subjective elements.

This thesis is about the pros and cons of using CBA and MCA as tools for, what might be called “accountable project evaluation in a sustainability context”, or “project evaluation in the context of accounting for sustainability.” Accountable project evaluation is more demanding than mere project evaluation, and project evaluation in a sustainability context is even more demanding of a rigorous approach.

Several authors have suggested that in the sustainability context the analyst-evaluator might use a combination of CBA and MCA (Van Pelt et al., 1990; Toman, 1998 and 1999; Adler and

¹ The terms Cost-Benefit Analysis and Benefit-Cost Analysis are equivalent. See Porter, 1995 for an earlier history of CBA.

Posner, 1999; Turner et al., 2000; Howarth et al., 2001). As MCA and CBA appear to be ‘worlds apart,’ it should come as no surprise that, until now, neither in theory nor in practice has the suggestion to combine methods been seriously worked out. This thesis explicitly considers the combination of CBA and MCA. After a careful examination of CBA and MCA performance within the sustainability context, the thesis tries to provide a theoretical basis for a combination of CBA and MCA (dubbed ‘MCCBA’) as well as to give insights into its practical applicability.

1.2 Central research question

The main aim of this thesis is to develop a general framework for accountable project evaluation in a sustainability context, which uses a combination of CBA and MCA. In addition to this central concern, it also tries to prove whether it is possible to apply this general framework in practice.

The central research question is:

How can CBA be combined with MCA to make a better tool for accountable project evaluation in a sustainability context?

To answer this central question, the following sub-questions will be addressed:

- What, ideally, are the demands that project evaluation tools have to meet, if on the one hand they are meant to be useful for decision-making in a sustainability context, and on the other, are to facilitate accountability? (chapter 2).
- How well does CBA performance meet these demands? (chapter 3).
- How well does MCA, the most relevant alternative evaluation tool, meet them? (chapter 4).
- Can a project evaluation approach be formulated to minimise the number of cons and maximise the number of pros of both CBA and MCA? What would be the essential aspects of such a Multi-Criteria Cost-Benefit Analysis (MCCBA)? (chapter 5).
- How applicable is this MCCBA approach in practice in various project evaluation cases? (chapters 6, 7, 8, 9, and 10).

1.3 The structure of this thesis

The thesis comprises four main parts. The introduction consists of chapters 1 & 2; part two consists of chapters 3, 4 and 5 and is primarily theoretical. In contrast, part three, consisting of chapters 6, 7, 8, and 9 is practical: it is mainly a discussion of three different case studies. The thesis concludes with part four, which is chapter 10; this chapter draws conclusions about the theoretical and practical value of the MCCBA approach to evaluation within the sustainability context.

When we consider the thesis structure per chapter, chapter 2 treats the demands that a project evaluation tool in a sustainability context has to meet. In chapter 3, the first chapter of part two, an analysis on how well CBA meets these demands will be conducted. In chapter 4 the same

process is undertaken for the most important alternative evaluation tools to CBA: the family of Multi-Criteria Analysis techniques. Chapters 3 and 4 will discuss that both CBA and MCA have strengths and weaknesses when used in the sustainability context. The insights from chapters 2 to 4 are synthesised in the final chapter of part two. In this chapter (5) the main theoretical conclusions are presented and a general approach to the use of CBA and MCA in a sustainability context is formulated. The approach, named MCCBA, combines CBA with MCA and better meets the demands which were identified in chapter two.

Part three illustrates cases in which the MCCBA approach has been practiced. Chapter 6 introduces these cases and highlights their differences in character. In chapters 7, 8 and 9 summaries of different research cases are presented then followed by a discussion in the context of Part two. The case studies are:

Chapter 7: Evaluation of the Dutch ecological main structure. This case is a typical public policy case. The evaluation was performed at only one spatial level, the highest decision-making level. Substantial effort was made to develop useful ecological indicators against the background of attempting to acquire comparable measurements of economic and ecological effects.

Chapter 8: Evaluation of the German flood barrier in the river Ems. This case is a mix of public policy and private enterprise strategy. Although here again the measurement of ecological effects is quite important, the major innovation was most likely that it clearly showed the strength of evaluation at different spatial levels.

Chapter 9: Design and evaluation of sustainable corporate performance policies with an illustrative evaluation of EurepGAP, a supply chain quality standard for agricultural products implemented by European retailers. This case has a strong focus on the value of causal (economic) analysis and, again, the importance of evaluation at a correctly identified high spatial/organisational level.

Part four has one chapter, chapter 10, which draws conclusions both from the theoretical part two and the application-focussed part three.

Although the order of treatment of theory and practice in parts two and three may seem quite logical, the reader should be aware that, at least historically, it is not. The main innovative features of the methodology were developed gradually, over the course of answering *practical* evaluation requests, and the methodology was developed further in a mix of academic and practical settings. The second part of this book outlines a selection of important cases in historical order. Throughout these cases, the CBA/MCA evaluation approach was developing. The cases therefore do not show instances of the *full application* of the new approach, but they do show aspects and instances of its – historic – development. To a large extent, the first part of the thesis underpins theoretically what had already been developed in practice.

This thesis actively acknowledges what Brent calls “the constant to-and-fro between theory and practice” (Brent, 1996, p. xiii). It tries to discuss CBA and MCA – even in the theoretical part – not as an abstraction but as a real practice.² The constant to-and-fro between theory and practice

² Compare Adler and Posner (1999) who see the opposite as more common for CBA and French (1998) who notices a somewhat comparable state of affairs for MCA. See also Bana E Costa (1997, p. 35): “...there was agreement that some popular multicriteria procedures lack a proper scientific basis.” Little and Mirlees (1994) overseeing project evaluation since the 1960s and 1970s at the World Bank, also express the rather sorrow state of affairs concerning the link between theory and practice.

hopefully leads us to a stronger connection between theory and practice.³ Looking at figure 1.1, one can see that the chapters in part two, although largely theoretical, contain strong elements of current practice and reflect on that practice as well. The chapters in part three, although (largely⁴) practical case studies, have directly influenced part one, and subsequently, part two.

1.4 A note on the style of the thesis

The thesis will closely scrutinise project evaluation in the context of accountability and sustainability. CBA and MCA's performance in this context will be discussed and a new combined approach: MCCBA will be formulated. The assessment of different evaluation methods could take a variety of forms, as it involves a substantial body of scientific literature⁵ and the issues involved are quite complex; both factors could easily lead one to a type of thesis that ultimately has minimal accessibility to the non-expert, due to jargon or by having complicated mathematical equations or technical details.

The current thesis is written with the notion that, ultimately, the key issues are non-technical. Therefore, it seems reasonable to avoid jargon, mathematics and other barriers wherever possible, without turning to the other extreme of being inaccessible to the expert by introducing new words for concepts that may have a long tradition in a specialist field. Subsequent chapters will focus on the non-technical essentials of the sustainability context of evaluation in the hope that this treatment may not only serve the needs of more theoretically-gearred scientists, but also the requirements of the practitioner-analysts, as well as those of interested laymen and evaluation stakeholders.

³ See for instance, Svenson (1998).

⁴ The exception being chapter 9, which has a more substantial theoretical part.

⁵ Several fields of scientific literature are relevant to this thesis and are actively treated: economics, welfare economics, CBA, decision theory, MCA, evaluation, sustainability, ecological economics, and corporate social responsibility (see reference list). The choice was made not to dig further into several other relevant fields (e.g., ecological indicators, public choice, life-cycle analysis, development economics, and financial assessment). Due to the 'publication explosion', which has occurred in nearly all pertinent fields, limits (in reading, let alone discussion) on rather practical grounds are necessary if the thesis is to be a project with not only a beginning but also an end.

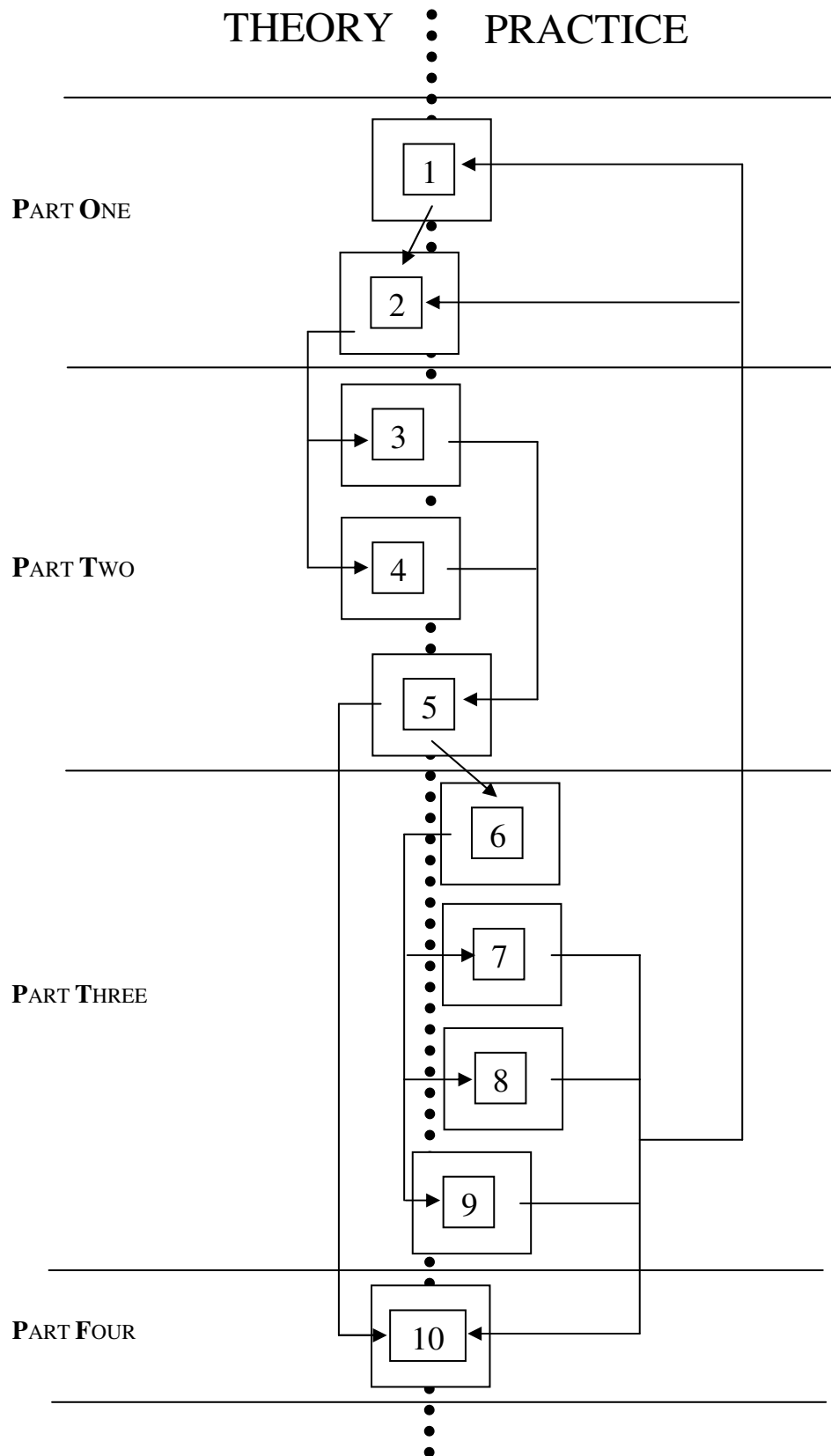


Figure 1.1: Graphical illustration of how the chapters in the thesis relate and how they are part of the “to-and-fro” between theory and practice.

CHAPTER 2: THE CHALLENGES FOR PROJECT EVALUATION TOOLS IN AN ACCOUNTABLE SUSTAINABILITY CONTEXT

2.1 Introduction

In which settings should the combination of CBA and MCA be applied? Because evaluation is a very broad research field, it may be clear to the reader that this thesis is not developing an ‘all purpose’ evaluation approach. Evaluation literature indicates that before making claims about the applicability of a method, one should try to be specific about both the setting in which a method is to be used and the purpose of the evaluation (Rossi et al., 1999, p. 94).¹ This chapter will sketch the challenges to an accountability-enhancing project evaluation tool that has to be used in a sustainability context.

In order to specify the characteristics required by the evaluation tool, section 2.2 will discuss the purpose of the evaluation and section 2.3 will treat what are, generally speaking, the types of projects to be evaluated. Section 2.4 synthesises the essential aspects necessary for the evaluation tool, while section 2.5 provides a brief summary.

2.2 The judgement perspective in evaluation

Three different perspectives in evaluation are often identified when discussing the entire field of evaluation research (Chelimsky, 1997; Patton, 1997;² see figure 2.1):

- the judgement perspective
- the enlightenment perspective
- and the development perspective.

The choice of perspective goes hand-in-hand with the intended purpose of the evaluation and with the types of questions being posed.

The first perspective is the *judgement* perspective. The aim of evaluation from this perspective is judging merit, worth or value. The evaluation is focussed on measuring results or efficiency.

The second perspective is the *enlightenment* perspective. Evaluation here aims at generating knowledge. The evaluation tries to acquire a more profound understanding in some specific area or field; its aim is generalised knowledge and ideas, often drawing from a body of existing evaluations. Enlightenment is, however, not always the primary aim, but it may simply occur as a by-product or accidental spin-off of evaluations done from the other two perspectives (Weiss, 1988 and 1990, p. 22).

The third perspective is the *development* perspective, which has the role of facilitating improvement of existing activities and organisations. Evaluation literature in general regards this facilitation in the light of providing (evaluative) help to strengthen organisations or institutions.

¹ Compare Fischhoff, 1991 about processes that stimulate exaggerations of applicability.

² Although the following terminology is more commonly used (respectively): the accountability, the knowledge and the development perspective. In this thesis judgement will be used, because the word accountability is already used here in a somewhat different, although related, meaning. The word enlightenment will be used (Weiss, 1988, p.176) because knowledge seems to be the purpose of every type of evaluation, while enlightenment has more the connotation of “*just for knowledge’s or understanding’s sake*”.

In the specific context of *project* evaluation, the development perspective is probably best understood as aiming at the generation of new and better alternatives.

Chelimsky specifies several dimensions in which the judgement, the enlightenment and the developmental perspective in evaluation differ. Two closely linked dimensions are crucial in this context:

- the relationship between the analyst/evaluator and the client/decision-maker, and
- the acceptability of evaluation findings to clients or users.

Development-oriented evaluation often has the evaluator as part of the team that works on the evaluated program. The relationship is close-knit, and this strong connection shapes the evaluation. The evaluator is sometimes referred to as ‘consultant’ or ‘critical friend’. Although projects may be evaluated, it is probably best to think in terms of policy programs that are being evaluated. The acceptability of evaluation findings is straightforward, as there is no threat being posed.

In judgement-oriented evaluation the relationship between the evaluator and the evaluated is more distant and the evaluator usually adopts an external perspective. The evaluator works with externally valid yardsticks measuring performance. Acceptability of findings may be difficult, as results *can* be controversial. The potentially problematic conditions of acceptability seem to be a crucial aspect of judgement evaluation.³ The client/decision-maker can request an evaluation, but there is some risk involved: outcomes can go either of two ways, pleasing or displeasing.

The enlightenment-oriented approach can take different forms and cultivate various close or distant relationships. Here, for instance (on the distant side), meta-analysis of many different evaluation studies can be performed to seek general agreement in outcomes. Due to the purpose of the evaluation, acceptability is less of an issue here. Findings that are disliked may simply be ignored.

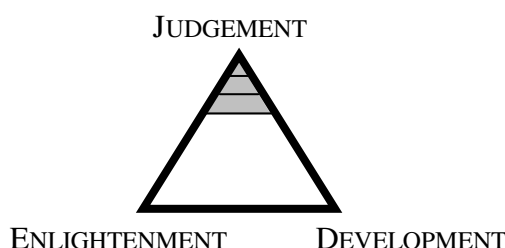


Figure 2.1: Three perspectives in the evaluation process

The three perspectives are not necessarily always mutually exclusive, and it can even be argued that more often than not in any particular evaluation elements from all three perspectives are present to some extent. This situation can be illustrated by placing any particular evaluation *within* the triangle of figure 2.1: it can be situated in the middle, balancing the three demands from different perspectives, or it can be closer to one or two of the perspectives.

The starting perspective of this thesis is the shaded upper corner of figure 2.1: judgement. The reason for this is of course that sustainability requires judgement about which project-alternatives should be rejected and which should be realised.

The traditions of both CBA and MCA have strong roots in this perspective.⁴ In evaluation for

³ One may notice that this possibly controversial aspect is closely related to the use of the word accountability in this thesis: open decision-making, with external stakeholders involved.

⁴ See Hellendoorn (2001) for a (Dutch) overview of different evaluation techniques and their main characteristics.

judgement, specifying the judgement criteria is “central and critical” (Patton, 1997, p. 66)⁵ and vital to our discussion below.

2.3 Type of projects

2.3.1 General characteristics

This thesis discusses specific types of projects to be evaluated; it is therefore important to identify the particularities of the types of project evaluation we consider. The projects:

- a) are new activities with uncertain outcomes
- b) are large and complex
- c) have impacts which carry on over time
- d) have economic and other impacts
- e) require accountability
- f) are evaluated upon sustainability.

Ad a)

The projects considered are often new activities. Examples include building a bridge where there was none, changing land-use to a new function, or introducing a new environmentally friendly product line. Evaluation is *ex ante*: it is beforehand [and not *ad itinere* (monitoring along the way) or *ex post* (looking back)]; this starting point logically implies that impacts are (to some extent) uncertain. When these uncertain effects are negative, one often speaks of risk. Evaluation has to handle uncertainty and risk.

Ad b)

Projects are government or business projects large enough and complex enough to set up a formal evaluation (they are not decided in an informal setting). ‘Large enough’ refers to the fact that evaluation itself involves costs that should be warranted by the potential benefits of evaluation. ‘Complex enough’ refers to the fact that project impacts cannot be identified and estimated easily. A systematic effort towards structuring alternatives and gathering relevant information is therefore useful in reaching a judgement on the best alternative. Evaluation should increase understanding of complex issues.⁶

Ad c)

Projects carry on over time and have a starting and ending point, which more or less excludes policy analysis as a subject of evaluation.⁷ Projects will often require investments, and therefore costs, at the beginning, with benefits being realised later. Impacts of projects are generally unevenly distributed over time, so the evaluation of this type of project means that the time dimension be handled.

Ad d)

⁵ The thesis examines evaluation experience and theory in order to derive a new improved methodology, which is drawn foremost from the knowledge perspective.

⁶ Understanding logically also implies understanding the limitations to the evaluation.

⁷ Although – other than evaluators in general – economists seem rather flexible in defining policy options as project alternatives.

The project has important economic aspects, but aspects are not exclusively economic. The economic effects are beyond the ‘merely financial’. Because CBA has a good record of capturing economic effects, we are especially interested in CBA as an evaluation tool. However, the projects considered will also often have economic, environmental (compare Pearce, 1998b), and other impacts.⁸ Therefore, the main impacts are not only financial or ‘economic’; in narrow economic cases, the evaluator’s existing toolkit seems to suffice. But in the evaluation process the analyst has a keen interest in assessing different aspects in a comparable way. What seems a logical treatment for one aspect may be unrealistic or irrelevant for another. Bringing different aspects in line somehow, is a major task for any project evaluation tool.

Ad e)

The projects require accountability, that is, they must have public or stakeholder consent and cannot, or are not, decided by an isolated individual decision-maker or a small group of individuals (management team). Decision-making is, to some extent, ‘open’ and can be actively open in the sense of inviting stakeholders to participate. It can also be tacitly open in that consideration has to be given to a wider audience. The type of project evaluation discussed here can be applied for public policy evaluation, that is, governments making policy decisions and discussing or accounting for these through interaction with various groups in society. Accountability may also be part of evaluation of business projects under the heading of sustainable corporate performance or corporate social responsibility. Because of the size and power of multinational companies, but also because of their vulnerability due to global competition and the importance of image and brand reputation, multinational enterprise decision-making is advancing towards more (active) stakeholder involvement: corporate decision-making is more “public”, and requires accountability to different stakeholders. The need for greater accountability to stakeholders has three implications for evaluation:

- First, evaluation has to be comprehensible in order to facilitate stakeholder participation.
- Second, stakeholder participation implies that the spatial or organisational scale of evaluation may no longer be exclusively limited to that deemed most relevant to the decision-maker. Stakeholders can ‘bring in’ other levels.
- Third, accountability and active stakeholder involvement seems to imply a striving for standardisation in evaluation. If consideration is given to the impacts of a project, accountability is greatly enhanced when the evaluation follows a fixed format, because standardisation can prevent manipulation of outcomes and may make results easier to understand.⁹ The standardisation of evaluation, to the largest extent possible, is therefore preferable.

Ad f)

Projects are evaluated in the context of assessing sustainability. Economic activity has its impact on the natural environment and on human development. The section below will delve further into the concept of sustainability and its implications for evaluation. We will show how the sustainability context sharpens or strengthens most of the aforementioned characteristics.

⁸ The author’s experience in project evaluation is largely in the field of spatial economics: land-use changes, spatial redistribution of economic activity, infrastructure development, and sustainable corporate performance.

⁹ Not least because one can build up experience over several projects in using a standardised evaluation approach. Compare Farrow and Toman (1999).

2.3.2 *Evaluating on sustainability*

2.3.2.1 The concept of sustainability

Although public concern about environmental damage caused by economic activity has existed for a long time (Edwards-Jones et al., 2000, p. 26), concern has progressively heightened in recent decades regarding the Earth's capacity to sustain economic activity. When the UN established the World Commission on Environment and Development in the 1980s, ecological impacts resulting from economic issues have since been seen in a broader sustainability context.¹⁰ Many governments, NGOs, citizens, and businesses have committed themselves to achieving greater sustainable development¹¹, and it has become "a high profile objective" (Pearce, 1998a, p. 69). But what exactly is sustainable development?

Defining sustainability and sustainable development is often seen as a difficult task, as in many contexts: 'a more sustainable world' has come to mean 'a better world', which can have many (normative) connotations and dimensions, from the economic and environmental to the social, cultural, historical, and ethical vantage points. The concept can sometimes be very vague (Sillanpää, 1998) or "somewhat obscure, often unique to the speaker" (Sarkar, 1997), which does not help to facilitate an evaluation.¹²

However, for our purposes, the essentials of sustainability will become clear enough. Furthermore, in discussing these essentials and the possibilities of different evaluation tools in handling these essentials, the reasons for the vagueness – that may occur sometimes – will be explained.¹³ The sections below will therefore define sustainability, or perhaps better, give it operational content, by identifying key elements of the concept (compare Pannell and Schilizzi, 1999, who follow a similar approach). These key elements together comprise 'the context of sustainability' for project evaluation.

It is noteworthy that the sustainability concept is on the fringe of science and public policy (Jacobs, 1991; Kamminga, 2001). Although from a scientific viewpoint it might be preferable to separate the normative from the positive aspects, this approach proves to be difficult in practice. However, there are at least two reasons why one need not worry too much about this situation. First, is that this state of affairs coincides with what philosophers of science see in many scientific fields, and perhaps may see in every scientific field if they look closely enough.¹⁴ Second, the focal point of the discussion here is project evaluation, which is often applied in a political setting and strives to be practical; so one nevertheless has to deal with the normative-political reality anyway (compare O'Connor, 2000, p. 172).

Finally a word should be added on the institutional context of sustainability evaluation. This

¹⁰ See the website of the UN Commission on Sustainable Development: www.un.org/esa/sustdev/csd/csd.htm.

¹¹ See for instance the Millennium goals of the United Nations (www.un.org/millenniumgoals), or the Council of the European Union (2006), or WBCSD, 2000.

¹² "Sustainability is at once extremely important and practically useless" (Pannell and Schilizzi, 1999, p. 65). Bromley (2001) sees sustainability as both "a fine idea" as it brings the fate of future generations to our attention and as "a hopeless concept", because it lacks operational content. Compare Little and Mirlees (1994, p. 213): "Sustainability has come to be used in recent years in connection with projects. This is more of a buzzword ... than a genuine concept. It has no merit. Whether a project is sustainable ... has nothing to do with whether it is desirable."

¹³ Running ahead a little: these reasons seem to lie in stakeholder involvement identifying mere 'concerns' and a lack of causal analysis preventing sustainability in a specific evaluation setting.

¹⁴ See for instance Feyerabend, 1975 and compare Sheldrake, 1997 for a reading on several interesting blind spots in science.

thesis addresses general issues in evaluating sustainability of projects and the related issue of accountability. However, at lower than global level (e.g., at EU or country level) and for subsets of sustainability issues (e.g., for environmental issues only), several detailed institutional contexts are relevant. For instance, in both the EU and the US, environmental impact assessments have long and well established traditions in which a specific legal/regulatory embedding of evaluation has developed by including several forms of stakeholder involvement in the evaluation (Thérivel and Partidário, 1996; Harrop and Nixon, 1999; Morris and Thérivel, 2001; and for instance Directive 2001/42/EC on environmental assessment of plans and programmes). Mutatis mutandis the same holds for the evaluation of transport infrastructure measures (SACTRA, 1999; Eigenraam, et al., 2003). Obviously we will borrow from literature and experience/case studies in these separate ‘branches,’ but our main focus is the general issue of evaluation of sustainability, which assures scrutiny of the integration of separate branches.¹⁵

2.3.2.2 A long-term view (f1)

The first key instruction in the sustainability evaluation is to adopt a long-term view.¹⁶ The Brundtland definition – probably the most accepted definition of sustainable development – (WCED, 1987, p. 43) is that:

“Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs.”

The Brundtland definition emphasises the future-oriented aspect of sustainability. There seems to be little disagreement between science and politics in this respect. Following the explanation of Pearce (1998a, p. 69), one may say that sustainable development can ‘endure’, it can ‘last’, which should be possible to ‘keep in being’.¹⁷ In business contexts a long-term view is sometimes seen as more than 10-15 years forward. Although one need not rule out this possibility in any specific evaluation context, many authors have much longer time periods of analysis in mind. For instance, Pearce (1998a) argues that 100 years might be taken as a minimum (!) time horizon.

Evaluation should (be able to) show long-term future impacts, which can be understood as strengthening the characteristics of point ‘a’ in section 2.3.1 above (ex ante evaluation, future oriented); and in point ‘b’ (impacts carry on over time).

2.3.2.3 A global (and local) perspective (f2)

Another key element that defines evaluation in the sustainability context is the global perspective: “sustainability calls for the globe as a relevant unit of analysis” (Schütz, 2000, p. 373; compare Grasso et al., 2003). The starting point of many analyses of (un)sustainability

¹⁵ Compare Meppem and Bourke (1999) for an analysis of difficulties in overcoming paradigmatic fixations in the ‘communicative turn’ that is needed in sustainability matters. Compare further, for instance, the EU difficulties in making progress in its Sustainable Development Strategy (SDS) and especially in balancing sustainability demands with the demands from the Lisbon agenda for growth and jobs (Council of the European Union, 2006).

¹⁶ See Munasinghe et al. (2001).

¹⁷ Compare Spash 2000; Spash sees two types of ‘economic’ definitions of sustainability: Ends-based definitions (e.g., non-declining consumption or utility) and means-based definitions (e.g., a non-declining stock of capital for “producing” well being).

lies at the global level (e.g., Clayton and Radcliffe, 1996; Rischard, 2002). The background for this state of affairs is at least twofold.¹⁸ First, consequences of many small scale actions (e.g., local or company level) are actually felt, and can be traced, throughout the world. This worldwide causality applies for all types of impacts, environmental, economic or otherwise. The global scale outlook is then necessary for not overlooking impacts.¹⁹ Second, size and scale of impacts has increased and many detrimental impacts have become highly visible on a global scale (Jepma et al., 1996; Nederveen Pieterse, 2000). The economic system has globalised at a rapid pace. Even small sized enterprises are involved in worldwide trade and production processes. For large multinational enterprises it is no longer ‘far off’ to take the impacts of their actions seriously at the global level. In management literature terms such as “global sourcing” and discussion on parts or sectors of the market economy as a ‘global village’ illustrate this earnest attitude. The challenge for evaluation is to show impacts on a global scale, since without the global scale, sustainability cannot be determined.²⁰ However, the challenge for evaluation is a bit more complicated. In the political field the slogan ‘think globally, act locally’ has become popular, which clearly shows at least one extra spatial level of analysis involved: the local action level (compare Curtis, 2003). Evaluation should help to guide decision-makers in their actions. Decision-makers in practice cannot *only* look at the global level to see the impacts of their actions; the spatial level at which a decision is taken, which in many cases is the administrative level, (in public policy) or a firm level (in corporate decision-making), will in practice play the most important role in the assessment of projects or policies.²¹ This situation implies a double challenge to evaluation: analysis at various spatial or organisational levels. Apart from the level closest to the decision-makers, the global spatial level is essential for a true assessment of sustainability.²² Use of the word global here is merely a standardised indication (to be on the safe side) of the maximum area in which effects are seen.

2.3.2.4 Economic development, environmental degradation and extreme poverty (f3)

2.3.2.4a *Three elements: Triple E*

The third key element is that sustainability involves at least three crucial components: economic development, environmental degradation and extreme poverty (Langhelle, 2000; Forte et al., 2001). Again, the report of the Brundtland Commission may be referred to, as it is often seen as “the key statement of sustainable development” (Kirkby et al., 1995, [quoted in Langhelle, 2000]). Immediately following the definition given above, the Commission identifies two key elements (not so often quoted): “It [that is, sustainable development] contains within it two key concepts:

- the concept of ‘needs’, in particular the essential needs of the world’s poor to which overriding priority should be given; and

¹⁸ In the next sub-section a third reason will be given.

¹⁹ It may be surprising to some readers that analysing at too low a level may lead to *overestimation* of impacts (apart from obvious chances for underestimation) as feedbacks may not be traced (Lakshmanan et al., 2001).

²⁰ See for instance Vogtländer (2002, p. 25): “Only when [emission] norms are set for the whole world, will problems like ‘export of environmental problems’ and ‘levelling the commercial playing field’ be resolved definitely.”

²¹ However, at below global levels the matching of Triple E scales will be a point of concern (Musters et al., 1998; Boisvert et al., 1998).

²² Compare Rischard (2002) for an analysis showing the need for improving the currently ineffective and slow global-problem solving on pressing global issues.

- the idea of limitations imposed by the state of technology and social organisation on the environment's ability to meet present and future needs.”

The notion that sustainability is about finding a good mix of economic development, environmental quality and social quality has widespread approval (Daly and Cobb, 1989; Elkington, 1997; Munasinghe, 2001; GRI, 2002; Pope et al., 2004 among others). There is, however, wide difference in the emphases given, and the social dimension in particular is treated quite differently by various authors. The interpretation of this thesis which, following Brundtland, focuses on extreme poverty and basic needs may show a ‘bias’, that some regard as common to the economist’s perspective²³. However, this interpretation can also be seen as choosing ‘minimum’ areas of concern whose importance seems undisputed.

The sustainability terminology, which has gained widest acceptance most quickly among business managers and consultants is Elkington’s ‘Triple P’ bottomline of “people, planet and profit” (Elkington, 1997). Elkington mainly addresses businesses, and argues that if corporations want to become sustainable, they should improve performance in the three areas of people, planet and profit. However, it seems that ‘people’ has a broad connotation, and the interpretation of ‘profit’ may be somewhat narrow. When we discuss here the three elements of sustainability as noted above, we will sometimes use the term ‘Triple E’ – paraphrasing Elkington – because it is conveniently short. To reiterate, Triple E stands for: 1) Economic development, 2) Extreme poverty²⁴ and 3) Environmental degradation. Evaluation in a sustainability context has to show Triple E impacts.

2.3.2.4b Triple E relations: separate concerns or closely related concerns?

The introduction of separate Triple E elements into a sustainability evaluation is tricky, however, as confusion is likely, unless the interrelation between the Triple E elements is clarified.²⁵ In the sustainability context there is obviously an interest in economic development that can be sustained over generations, but how do extreme poverty and environmental degradation relate to it? Are they simply ‘added concerns’ or are they strongly dependent upon each other?

Figure 2.2 illustrates the basic options. The Triple E elements could be unconnected (I) and could therefore warrant interest in their own right. Conversely, they can be very strongly connected (IV): an interest in one necessarily leads to interest in the others. All three elements can be somewhat related (III), or the situation can be different for each link (II).

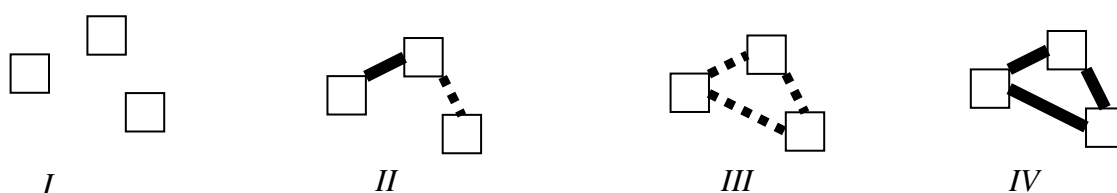


Figure 2.2: Four basic options of connections between the Triple E elements

²³ As Serageldin (1993) argues. See Munasinghe (1993), Pezzey (1992) and Hanley (2000) for a number of economists' views on sustainability. See Cernea (1993) and Rees (1993) for examples respectively of a sociologist and ecologist's perspective.

²⁴ The focus on the extreme poor rather than people in general provides a clear link with Rawlsian justice ideas (see for instance Langhelle, 2000).

²⁵ This situation mirrors the confusion around the Triple P approach to business performance.

As will be seen in chapters 3 and 4, this issue is crucial for the purpose of evaluation, because, if the relation is very strong as in situation IV (and partly in situation II), then there is really only one factor or aspect to be evaluated; and evaluating all three as separate concerns would easily lead to a form of double counting (Vogtländer, 2002).

2.3.2.4c Triple E – partly elements in their own right

The first thing to be recognised is that, at least partly, concern about extreme poverty and environmental degradation are part of the quest for sustainable development in their own right, quite distinct from economic development. Both extreme poverty and environmental degradation are then seen as separate ‘performance areas’ whether for the global system as a whole or for (e.g.) multinational enterprises.

To illustrate this point, we can observe that many people care about whales and elephants without the animals being an economic resource to them, and many think that nature should be respected, regardless of its economic use. Likewise, it is commonly believed that the extreme poverty of fellow human beings should be abated, without necessarily taking steps towards a positive impact on further economic development. (A case in point is fundraising in the West to combat African famine.)

2.3.2.4d The three relations between the Triple E elements

However, it is obvious from the sustainability literature that the Triple E elements are not only considered in their own right. On the contrary, the sustainability debate is strongly focussed on the interrelatedness of the Triple Es, and the consequences of this (WCED, 1987; Pezzey, 1992, p. ix; Jepma and Munasinghe, 1998; and for instance, Krabbe 2001; Turner et al., 2000; Munasinghe et al., 2003). The three relations are briefly considered below.

For the link between *economic development and environmental degradation*,²⁶ two aspects are worth mentioning. First, it may be obvious that economic development means the use of environmental resources. Economic development implies the transformation of resources into products and services; production processes require energy and produce waste. Due to economic globalisation, these simple and direct links between economic development and environmental degradation may no longer always be directly visible, because production activities for one good may occur in many different places worldwide. In this case the visibility of the link at different scale levels is problematic,²⁷ but the nature or the mechanism of the relationship is not.²⁸

Second, a major aspect relevant for this relation is the notion of unexpected and highly unwelcome limitations on economic development being posed by environmental limits. Brundtland discusses the limitations of the environment to support present and future needs. Sustainability literature widely shows that recent economic development has had a very strong impact on the environment, and several authors predict that – on the basis of past performance, current trends and estimates of environmental frontiers or feedbacks – this hampering of (further) economic development may occur on a much wider scale in the future. (Clayton and Radcliffe, 1996; Jacobs, 1991; Meadows, 1972).

²⁶ See Cleveland et al., 2001 for a useful introduction to the theme.

²⁷ For instance, Hart argues that to some extent more stringent environmental standards in developed countries have caused a movement of production processes not meeting these standards to developing countries at the expense of the environment in these countries (Hart, 1997, p. 68). Compare also Beladi and Frasca (1996).

²⁸ As for instance Siemons (2002) notices, this also implies that many environmental issues are already incorporated into the monetary economy. As will be discussed in part two, this fact in turn implies that – to avoid double counting – combined monetised assessment and physical/environmental assessment cannot take place without an analysis of the interrelations.

Again, causality here as such need not be too complex, as can be seen in many examples of the ‘tragedy of the commons’ type.²⁹ The cumulative size of many small-scale economic choices becomes too large and leads to over-exploitation of environmental resources (Ophuls and Boyan, 1992; Gordon, 1998; Schultz and Holbrook, 1999). For this type of problem adopting a higher spatial or organisational level of analysis or using a longer time perspective is enough to clarify the simple causal relationship.³⁰

However, causality can be complex too, if one for instance addresses the problem of greenhouse gases (Van den Bergh, 2004).³¹ This causality may then be uncovered by analysing developments over longer periods of time, by examining impacts on higher spatial levels, by finding the mechanisms of environmental change, or by a combination of these, and in the end, even this may not be sufficient to depict the relations.³²

The relationship between *Extreme poverty and environmental degradation* is often very sharp and clear on a local or regional scale. Extreme poverty, or “short-term survival pressure”, is often the basis for strong population growth, while strong population growth leads to pressure on the environment. The literature mentions, for instance, heightened pressure nowadays towards deforestation for energy, with general disregard for the longer term consequences for water regulation. Wood shortages also lead to the burning of dung for fuel, which, according to Hart, is “one of the greatest – and least well known – environmental hazards in the world today” (Hart, 1997, p. 69).

For evaluation purposes we can emphasise that the first situation here rather resembles the former link where a more distant perspective may clarify the relation; whereas the burning dung for fuel case shows that extra knowledge about problems of differing complexity may be involved. Both examples also show that, although causality may be clear, this has little bearing on the availability or viability of alternative courses of action.

The link between *Extreme poverty and economic development* is really about the relation of poverty of some people and the development of others. So, in a way, it is about interaction within the economic system. For expositional purposes it may be useful to distinguish the ‘market economy’ from the ‘subsistence economy’ (Hart, 1997).³³

²⁹ See for instance Lundgren and Burnet (1999) and references therein for an introduction and some recent debate – included is also a short commentary by the author of the influential 1968 article ‘The Tragedy of the Commons’: Garrett Hardin. See Gordon (1998) for a brief discussion of an example of the tragedy of the commons type: the example of the North American codfish.

³⁰ Formulated the other way around: causality may be blurred when a low level and a short-term view is adopted in evaluation.

³¹ Compare Aguilera-Klink et al. (2000) for the subtle socio-economic complexities in defining and experiencing water scarcity due to various economic developments.

³² A complete account of these possibilities is not appropriate to this thesis, but it may be fruitful to specify a number of important problems. According to Ayres (2001), the most immediate problems where the environmental limits hamper economic development are the supply of water (both for irrigation and drinking in cities in several regions in Asia) and wood (for fuel and charcoal production in Africa and Asia). The shortage of wood leads to uncontrolled deforestation, which in turn leads to soil erosion and increasingly catastrophic floods. Furthermore, Ayres sees a major problem in the reduction of toxic-waste assimilative capacity of the earth (due to topsoil loss and degradation and loss of biodiversity) while the demand for this ‘service’ of the environment is rising. The last problem posing binding limits is climate warming and its associated dangers.

³³ The subsistence economy is not necessarily a situation with extreme poverty. In this distinction the market economy consists of developed and emerging economies. The survival economy is the traditional, village-based way of life in rural areas of developing countries. According to Hart, about one-sixth of the world’s population (one billion people) live in developed countries of the market economy. They account for more than 75% of the

Concerning direct links, one can think of the lack of well defined or distributed property rights in the subsistence economy (which gives room to economic exploitation by the market economy) when for instance indigenous people in rainforests do not have effective ownership, and hence no power to stop forest exploitation.³⁴ One can also think of the ‘outsourcing’ or shift of production processes to low-wage countries. This outsourcing may directly involve extremely poor people in the production process. The impact this will have on the directly involved will probably be a relief of poverty in the short-run. The longer term impact of outsourcing depends (among other things) on the terms of involvement, the continuity of the labour demand, its impact on the former – possibly subsistence – economic structure of the community, and the families involved.

Infrastructure development is worth mentioning with regard to indirect links between economic development and extremely poor people, that is, infrastructure to support exploitation of woods or power generation activity to facilitate mineral extraction. The impact of these indirect links can go both ways: it can offer more opportunities to reduce poverty (sometimes for a short time only: see Jepma, 1995), or it can disrupt ecosystems upon which the survival economy relied, and thus lead to more poverty (Van Soest, 1998; Hart, 1997). As may be obvious from the above, the link between economic development and extreme poverty clearly involves justice issues.³⁵

2.3.2.4e *Triple E relations: facts and values*

Finally, it should be pointed out that the causality may be quite complex between empirical matters and value judgement. If an empirical analysis indicates that the tropical rainforest, or the codfish or the oil reserves will disappear within two decades, some people may find this very disturbing; others may not care.

Or, and this seems to be the more important case, citizens and policy makers may have no available alternative actions, at least in the short-run. The analysis may be clear on a global or other high spatial or temporal level, but on a micro level economic actors may not have much room for alternative actions, even if they are (made) aware of the relationships.

Formulated more generally – even if the empirical causality is clear – this fact in no way guarantees that the behaviour leading to it can be changed. Naturally there will be an interaction between knowledge of facts and values, but empirical clarity does not readily imply consensus about values.³⁶

2.3.2.4f *Conclusions about Triple E relations*

The interrelation of Triple E elements was briefly touched upon above, and now three conclusions may be drawn:

First of all, we have seen that Triple E elements are partly important in their own right. Either

world’s energy and resource consumption and bulk of waste creation. The survival economy part of the world is occupied by about half the world’s population – three billion people (Hart, 1997).

³⁴ Ayres sees the equitable distribution among the world’s inhabitants of tradeable emission rights as a possible solution in the area of assimilation of carbon (Ayres, 2001).

³⁵ Compare Fusco Girard and De Toro (2001, p. 402). “Social integrated evaluation processes are a good exercise to overcome our post-modern culture and identify new values and new priorities, in order to pay attention to those who have no voice, to poor people and to marginal people, and then to give sustainable development a ‘strong’ meaning, linked to unsolved justice problems.”

³⁶ In the scientific arena this is illustrated by the debate on notions of strong and weak sustainability. Strong sustainability takes the position that critical parts of natural capital should not be depleted. Weak sustainability allows substitution between natural capital and man-made and human capital (Pearce and Atkinson, 1998). Scientists may differ about which parts of natural capital they deem ‘critical’.

there is no relationship or the question of their relationship is not seen as relevant (see figure 2.2 - I).

Second, some of the links between Triple E elements are really rather direct and simple in terms of the mechanism involved.

- In some cases the links are direct and simple, and visible or clear to everyone involved.
- In other cases these direct and simple relations are not always visible on a local scale or low spatial level,³⁷ but may rather easily become clear when a (somewhat distant) perspective is adopted. This state of affairs strengthens the importance of evaluation at different spatial/organisational levels, and of the inclusion of the global level. Likewise in another set of cases these direct and simple relations are not always visible from a short-term perspective, but may rather easily become clear when a long time frame is considered. Figure 2.3 illustrates this situation. With this type of link one should be aware that, although clear links may be identified at high spatial levels or from longer time perspectives, this fact as such does not imply the existence of an agreement on possible mechanisms of change. This fact underlines the importance of ‘economic realism’ on a micro level, as the ‘distant’ understanding has to relate with the real-life mechanisms involved. It has to relate with the alternatives for actions that are available or absent at the micro level.

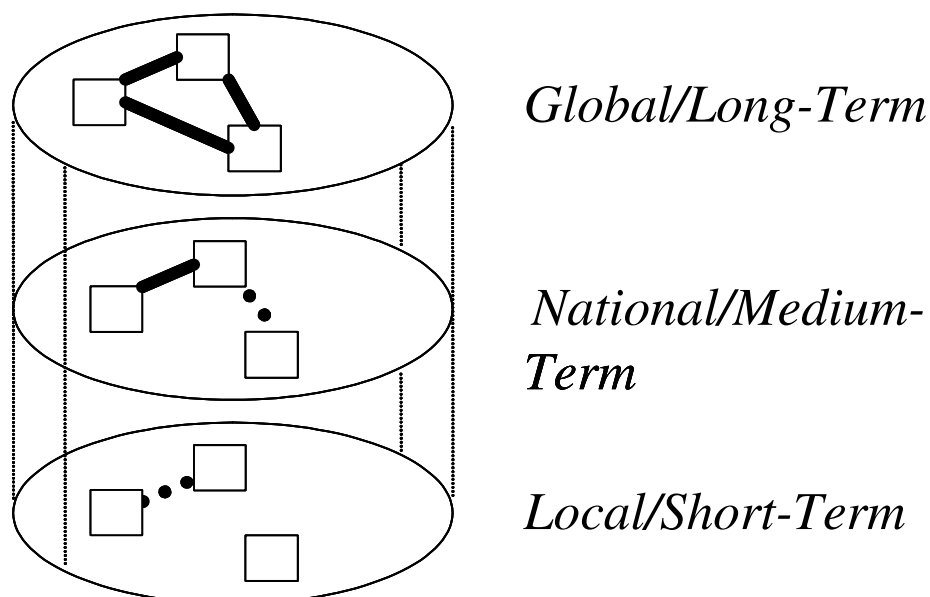


Figure 2.3: Illustration of how a higher spatial level and a longer term perspective may give a clearer view on the interrelatedness of Triple E elements.

Third, some of the links are complex and uncertain. Limited knowledge may indicate important risks in the interrelatedness of Triple E elements, or positive feedback relations, but the relationship remains complex and uncertain. This type of relation may occur in both the socio-economic and the ecological-environmental systems³⁸. (Figure 2.2. III).

³⁷ Although of course at regional or local levels we may see *some* strong links; and when we think of traffic congestion or noise, the links may be especially strong at the local level.

³⁸ Several authors believe that system thinking may be helpful as tools of analysis (Clayton and Radcliffe, 1996). System thinking would stress that social-economic and environmental-ecological systems are complex adaptive systems that are living and learning. These systems have numerous thresholds, feed-back structures, capacities to accrue and interpret information (Schütz, 2000). Over and above that, the interrelations between the several co-

At this stage it is sufficient to acknowledge that evaluation may on the one hand help to better understand the links; it may increase existing knowledge. Conversely, in as much as substantial uncertainty and complexity remains, there may be no easy way of objectively presenting the essentials of these links, the more so as the analyst and evaluation stakeholders alike participate in this open system and are creatively involved in forming and defining it (Schütz, 2000).

2.3.2.5 Stakeholder participation (f4)

2.3.2.5a Two reasons for stakeholder participation

The last key element central to the sustainability context of evaluation is the need for broad stakeholder involvement in decision-making on sustainability.³⁹ There are two reasons for this need.

First, as an old Dutch saying goes: ‘two know more than one’. Around complex and uncertain issues –such as finding routes towards a more sustainable situation (Pearce, 1998a, p. 69)⁴⁰ – as much knowledge and experience as possible should be activated.⁴¹ Naturally time and budget constraints will impose limits on this process.

Second, commitment may also be relevant, even if the route to greater sustainable development in a certain field is not particularly complex from a knowledge perspective. Decision-makers may have difficulty attaining enough commitment and (public and political) recognition of the problems and facts (Daly and Cobb Jr., 1989, p. 355). For many political decisions, commitment by stakeholders is deemed essential in itself as a logical part of democratic decision-making. Furthermore, in many situations stakeholder commitment is required, because action will be impossible or severely frustrated without it. Stakeholder involvement in the decision-making and evaluation process may then be crucial for establishing this commitment.⁴²

Therefore, both for reasons of activating knowledge and for assuring commitment, evaluation in a sustainability context should allow for stakeholder participation.

2.3.2.5b Various groups of stakeholders

As previously seen, the judgement perspective in evaluation often has evaluators who are somewhat distant from that which is being evaluated. Meanwhile, our interest here is in decision-making that is open and accountable to different stakeholders. The relationship cursorily examined above between evaluator and evaluated can be seen as one of the many evaluator-stakeholder relationships. Numerous different stakeholder groups may be relevant to a given project evaluation. Many project evaluation cases in recent years indicate an increased role

evolutionary systems in the world are very complex. However, we need not assess at this stage how far system thinking can help us understand these complexities.

³⁹ See for instance Hisschemöller et al., 2001; Ward et al., 2003; and Finco and Nijkamp (2001, p. 301) who see a “close interplay of all stakeholders” as the key for success in urban sustainability – worldwide.

⁴⁰ The lack of ‘a global government’ is a complicating factor in this, as many problems would be easier to handle if there was one (Tinbergen, 1970).

⁴¹ As Munda (1996, p. 166) expresses it when discussing CBA and Integrated Environmental Assessment: “Because of the deep uncertainties in evaluation methods, it is a case of ‘post-normal science’. In such cases, the traditional subject-specialty expertise is inadequate for quality peer review. Quality assurance therefore requires ‘extended peer communities’.”

⁴² In environmental science the three concepts of feasibility, effectiveness and acceptability are sometimes used. Project proposals may be (technically) feasible, they may be effective (when implemented they reach their goal), but acceptability (to those affected) may be lacking.

for the general (lay) public as judges of merit or worth, that is, of non-experts with little time or inclination to scrutinise closely; this new role seems to be part of a general trend in the relationship between science and society (Nowotny et al., 2001).

Rossi et al. (1999) distinguish as many as nine stakeholder groups even without attempting to be complete:

- 1) policy makers and decision-makers
- 2) program or project sponsors
- 3) evaluation sponsors
- 4) target participants or beneficiaries of the project
- 5) management involved in the project
- 6) personnel involved in the project
- 7) competitors for project funds
- 8) contextual stakeholders in the environment of the project (e.g., citizen groups, public officials)
- 9) the evaluation and research community.

For defining stakeholder groups one can turn to the management literature also (see for instance Henriques and Sadorky [1999] for another list of coincidentally also nine stakeholder groups).

These stakeholder groups may directly participate in the evaluation from the start, but they may also become interested in the evaluation process along the way or in its results at the end. Furthermore, the entire list as such is a list of ‘potential’ participants or ‘potential’ audiences; they need not all be actively relevant in a specific evaluation.

However, it is noteworthy that long lists of stakeholders are standard to many types of evaluation. The sustainability context will thus often be even more complicated. Consequently, the greater the number of stakeholders involved, or the more stakeholder decision-makers think they should be made accountable to, the greater the conflicts of interest between them will be; the greater the differences in (cultural) value systems, and the greater the potential mistrust and occurrence of strategic behaviour. In the sustainability context even the simple identification of stakeholders may prove to be complicated.

2.3.2.5c Minimum requirements for stakeholder participation

Evaluation literature in recent years has focussed largely on the relationship between evaluator and stakeholders. There seems to be a general trend towards more stakeholder participation in evaluation (Guba and Lincoln, 1989; Renn et al., 1993; Patton, 1997; Voogd, 2001). If one considers the enormous diversity among stakeholders and the number of people potentially involved, then to specify active participation in general is not a simple task (Rauschmayer, 1999; Gregory, 2000; Hisschemöller et al., 2001). The converse is true, as it is quite often impossible to specify in advance who will be interested, especially since involvement will also depend on evaluation outcomes and dissemination of findings.

However, at this stage it is not the nature and specifics of these relationships which need attention. Important is the fact that the process or results of any given case of project evaluation may arouse the interest of a wide range of stakeholders who differ markedly among themselves. They will differ firstly in their perception of the problem, and secondly they will differ in background, training, character, intellect, and so on. As a minimum requirement then, one very important aspect of project evaluation is that its results and methods be easily understood.⁴³

⁴³ See Lichfield (2001) who observes the development of a common language in participatory evaluation as one of the main implications. See also Niculae and French (2003) for development of evaluation tools in the e-democracy context.

One element that may enhance this understanding is the inclusion of extra scale or organisational levels in the evaluation to show outcomes relevant for different stakeholders. Another element that may help stakeholder understanding is standardisation of methods. More elements may be relevant (discussed in the following section) but it may be clear that without comprehensibility to stakeholders of both methods and results, there seems to be little room for any meaningful participation that goes beyond the involvement of experts.⁴⁴ Evaluations in the sustainability context should be credible to stakeholders (Chelimsky, 1997) and the possibility to understand outcomes should therefore be maximised.⁴⁵

2.3.2.5d Conclusion

The ideal project evaluation tool should allow stakeholder involvement, which may have several different forms. The involvement may take the form of very active participation, but it may also be more passive as, for example, listening audiences, which can nevertheless be a serious political power. The clarifying judgement that evaluation usually offers to experts should *also* be offered to non-expert stakeholders, meaning that evaluation results and methods (as a minimum requirement) should be easy to understand by non-experts.

2.4 Synthesis of the essential characteristics

2.4.1 Four key items

We have thus far identified several important characteristics that an ideal project evaluation tool should have in the accountable-sustainability context. This section will present a short overview and summary of the elements, both the general characteristics (a to e) and the sustainability-related characteristics (f1 to f4); see figure 2.4 below. In brackets – below the elements – the implications for the project evaluation tool are briefly stated. We have already noticed that the different elements are related, and they often strengthen or sharpen each other. In figure 2.4 the major strengthening and sharpening relations are shown. The arrows indicate when separate elements add to the same key item. If two elements are on the same row they imply a strengthening or sharpening relation, which implies their easy ordering as a single element.

⁴⁴ Compare Ashby (1980, p. 1180): “How can participants all have equal access to the necessary data? And how can they be educated – without the education being rejected as propaganda – to understand the data if it is disclosed to them?” On the value of expert models versus public participation in integrated assessments, see also Hisschemöller et al., 2001.

⁴⁵ Voogd 2004, p.227: ‘...evaluation...should always be judged as a partisan activity. This implies that transparency is very important, often a fundamental weakness of formal evaluation methods.’

General characteristics	→	Key items of accountable evaluation in a sustainability context:	←	Sustainability related characteristics
Economic impacts and more [d] (comparable)	→	I Triple E assessment	←	Triple E assessment [f3] (increase understanding of impacts, both separately and in their interrelatedness; show impacts comparably; analyse at global and lower spatial levels; relate to real-life economic mechanisms; show limited knowledge where relevant; values and facts may diverge)
		II At global and decision-making levels	←	Global (and local) perspective [f2] (analyse global and other spatial levels)
Time dimension [c] (show impacts over time)	→	III Over (a long) time	←	Long-term view [f1] (handle long-term impacts)
Accountable and open to stakeholders [e] (comprehensible, standardised, more spatial levels)	→	IV Which increases understanding of all (also about limitations)	←	Stakeholder participation [f4] (comprehensible, standardised, more spatial levels, worldwide involvement challenge, wide value and interest differences)
Ex ante [a] (uncertainty inevitable)	→			
Large and complex: formal project evaluation [b] (increase understanding and show limitations)	→			

Figure 2.4: Key items of accountable evaluation in a sustainability context

In the middle column of figure 2.4, four key items of accountable evaluation in a sustainability context are identified. They can be combined into one sentence:

The key items of accountable evaluation in a sustainability context are Triple E assessment at global and DM levels over (a long) time, which increases everyone's understanding (also about limitations).

I Triple E assessment – The aim here is to use the judgement perspective in evaluation. Deriving criteria for judgement is then central and critical. The elements ‘economic impacts and more’ and ‘Triple E assessment’ provide more content to the criteria. They overlap: the first is more open and the second is more specific and defines the minimum aspects to be covered. The first Key element can then read: the evaluation facilitates Triple E assessment. If one prefers to allow for clear judgement and to understand the links between these impacts, then these assessments or measurements of impacts should of course be as comparable as possible across different dimensions.

II At global and decision-making levels – The characteristic ‘global perspective’ adds further

content to the type of judgement being aimed at, as it is additional to the most relevant decision-making level(s). The global level can also be interpreted as an increasingly common impact level, as projects tend to have more impacts across the globe.

III Over (a long) time – Here the element ‘long-term view’ strengthens the ‘time dimension’ element.

IV Which increases understanding of all (also about limitations) – The evaluation should increase the understanding of all parties concerned.

Not only should experts or decision-makers have increased understanding, but broad ranges of stakeholders should too. The ideal project evaluation tool will allow for fruitful stakeholder participation. The elements ‘accountable and open to stakeholders’ and ‘stakeholder participation’ strengthen each other and lead to this extra challenge for the evaluation tool. The element ‘large and complex: formal project evaluation’ specifies the general background. The increased understanding relates to the limitations as well. Evaluation in the context of accounting for sustainability will have various limitations and uncertainties that should be clarified. The elements ‘ex ante’ and sometimes ‘complexly related systems’ – provide a general warning about possibilities and limitations: ex ante assessments of complex systems are performed about which there is limited knowledge. This general warning is greatly strengthened by the stakeholder participation elements.

Accountability and sustainability = sustainability (for short)

Although accountability and sustainability are distinct areas, we have seen that the sustainability context is closely connected to accountability to stakeholders. This thesis will often discuss the evaluation context of striving for accountability to stakeholders while assessing projects with Triple E long-term and global impacts. One might regard this context as the accountable-sustainability context, or the context of accounting for sustainability. In order to allow for succinct formulation, this wordy terminology is shortened to: the sustainability context.

2.4.2 More on limitations, judgement and uncertainty

2.4.2.1 Limitations and judgement

Four key items were identified in the discussion above that are important for a project evaluation tool used in a sustainability context of decision-making. In order to increase the understanding of the key items, we notice that, generally speaking, these four items can relate to two central concerns: limitations and judgement.

Every evaluation technique and every estimate of impacts has its limitations, but as seen above, evaluation in the accountable-sustainability context may have more than the average number of limitations. These should be (actively) clarified, as it is at least as important to know what is neglected as it is to know what is taken into account (Fischhoff, 1977).

Relating this discussion back to the three evaluation perspectives – in section 2.2 – it may be clear then that the evaluation context at hand is still judgement-oriented, but that already at this stage there may be some doubt about whether the upper parts of the shaded triangle of figure 2.1 can be attained.

It may now be fruitful to discuss in greater detail the different types of uncertainty that different knowledge elements within an evaluation may exhibit.

2.4.2.2 Uncertainty

Evaluation in the sustainability context and uncertainty as such seem to be closely connected (Lempert et al., 1996; Howarth, 1995).⁴⁶ The extensive literature about uncertainty and choice in general yields various classifications of types of uncertainty (Luini, 1999; Beroggi, 1999). For our purposes it is important to outline four areas of knowledge, namely:

- Awareness of certainty
- Awareness of soft uncertainty
- Awareness of hard uncertainty
- Unawareness of ignorance.

The terminology largely follows Vercelli (1999).⁴⁷ First, there is certain knowledge, about which little needs to be said, except that the evaluation process may *establish* what is certain,⁴⁸ and that evaluation will involve a *structuring* of certain knowledge.⁴⁹ Soft uncertainty is defined as the area of uncertainty in a familiar world (theoretically, roulette games or horse races function as examples). The area of soft uncertainty partly consists of the area which in formal decision theory is called risk: uncertainty with known probabilities. However, it may also include risk in a more real life setting. Risk in a real life setting will be frequently somewhat more loosely defined and may involve several probabilities.⁵⁰ In the sustainability context a complete empirical risk assessment may typically require several steps, which may necessitate estimates of probability.⁵¹

However, lack of certainty not only relates to empirical matters as such, but also to *beliefs* about empirical matters (Vicelli, 1999, p. 238). Different stakeholders may have different beliefs; experts and non-experts may differ (Pollak, 1998), but experts may also disagree among themselves (OECD, 1983). Because of the importance of beliefs about empirical matters, it seems useful to define risk in a more general and less formal way by making it refer to a situation in which there are either known probabilities or where there is a ‘familiar world’. Familiar is defined here as a situation where there is substantial knowledge about the mechanisms involved.

Hard uncertainty thus refers to knowledge areas where there are either unknown probabilities or in which the mechanisms involved are unfamiliar.⁵² Still, in this uncertainty area, there is awareness. Decision-makers and stakeholders know that they have incomplete knowledge.

The final category that Vercelli distinguishes is unawareness of ignorance. This category seems to be especially useful if the evaluation process is considered. Different stakeholders may have different levels of unawareness and may be unaware of different areas, and

⁴⁶ Compare a quote of Funtowicz and Ravetz in 1991 (quoted in Froger and Munda, 1998, p. 173): “ the facts are uncertain, values in dispute, stakes high, and decision urgent. In this way it is ‘soft’ scientific information which serves as inputs to the ‘hard’ policy decisions on many important environmental issues.”

⁴⁷ However, our definition of soft uncertainty is less formal than Vercelli’s.

⁴⁸ It need not be obvious at the start, as evaluation may bring awareness of this certain knowledge.

⁴⁹ It is not a mere collection of certainties; they are structured, which is the reason why in figure 2.5 we have given this area some structure too.

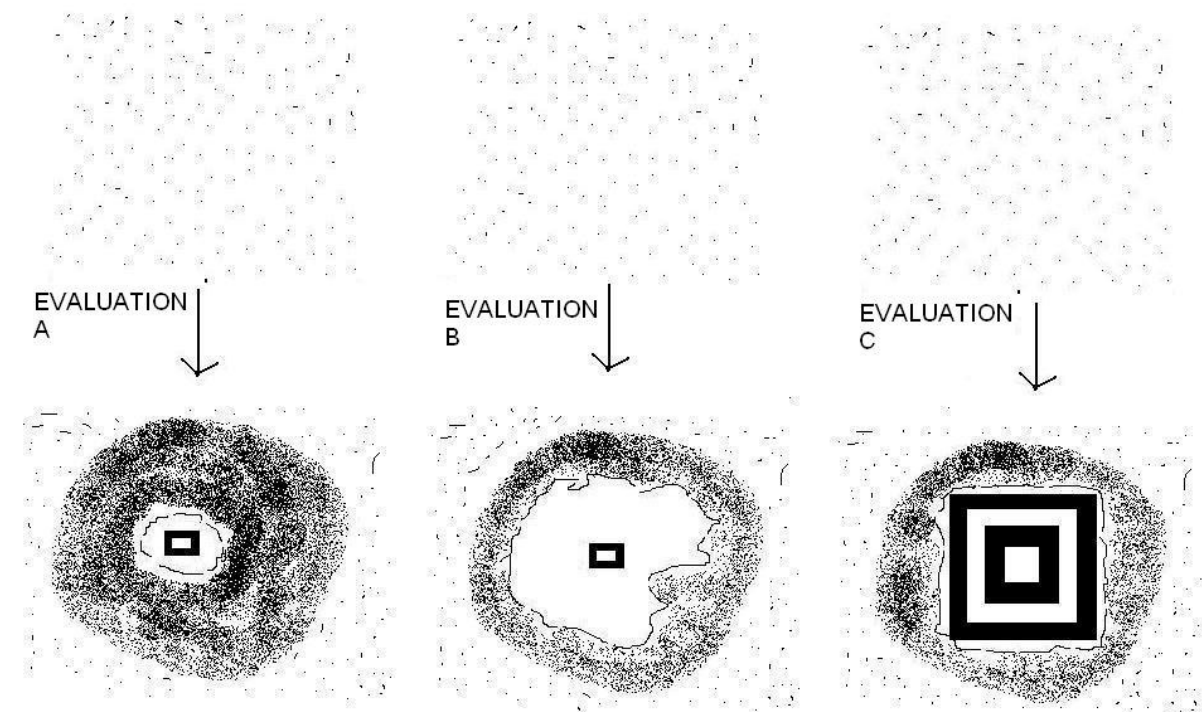
⁵⁰ E.g., Covello and Merkhofer (1993) define risk as a characteristic of a situation or action with more than one possible outcome; it is unknown which outcome will occur and at least one outcome is undesired.

⁵¹ Covello and Merkhofer (1993) in this respect make use of the ‘risk-chain’, which analyses first the risk source release processes (e.g., the likelihood of radioactivity release from a nuclear power plant); second, the exposure processes (how many people live in the vicinity of the plant); and third, the consequence processes (what effects may be experienced by exposed individuals).

⁵² Naturally there may be a grey area between these categories. For instance, risk defined in a real life way (following Covello and Merkhofer, 1993) may refer to both soft uncertainty or hard uncertainty, mostly depending on the familiarity of the risks involved.

furthermore their (un)awareness may change as a result of the evaluation.

Figure 2-5 illustrates how the process of evaluation may result in different areas of certainty, uncertainty and unawareness. The initial state of affairs before the evaluation for a non-expert stakeholder may be one of complete unawareness of the impacts of a project – this is naturally an extreme case. Different evaluations may result in different situations. There may be a small certain knowledge area, another small area of soft uncertainty and quite a large area of hard uncertainty (the situation to the left in Evaluation A). There may be a small certain knowledge area and a small hard uncertainty area, but a relatively large soft uncertainty area (the situation in the middle in Evaluation B). Finally, the evaluation may result in quite a large area of certain knowledge (the situation to the right in Evaluation C). This last situation is of course in general the most preferable. However, looked upon solely from the perspective of limitations of evaluation in the sustainability context, the most crucial aspect here is the amount of clarity achieved among decision-makers and stakeholders about the size of the areas. The analyst evaluator has a crucial role to play in communicating these ‘results’.



Legend:



Figure 2.5: Illustration of how, from an initial state of unawareness, three different evaluations may result in various divisions of knowledge areas.

As will be seen below, the importance of these distinctions lies in the variety of possibilities the different situations give for useful value judgements.⁵³

⁵³ See for instance Lind (1995) who argues that CBA is not very useful in the global climate policy debate, as this context is one with a high degree of uncertainty. See also Lempert et al. (1996) who argue that in a ‘large

Knowing the limitations and possibilities of techniques and estimates of impacts is important to the evaluator, but it is perhaps even more important to the stakeholders who participate in the decision-making process (compare Van den Bergh, 2004).

2.5 Short summary of evaluation demands in a sustainability context

Figures 2-4 have presented a synthesis of the most important items for accountable evaluation in a sustainability context, and the section above has explicated the issue of limitations to judgement. Before proceeding, the findings thus far can be summarised. The evaluation tool used in the accountable-sustainability context should:

- provide **judgement** about impacts on economic development, environmental degradation, and extreme poverty at global and DM levels over (a long) time, which increases understanding of all (stakeholders)
- make **clear** what the **limitations** of the evaluation are.

Evaluation tools will differ as to how well they can meet these demands.⁵⁴ The following chapters will assess the relative merits of both CBA and MCA. Chapter 3 begins with CBA while MCA is discussed in chapter 4. In chapter 5 a mix (MCCBA) is developed, which will try to combine the best of both techniques.

uncertainty space' best-estimate approaches have limited use. Compare further Nicolae and French (2003) about different policy or decision-making contexts (chaos, complex, knowable, known) and the X-space: where there is little agreement on the categorisation of the context.

⁵⁴ The operational judgement concepts should be conceptually clear and should allow for straightforward measurement even over long periods of time and on a global scale. The combinations of judgement criteria for all three Triple E elements should be comparable qua concepts and measurements and should preferably have a standardised measurement. One might be tempted to operationalise these demands into formal criteria and follow a multi-criteria approach for determining the best suitable tool. But this approach is not followed here as the readability would suffer quite seriously. The reader might compare Covello and Merkhofer, 1993, p. 239-265. They specify 19 criteria for judging evaluation methods, structured in a multicriteria tree. These are split into two groups: internal and external (resembling our expert and non-expert stakeholders). The internal group has three sub-groups: logical soundness, completeness and accuracy. The external group also has three sub-groups: acceptability, practicality and effectiveness. See also Smith (1986).

PART TWO – THEORY

CBA, MCA AND THEIR INTEGRATION IN A SUSTAINABILITY CONTEXT

CHAPTER 3: CBA IN A SUSTAINABILITY CONTEXT

3.1 Introduction

3.1.1 *The discussion of CBA*

Cost-Benefit Analysis (CBA) is an evaluation method that has theoretically largely been developed within economic science. It compares current and future costs and benefits of projects or policy alternatives in monetary terms. In principle, it can be applied to any resource allocation in an economy and some even claim that “there is no problem, public or personal, to which its broad ideas could not be applied” (Layard and Glaister, 1994, p. 1).

There is a long list of books stretching over several decades in which it is explained how to perform CBA, and which theoretical concepts are, or should be, followed.¹ Contrary to what some readers might expect, the literature is by no means unanimous. The major issues identified as relevant in the context of accounting for sustainability are and have been discussed continuously, although the ‘sustainability terminology’ in use nowadays may not have been used then (see table Table 3.1).

Table 3.1 Aspects of accountable-sustainability context explicitly addressed by different CBA authors of textbooks (although terminology may differ)

Accountable-sustainability aspects	Authors
Environmental issues	Abelson, 1979; Hanley and Spash, 1993; Johansson, 1993.
Distribution of impacts (especially in underdeveloped countries)	Dasgupta, et al., 1972; Squire and Van der Tak, 1975; Brent, 1996, 1998.
Clarification of CBA to non-experts	Sassone and Schaffer, 1978; Mishan, 1975; Boardman et al., 2001.

Obviously this debate is not limited to books and textbooks; the list of articles discussing major sustainability issues is even longer, as will be seen throughout this chapter. The reader is forewarned: debate is at times fierce.² Critics argue for instance that CBA is ‘stupid’ (Richardson, 2000) or only ‘a technician’s guess’ (Junger, 1979). Scientifically-oriented defenders usually argue that critics do not recognise the limitations and possibilities of CBA correctly (e.g., Rosenthal and Nelson, 1992). Interestingly, the practice of CBA seems to have had much of its own development only loosely connected with scientific debate (Krutilla, 1981; Zerbe, 1998; Adler and Posner, 1999).

Still, the literature portrays a clear enough picture as to what CBA is, how it is applied, and its major strengths and weaknesses.

The discussion here will assess the merits of CBA in a sustainability context and thus requires

¹ Harberger, 1972; Dasgupta, et al., 1972; Squire and Van der Tak, 1975; Mishan, 1975; Sugden and Williams, 1978; Sassone and Schaffer, 1978; Ray, 1984; Hanley and Spash, 1993; Johansson, 1993; Zerbe and Dively, 1994; Layard and Glaister, 1994; Nas, 1996; Brent, 1996, 1998; Boardman et al., 2001.

² Wildavsky introduced the two extreme positions in 1966 (!): “One can view cost-benefit analysis as anything from an infallible means of reaching the new Utopia to a waste of resources in attempting to measure the unmeasurable” (p. 293).

addressing fundamental issues.³ As seen in chapter 2, the sustainability context takes one to the essential preconditions and extreme outcomes of the economic system: global and long-term Triple E impacts. Therefore, the discussion of CBA has to be fundamental as well; even more so, since the debate about assessment tools in the sustainability context has revealed much confusion about these fundamental issues.⁴

However, there is danger in following a thorough-going approach, as in practice, any specific evaluation will naturally be limited in scope. It will focus on particular impacts and concerns or on small changes within only segments of the economic system. Since this chapter is general, our discussion of issues and problems of CBA could potentially suggest a fundamental problem with CBA and thus undermine its practical use altogether. The same may hold *mutatis mutandis* for the discussion of MCA in chapter 4.

However, the intention of these chapters is by no means to refute CBA or MCA to the extent that their practical value is ignored.

We can nevertheless be aware that there is a fundamental problem for example, with using a screwdriver to nail something down or using a hammer to screw something down. In other words, in practice, any evaluation tool will have limitations, depending on what it is to be used for. The reason for addressing these problems is to clarify a method's potential practical use and its limitations. Our key concern here is to track the extent to which CBA is useful in a sustainability context (compare Farrow and Toman, 1999).

Before outlining a detailed assessment of CBA, it is necessary to elucidate part of CBA's background in economic science. As will be seen later, the economic background is by no means without complications and one may even challenge the often made 'claim' that this scientific background is the foundation of CBA. Still, a textbook-based explication may put forward an initial idea of CBA's strengths and weaknesses; it may put the more fundamental criticism into a proper perspective and introduce the importance of economic market related causality as an important building block of CBA.

3.1.2 General background of CBA in economic science

CBA textbooks diverge as to whether a formal or informal approach in presenting the background is taken (compare for instance Johansson, 1993 with Boardman et al., 2001). A more formal presentation requires involvement of the reader in technical, graphical and mathematical formulation. Although there is no objection as such to this type of approach, the value for the purpose of this thesis will be either rather limited or negative, depending on the reader. Because the essentials needing to be understood are non-technical, it seems reasonable to follow the informal approach (compare section 1.4).

To grasp the general background of CBA in economic science, the following topics have to be discussed (Johansson, 1993):

- a) the behaviour of individual households and individual firms - the notions of

³ Compare the discussion of House (2000) on limitations to CBA, which easily addresses World Bank CBA practice as well as fundamental issues on justice, fact-value distinctions and rationality.

⁴ One reason for this confusion is the difficulty of multi-disciplinary debate: understanding a 'soft' social science from a 'hard' scientific perspective.

individuals or individual households choosing to consume different goods and maximising utility (subject to budget constraints) and the notion of firms being able to produce various (amounts of) goods and maximising profit (using labour and capital inputs)

- b) competitive markets and their possibilities to achieve a general equilibrium that is socially optimal to some extent (strong or weak Pareto optimality)⁵
- c) CBA as a tool for evaluating welfare changes in the real world where the existence of public goods, externalities and various forms of market failure may prevent the market mechanism from being a socially optimal allocation mechanism⁶

Ad a The behaviour of individuals and firms

Individuals or individual households are assumed to be maximising utility. On consumer markets individuals demand certain goods – at different prices they demand different quantities of these goods, depending on the individually derived utility it offers them. They are furthermore supplying labour and receive an income in exchange.⁷ The amount of labour they offer and the wage or income they want to receive in return also depends on the utility they derive from various options. The income received of course gives a budget constraint for the demand for consumption goods. The savings from income may supply capital to firms.

Individual private firms – ultimately owned by households – are producing consumer goods, and have several possibilities to produce more or less and switch between different types of goods. They use labour and capital for the production process, and may use them in various amounts. Firms try to maximise profits, which is the difference between their sales revenue and their wage and capital costs.

Notice that it is no coincidence that this discussion begins with individuals: “We assume that, as a rule, each individual is the best judge of his own welfare and chooses to maximise his welfare” (Ng, 1983, p. 12).

Ad b The market mechanism

The supply and demand for goods, labour and capital are confronted with each other in competitive markets. The issue of perfect competition in particular has to be addressed, as in the multi-disciplinary sustainability debate (Ring et al., 1999), confusion continues about the merits of the market mechanism and economists’ or CBA-analysts’ relation to it. The ideal of a perfectly competitive market is important because welfare economics asserts that the more reality meets the ideal, the more market values can be trusted as indicators of ‘social optimality’ or social welfare. But what is perfect competition? First, we will list the requirements of *pure* competition (Blaug, 1985, p. 594):

- the market has only small firms and consumers that cannot individually influence prices and furthermore
- all prices of homogeneous products and production factors are uniform throughout the economy.⁸

⁵ The ‘weak’ Pareto criterion indicates that a policy change is desirable if everyone is made better off. The ‘strong’ Pareto criterion asserts that a change is desirable if at least someone is made better off while no one is made worse off. Needless to say, perhaps in a real life setting the weak ‘criterion’ is indeed rather weak while the strong criterion is in fact weak too.

⁶ The notion of a social welfare function which can be maximised under various constraints and which provides further possibilities for ranking different situations (especially those with both winners and losers) than the Pareto optimality criterion provides is not discussed here, as it would take us too far at this moment.

⁷ Government taxes may have to be paid from this income.

⁸ Note that no spatial dimension is involved here; markets are just markets, not markets somewhere.

These are necessary conditions but not however sufficient for optimality. For competition to be *perfect*, more conditions have to be met:⁹

- all production factors must be perfectly mobile to prevent more than normal profits¹⁰ and
- all economic agents must have perfect knowledge of available alternatives.

Under the ideal circumstances of perfect competition the market mechanism assures that the various so-called ‘marginal conditions’ that economic science has developed for optimal allocation of resources do occur,¹¹ and that they are ‘socially optimal’ to some extent. In the equilibrium reached on a specific market, the individual marginal cost of the supply equals the individual marginal benefit of the demand, and there is no divergence between the marginal ‘social’ costs and ‘social’ benefits on the one hand, and the aggregation of individual costs and benefits as reflected in the aggregated supply and demand curves on the other. Put another way, under ideal conditions the many individual judgements of consumers and producers result in market outcomes that can be seen as socially quite optimal. The notion of Pareto optimality is often used in this respect. The market equilibrium is Pareto optimal: it is “impossible to improve the situation (welfare/profits) of one agent [person or household or firm] without worsening the situation of some other agent(s)” (Johansson, 1993, p. 11).

Ad c CBA as a tool for evaluating welfare changes

The aim of CBA is to determine the welfare impact of a certain change by assessing the impact the change has had on the welfare of all individual consumers, which can then be measured – and valued – by the change in prices and quantities in markets. CBA uses monetary valuation of impacts¹² and costs¹³ and adds them up.¹⁴ A positive sign does not directly indicate that the project or policy change is Pareto optimal, as there may be both winners and losers involved.

To achieve a more practically relevant criterion for social optimality than Pareto optimality, CBA theory utilises the Kaldor-Hicks criterion, or *potential* Pareto improvement (PPI): “a policy should be adopted if and only if those who will gain *could* compensate those who will lose and still be better off” (Boardman et al., 2001, p. 29). The actual compensation of winners to losers need not take place; compensation should be possible in principle.

Determining net benefits or net costs and determining whether compensation is possible in principle is greatly helped through market related information; the more money values reflect welfare properly, the more suitable is the CBA analyst’s toolbox. On the other hand, the more perfect markets are, the less need there is for evaluation on a project basis.¹⁵

The importance of the perfectly competitive market mechanism to CBA lies in the causality and the related transmission of welfare impacts through the economic system. Actions and

⁹ Various authors have added to this list the condition that production factors should be privately owned (see Nas, 1996, p. 19) but this condition seems unnecessary: the factors of production may be collectively owned too (Blaug, 1985).

¹⁰ That is, there should be no entry barriers.

¹¹ For optimal exchange, production, output composition, intensity of factor use, and inter-temporal optimality of these see Blaug, 1985, p. 593.

¹² In terms of willingness to pay.

¹³ In terms of so-called ‘opportunity costs’ – see below.

¹⁴ Using discounting to bring costs and benefits spread over time to one aggregate value (see below).

¹⁵ In as much as it concerns non-private CBA of course, as private CBA (firms evaluating projects on their private benefits and costs) will remain relevant.

impacts are causally linked. If a producer finds a way of producing more efficiently, competition assures that prices will be lowered, that consumers will benefit through lower prices. This in turn leaves them with more money to spend on these or other goods, which may either increase prices or increase the output of these goods, which will lead to more or dearer labour being asked for the production of these goods, etc. Although the impacts can ultimately be manifold and complex, and CBA takes the (final) welfare change of individuals/consumers as the best point of measurement, perfect competition may greatly reduce the workload of welfare measurement. As perfect competition assures that the initial effect in the first step of the causal chain (the first-order project impact) reaches the final individual consumers without distortions, welfare measurement can be restricted to these so-called direct effects.¹⁶

The economic theory that acts as a background is a-spatial: the market actors can be anywhere in the world. It may be clear then that economic theory backing up CBA easily supports the global level mentality of the sustainability context¹⁷. However, in CBA practice the question of geographical scale often arises, that is, the difficulty that often part of the impacts are outside the area deemed most relevant – by stakeholders or decision-makers – for analysis; e.g., international spillover effects in a national analysis. This complicates the CBA evaluation.

However, for some impacts even ideally, a perfectly competitive market system might have problems reflecting welfare impacts correctly. Economic literature has identified three issues. Two are particularly relevant to the sustainability context: 1) the existence of externalities or more generally non-market interdependence of welfare and 2) the provision of public goods. The last is of a more general character: 3) the market system may fail to be non-perfect or non-pure.

Externalities and non-market interdependence of welfare

Externalities refer to impacts that production or consumption can have on the welfare of others – other than those involved in the market transactions – without them being priced through the market. Classic examples are environmental effects from production such as pollution and noise; but the benefits from the discovery of knowledge or the making of inventions may also be cases. Externalities are therefore highly relevant in the sustainability context. The debate around the Coase theorem has shown that the detrimental consequences for social welfare can be ‘repaired’ by bargaining between the causers and the affected, but only if transaction costs are reasonably low, property rights are well defined, and preferably only a few parties are involved (Aivazian and Callen, 2003).

More generally an externality involves what Blaug calls ‘non-market interdependence’. This type of interdependence is relevant on the consumption side too when for instance the market value is caused by a preference to ‘keep up with the Joneses’. In such cases simultaneous reduction of the consumption of all consumers may not influence overall welfare levels, while a comparable reduction of only part of the consumers would have reduced it.

The provision of public goods

The second source that may create a divergence between private and social costs – and thus

¹⁶ Formulated differently, one can measure *either* direct *or* indirect effects *or* – in the case of market imperfections direct plus *additional* indirect effects. See Rouwendal and Verhoef, 2003; Bröcker, 2003; Elhorst et al., 2004; see also section 3.2.3.1c.

¹⁷ Although one might also argue that economic theory has underestimated or neglected the problems of the aggregation process (to the global level); see Van der Veen and Otter, 2002.

limit the social optimality of the market mechanism – are public goods. Most goods sold on markets are private goods; they can be consumed individually. However, for some goods consumption is necessarily ‘joint and equal’ (Blaug, 1985, p. 596). National defence is often used as an example; flood control is another. If it is provided it is provided to the whole nation and everyone benefits. There are very few examples of pure public goods, but several have some degree of public-ness because they are either non-excludable in their provision or there is non-rivalness in their consumption.¹⁸ The difficulty it provides for market functioning is that individually everyone will then prefer to evade payment (i.e., a prisoner’s dilemma).

Market failure or real world markets

CBA is active in the real world in which the markets are not perfect and most CBAs intend to measure national rather than global net benefits. In this real world setting the competition is less than perfect and there are varying degrees of what economists call ‘market failure’.¹⁹ Market failure may refer to economies of scale leading to monopolistic firms, entry barriers to certain industries, price inequalities throughout the economy, and very imperfect information.²⁰

Generally speaking, CBA attempts to combine outcomes from the real world competitive markets with additional monetary estimates of welfare impacts seen as not being accurately reflected in market interactions.^{21 22}

3.1.3 Two-steps

This explication of CBA’s background in economic science has given a first impression of the CBA evaluation style. The remainder of this chapter will assess how well CBA meets the demands of the accountable-sustainability context – as specified in chapter 2 and will follow a two-step approach.

In step 1 (section 3.2) the stages of CBA are discussed. CBA’s performance in the sustainability context will be assessed here, provided that it is clear-cut and can be handled without lengthy

¹⁸ Non-excludability – It is available to all if it is available to any. Non-rivalness – consumption by one does not impede consumption by others. See Garrod and Willis, 1999 (p. 17-21) for a classification of goods (following Winch) on the basis of three criteria: opportunity costs of consumption, property rights of the producer and property rights of the consumer.

¹⁹ E.g., Boadway and Wildasin (1984); see Zerbe and McCurdy (1999) for a critical assessment of this concept.

²⁰ Especially the last condition makes quite it clear that the analyst will rarely meet completely ideal conditions, as there would be little work remaining if this condition were met. Even more disturbing perhaps is that empirical work has found that a hierarchy of consumers seems to exist with varying – high and low – degrees of reception, understanding and use of consumer information (Hyman and Shingler, 1999).

²¹ In the real world setting CBA often partly uses direct *and* partly indirect effects, as relying on only one type of impacts may give an incomplete picture. Making use of both types of effects leads to a need for extra attention in avoiding double counting. See Rouwendal and Verhoef (2003) versus Oosterhaven and Elhorst (2003), and Elhorst et al. (2004).

²² At this point the theory of second best does come in. According to some, the most disturbing objection against using market values for social welfare estimates is Lipsey and Lancaster’s theory of second best. Their theory postulates that if there are at least two markets in which the optimum conditions are not satisfied, then there is no way of telling whether breaking down imperfection in only one market will improve welfare. As CBA practice is typically part of the ‘piecemeal welfare economics’ founded on the assumption that such a focus on one market is useful, their theory is rather disturbing, at least theoretically. In this chapter it is argued that the ultimate basis of CBA is not theoretical, but rather the consensus of stakeholders that everyday economic reality does matter, which seems to make this theory somewhat less disturbing.

discussion. In step 2 (sections 3.3 to 3.5) topics of problematical CBA performance warranting greater discussion will be handled.

3.2 The stages of CBA

3.2.1 Introduction

In a discussion of the structure of CBA, largely following the stages as explained by Hanley and Spash (1993), one can detect seven stages:²³

- Stage one: Define the project (or policy) and impact population
- Stage two: Identify project impacts
- Stage three: Quantify relevant impacts physically
- Stage four: Value relevant effects monetarily
- Stage five: Discount costs and benefits
- Stage six: Apply the Net Present Value test
- Stage seven: Perform sensitivity analysis

A short general description of every stage will first be given on the basis of existing literature; secondly, an assessment of the pros and cons of the stage will be provided, with special attention given to the demands of the sustainability context.

3.2.2 Stage one: Define the project (or policy) and impact population

3.2.2.1 General description

3.2.2.1a Specification of project alternatives

The first major concern of CBA is to define which action or investment is to be evaluated; that is, project alternatives need to be specified. There will usually be several different options for tackling a problem, and the evaluation aims to choose among them. A very important aspect of CBA is that defining alternatives will also mean defining a baseline (or ‘without’) alternative of the evaluation. A baseline alternative is often a ‘business-as-usual’ scenario to which the performance of the other alternative(s) is compared. The without-alternative should not be regarded as a do-nothing alternative, unless doing nothing is a real option.²⁴ Specification of this alternative in CBA is taken very seriously, as it determines many of the outcomes and provides a basis for distinguishing impacts that may be different before and after the project from changes that are *attributable* to the project (Zerbe and Diveley, 1994, p. 3; Jorgenson, 1993). Specification of the without-alternative is moreover closely related to CBA’s fierce attempts to avoid ‘double counting’ of effects (Sassone and Schaffer, 1978, p. 32; Eijgenraam et al., 2000).

Definition of the project alternatives determines the boundaries of the analysis. Preferably the

²³ Hanley and Spash actually identify eight by splitting stage two in two parts.

²⁴ Notice that defining alternatives often requires building scenarios for future behaviour and future circumstances. Which behaviour and which circumstances comprise the without-alternative and which comprise the with-alternative(s) is open to discussion (Sassone and Schaffer, 1978). Compare Eijgenraam et al., 2000, who argue to make the without-alternative a ‘best alternative action’. Whether this is very practical in most cases is arguable, but it does highlight that to ‘do-nothing’ is usually not a realistic option.

analyst's utmost creative effort should go into finding and specifying alternatives. Although limiting the analysis may be a natural and sensible thing to do, in practice this does not make it very easy or uncontested. Much public discussion about CBA outcomes seems to focus on the alternatives taken into consideration, and over the course of the analysis analysts may decide that they should have considered other or more alternatives.²⁵ CBA literature yields little guidance however on how to deal with new alternatives that may emerge over the course of the evaluation.

3.2.2.1b Determining the scale of the analysis: The impact population

The second major concern at this stage is to consider the population of gainers and losers. Sometimes law will define this population, but more often some choices can be made here. CBA can be performed at a company, regional, national, or global level, to mention only a few possibilities. The decision-maker can play an important role in choosing a level, but if given the opportunity to decide, the analyst will not generally limit the analysis to a small spatial scale. As noted above, CBA and CBA theory start out from a-spatial markets, which could imply *global* markets. In practice the analyst should choose, in line with CBA textbook theory, the 'population area' as the area in which *important* impacts can be observed. Following this rule will often imply that the CBA analyst takes in a larger relevant area of analysis than a decision-maker will.²⁶ In general, choosing an impact population smaller than one in which impacts are observed is problematical for CBA, as the chain of passing impacts through the market mechanism is broken down. Such a choice could hamper the measurement of impacts on final consumers. But once the most appropriate impact population is selected, costs and benefits are normally analysed, regardless of to whom they accrue.²⁷

3.2.2.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.2.2a CBA literature silent on incorporation of newly emerging alternatives

Stage one makes clear that any CBA evaluation implies the comparison of alternatives. Judgement is therefore confined to the relative performance of a group of alternatives, which underlines the importance of choosing the alternatives. In a setting with many stakeholders the considered alternatives may always be contested, and alternatives may be 'brought in', or argued for, which will change the relative performances strongly. Although CBA literature acknowledges the importance of considered alternatives, there is scant information on how to generate alternatives or incorporate newly emerging ones.²⁸

3.2.2.2b Measurement at a global scale is natural to CBA: Other scales can be included.

The sustainability context demands clear measurement at global and other scale levels (Turner et al., 2000; Holmes et al., 2004; Grasso et al., 2003). The reason for this is, as was seen, that projects often have global impacts even if they seem only to be small-scale local actions. An essential and strong point of CBA, which is often highlighted, is that it does not automatically concede to a decision-maker's scale of analysis, e.g., government agency or private firm, but rather adopts the 'perspective of society' (Nas, 1996, p. 5). The general description of stage

²⁵ Hanley and Spash give an example of an analysis of the building of a new nuclear power station that spread over into an appraisal of UK energy policy, a choice that was later seen (by some) as unwarranted.

²⁶ In public policy there is the strong tendency to examine only costs and benefits from the perspective of the government authority: CBA requires the perspective of society in order to be adopted (Nas, 1996, p. 5). Compare also Oosterhaven (1982) for a combination of different spatial levels in CBA.

²⁷ Although, as will be seen, attention is often given to distributional impacts.

²⁸ The lack of attention to the generation and emergence of new alternatives is one reason why Richardson (2000) speaks of the 'stupidity' of CBA.

one showed that the notion of choosing a relevant spatial level or ‘impact-population’ level is natural to CBA. Analysing at lower levels than the impact level may raise extra problems for CBA, despite that it is often practiced; it may require that extra and different types of data be gathered.

3.2.3 Stage two: Identify project impacts

3.2.3.1 General description

3.2.3.1a How to identify impacts in a CBA

The judgement criterion of CBA is the impact of a project on social welfare or the utility levels of all the individuals in an impact population. In stage two CBA proceeds to identify all the project impacts and organise them properly.

To determine which of the impacts should be taken into account, and in what way, CBA literature uses its theoretical base of welfare economics. Utility, whether social or individual, is a central concept in welfare economics. Those impacts that influence peoples’ utility are accounted for in a CBA. Hanley and Spash remark on this central issue: “[A]ssume society is interested in maximising the weighted sum of utilities across its members. These utilities depend upon, amongst other variables, consumption levels of marketed and non-marketed goods... The aim of CBA is to select projects, which add to the total of social utility, by increasing the value of consumables and nice views by more than any associated depletion in the levels of other utility-generating goods. ... Thus what are counted as positive impacts [benefits] will either be increases in the quantity or quality of goods that generate positive utility or a reduction in the price at which they are supplied. What we count as costs will include any decreases in the quality or quantity of such goods, or increases in their price” (Hanley and Spash, 1993, p. 9-10). Although CBA literature is filled with the terminology of goods, prices and costs, it is not limited to merely tangible impacts; both tangible and intangible costs or benefits should be incorporated (Nas, 1996).

Stage four returns to this issue, but for now the process of distinguishing between types of costs and benefits and properly ‘organising’ them in a CBA should be clarified. Two conceptual elements are central to this process: opportunity costs and transfer payments (Mishan, 1975; Hanley and Spash, 1993; Nas, 1996; Scheffer and Sassone, 1978). Both are more easily understood if the general background of CBA in economic theory and more specifically, the notion of perfectly competitive markets, (see section 3.1.2) are kept in mind.

3.2.3.1b Identifying impacts: Opportunity costs

Opportunity costs are important when identifying costs and benefits in CBA. Two examples (Nas, 1996, p. 61) may help to explain them. If one considers expanding or demolishing a recreation centre, a CBA regards the historic costs of once building it as irrelevant to decisions upon a new allocation of resources. The alternative returns the resources would have generated now and in the future are what is relevant. Likewise the cost of labour involved in a certain project is not the wage the workers will receive, it is the output lost elsewhere in the economy as a result of workers being relocated to the current project. Therefore, the opportunity cost is what is relevant: the value the resource could have generated in its next best use (Nas, 1996, p. 61). The opportunity cost is usually regarded in terms of a specific alternative; usually it is the ‘without’ situation which specifies the opportunity cost (Mishan,

1975).²⁹

3.2.3.1c Identifying impacts: Transfer payments

Exclusion of transfer payments from CBA is also a very important step when identifying the costs and benefits in CBA. It is not so much pecuniary effects (or transfer payments) but real output effects that count. Real output effects are changes in physical production possibilities and the welfare effects that go with them. Examples of mere pecuniary effects that should not be taken into account as such are reductions of indirect tax revenues, additional unemployment benefits becoming payable, and increase in property values because of the project. Apart from specific situations, CBA treats these flows as mere redistributions of money within the population considered and no real consumption of resources is taking place. The exclusion of transfers is closely connected with CBA's attempts to avoid double counting of impacts.

3.2.3.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.3.2a Problems with identifying costs and benefits

This stage reveals essential reasoning being followed in CBA, which aims to identify costs and benefits properly, and is perhaps the most important and fundamental part of the conceptual clarity of CBA. Both in the case of opportunity costs and the case of transfer payments, notice that CBA has strict ideas on what should count and on how impacts should be measured, and it is hard to imagine how these ideas would have arisen if an understanding of the functioning of the market mechanism were not somehow involved.³⁰

Despite the fact that real markets are often far from perfect, CBA usually relies to some extent on the outcomes of the market mechanism but simultaneously CBA's utility base acknowledges the relevance of any impact, whether it is expressed through the market or not. In other words, identification of costs and benefits on the one hand may seem to relate to utility only while on the other hand, market related notions are seriously involved. Because of this two-sided situation, two related and rather fundamental aspects can be identified in this stage:³¹ the problem of social aggregation and the problem of broad utility versus narrow economics.

3.2.3.2b Fundamental problems of identifying costs and benefits: 'Social' aggregation

First of all there is a tension between utility to the individual, and social utility for a complete impact population. The quoted text above speaks of 'the weighted sum of utilities across its members' but it may prove to be a thorny if not impossible exercise to give this summation any operational content. This problem is of a quite general nature and not specifically confined to the context of accounting for sustainability, as can be seen from the extensive debate on the difficulties of constructing a social welfare function that is not a mere theoretical construct

²⁹ Opportunity costs naturally require an (often implicit) scale of reference (compare Hill, 1999).

³⁰ It is obvious that various stakeholders may feel that demolishing a recreation centre is a relevant cost as such, and that often 'mere' transfers are not seen as 'mere' redistribution but as real impact on social welfare, which may show that what to count as transfers and what not to count is not indisputable. Jones and Cullis (1996) for instance argue that prestige and power transfers from voters to politicians are transfers that should be included in CBA. Compare also Zerbe (1998, p. 420): "Costs and benefits are necessarily psychological constructs; they are a matter of subjective emotion."

³¹ Although one might argue that this topic should be treated in a later stage when measurement is discussed.

and the related problems of finding satisfactory social aggregation procedures.³² We will treat this subject in more detail in section 3.3. Furthermore, and more actively relevant in the specific context of sustainability, is the debate about distributional weights and income distribution, which will be discussed more deeply in section 3.4.

3.2.3.2c Fundamental problems of identifying costs and benefits: Relating 'broad' utility and narrow 'economics'

The second problematical aspect is that utility is a broad welfare concept. It is roughly synonymous to, say, happiness. Triple E aspects can easily fall within the scope of this broad concept, as long as they cause someone to become 'more or less happy' (Hanley and Spash, 1993, p. 10). However, notice that although CBA expresses impacts in monetary values, and although economics as a science is commonly associated with expertise on production or consumption activities and markets, the concept of utility hardly has anything monetary, everyday economic, financial or market related about it. Consequently, at this stage (and at stage four, discussed below) there may be a significant gap between the all-encompassing concept of utility and narrower 'economic, financial or market related' concepts. From the outset it creates a tension for any practical CBA to bridge the gap between the ideal and the operational measurement possibilities. The tension calls for clarification of the terminology concerning 'economic' and for further investigation into the subject. These issues will be addressed in section 3.3.

3.2.4 Stage three: Quantify relevant impacts physically

3.2.4.1 General description

This stage determines the physical amounts of costs and benefits and their distribution in time. Quantification of the elements deemed relevant can range from the very simple to the very difficult. Determining the amount of cement used in the construction of a bridge may be simple and straightforward. However, determining the local environmental impact may be far from easy, and by the same token, estimating the number of cars crossing the bridge ten years after its construction may also be difficult. CBA needs physical quantification of impacts before values can be placed on them. If possible, probabilities or – a range of – expected values may be determined for different impacts.

CBA theory as such gives little further guidance for this stage, except when discussing specific case studies. The reason for this lack of general guidance is probably because this stage is common to many types of evaluation. As Patton argues, evaluation is data-based, and formal judgement-oriented evaluation cannot function without systematic data collection (Patton, 1997, p. 25). This stage in the analysis may require development or use of variously complex models, which are often at the core of CBA, but they are often at the core of MCA as well (Merkhofer, 1987, p. 59).³³

³² See Layard and Walters (1994) for a discussion. One important focus in this debate is Arrow's impossibilities theorem (see for instance Yew-Kwang Ng [1983] for a short general description. Sen [1982]) and Arrow and Raynaud [1986] give some further interesting views).

³³ For instance the use of spatial economic general equilibrium models is sometimes proposed as the preferred way to generate CBA outcomes (Knaap and Oosterhaven, 2003).

3.2.4.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.4.2a Empirical facts about (Triple E) impacts, their (un)certainty and relations become clear

The earlier stages of CBA have provided some specific structuring and framing of the evaluation to guide stage three. But apart from this structuring, stage three could represent any type of evaluation. The generality of gathering empirical data should not suggest that this stage is somewhat unimportant. On the contrary, this stage will often ultimately yield the most important information of all.³⁴ In this stage the physical size of the Triple E impacts may become clear. The varying degrees of certainty attached to different physical impacts may be established³⁵ (compare figure 2.5) and the causality between the different impacts may be disentangled: some links may be simple and direct (sometimes especially at higher spatial levels) others may be complicated and fuzzy. Naturally beyond this stage, values – about the found empirical facts – become more prominent.

3.2.5 Stage four: Value relevant effects monetarily

3.2.5.1 General description

In CBA all impacts are valued in monetary terms. For traded goods and services market prices carry valuable information about relative scarcity. Through the price of materials, capital and labour, the related impacts can be valued directly (or with relative ease). According to Hanley and Spash, the analyst then has a further three tasks: 1) predict prices for value flows in the future, 2) correct market prices where necessary and 3) calculate prices where none exist.

Ad 1) prices should be projected in real terms: they are corrected for inflation so that price changes reflect real production possibility changes.

Ad 2) market prices in equilibrium should show equality of marginal social cost and marginal social benefit. Market failure, like imperfect competition (in its extreme form: monopolies) or government intervention can prevent this from happening. The analyst may then decide to work with so-called ‘shadow prices’; the analyst estimates prices that are supposed to better reflect true marginal social costs and benefits.

Ad 3) very often the analyst has to put monetary value on goods or services that are not traded in a market. Different techniques are available for this. There is intense debate about the applicability of these techniques in different settings however.³⁶ In the sustainability context the major debate is about the usefulness for assessing environmental impacts monetarily.

3.2.5.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.5.2a Market values are the basis of CBA valuations of impacts

CBA measurement aims at monetary valuation of all Triple E impacts. The starting point for this measurement is monetary market values. The values people hold are taken to be reflected in their behaviour when they buy groceries, book vacations, hire new personnel, or trade in factory supplies. The monetary values that the CBA analyst actually uses should reflect social welfare

³⁴ It is one of the ‘two arms of evaluation’; see Rossi, 1999, p. 21.

³⁵ Compare Gollier (2001) for a discussion of uncertainty and the merits and demerits of precaution in relation to CBA.

³⁶ See for instance Drèze and Stern, 1994.

changes. But as already mentioned when discussing stage two, there is tension between the broad welfare changes that may be identified and the narrower market related welfare changes that can be measured. Quite often the analyst will process or adapt market values to better reflect changes in social welfare. To understand the rationale and the possibilities for this procedure, the concept of welfare and its relation to market values will be discussed further in section 3.3.

3.2.5.2b Valuing environmental impacts often requires estimation of artificial market values

Valuing environmental impacts often requires calculating money values without using market related data. Section 3.5 will discuss assessment of environmental degradation impacts more deeply; the available techniques, their potential and the difficulties surrounding them are treated there.

3.2.6 Stage five: Discount costs and benefits

3.2.6.1 General description

3.2.6.1a Discounting is natural for market related impacts

If all cost and benefit flows in time are expressed in monetary terms then in CBA they are all converted to present values: an amount of money spent today that can be seen as the equivalent of the net costs spread over time, or an amount of money earned today that can be seen as the equivalent of the net benefits spread over time.

For a number of reasons, people value equal money values at different points in time differently. People equate a higher value to having €1000 today rather than having €1000 one year (or 10 years) from now. In CBA future money values, costs or benefits alike, are converted to present values, using a discount rate. If the discount rate is 4% per year, then €1000 today is valued equally to €1040 received a year from now. The higher the discount rate the lower future money values are valued relative to current money values, so a lower discount rate makes future benefits and costs more important; a higher one makes them less relevant.

The level of the discount rate(s) has several theoretical explanations: the level of risk, the amount of time preference and the (marginal) productivity of capital. A very important aspect to be noticed right away is that despite the often very technical discussions surrounding this issue,³⁷ discounting is a perfectly natural thing to most people in as much as it relates to economics (understood in the sense of everyday economic activity). Most people are familiar with borrowing and lending money at a certain interest rate, either as consumers or as producers. The time periods can in principle be quite long too, although in actual economic life longer term lending and concomitant interest lending rates are most common for time periods of 1 to 10 to 30 years. A time perspective of 100 years ahead, relevant to sustainability assessment is uncommon in most real life economic transactions.

3.2.6.1b Practical approach to choosing a discount rate: Government prescription

There has been and still is considerable debate about discounting in CBA (e.g., Moore et al., 2004). Authors who write about discounting in CBA, usually start from real life lending rates that are familiar to consumers or producers (Stiglitz, 1994) and then argue for a different one in public policy evaluation or in a sustainability context,³⁸ Stiglitz (1994) for instance argues

³⁷ Which exceed what the average stakeholder can follow; and even trained economists have difficulty convincing each other.

³⁸ A third approach attempts to establish a social rate of time preference.

that the discount rate may differ from project to project and that the choice of the proper rate should be made after a thorough analysis of, for instance, the distributional consequences of the project.³⁹ Depending on the outcomes of this analysis, discount rates may be set lower, equal, or higher than consumer or producer lending rates.

In the setting of public policy evaluation there seems to be consensus on adopting a somewhat lower discount rate than current long-term market interest rates (see for instance Quiggin, 1997; Nas, 1996). The logic behind this is easy to understand. As public policy often tries to redress some kind of market failure, it would be somewhat awkward to expect market returns from public policy measures.

Because choosing a discount rate is crucial to CBA, and because this choice is often debated, at the national level governments have increasingly *prescribed* the use of a specific discount rate.⁴⁰ These rates are often just below or at the lowest long-term market rates (e.g., long-term government bonds, see Layard and Glaister, 1994, p. 29 and Quiggin, 1997). Naturally this procedure of standardised government prescription of a discount rate strongly increases comparability of results across projects as well as transparency and accountability in CBA evaluation (Adler and Posner, 1999; compare Squire and Van der Tak, 1975, p. 9).

3.2.6.1c Arguments for lower discount rates in the sustainability context

The sustainability context, with its attention for future generations and long-term impacts, has further deepened the debate about a proper discount rate in CBA (Heal, 1997; Quiggin, 1997; Lind, 1997; Tóth, 2000; Goulder and Stavins, 2002). Some argue that the social discount rate should be declining⁴¹ and for instance Heal (1997, p. 7) argues that for very long-lived projects the discount rate may even be zero.

The efforts to argue for low discount rates should come as no surprise: from the (political-normative) perspective of taking long-term impacts seriously, there is a 'need' to find theoretical reasons for a low discount rate, as can be seen in figure 3.1 below. As mentioned in chapter 2, a long-term analysis of effects might require a 100-year period as a relevant period of analysis. If costs and benefits are to be taken seriously into consideration over such a long time period very low discount rates have to be used. Figure 3.1 shows that already with a discount rate of 4%, there is hardly any impact of a cost or benefit of €1000 after 100 years, and even after 25 years a €1000 cost or benefit will only have a discounted impact of some €400. With a 10% discount rate already after 50 years there is no longer any real effect.

³⁹ "These results may be frustrating to those who seek simple answers, but such are not to be found" (Stiglitz, 1994, p. 155).

⁴⁰ There is theoretical support for this going back to the work of Feldstein (see Zuidema, 1982).

⁴¹ For references see Tóth, 2000.

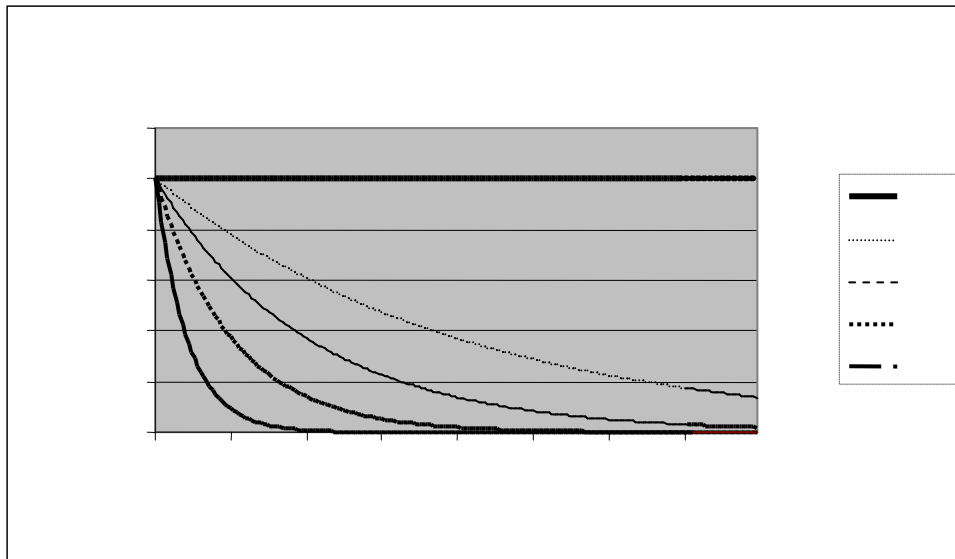


Figure 3.1: The impact of several discount rates on the long-term effects

For instance Sen (Nas, 1996, p. 140-141) distinguishes three arguments for the use of a social discount rate lower than that of the current generation in CBA. The *super responsibility argument* states that governments can be seen as having a special responsibility towards future generations. This issue was already touched upon above, when it was said that government has the perspective of looking beyond market realities; this argument is quite common for any public policy role of the government. The *dual-role argument* asserts that individuals as citizens may be more concerned about future generations than individuals as consumers of market goods only; this is an important aspect in the sustainability context⁴² and more of this will be seen when discussing the use of individual willingness to pay valuation of environmental degradation. The *isolation argument* states that individuals would participate in collective savings⁴³ by all, something that they would not do in isolation. This latter point touches upon the issue of the provision of public goods and the prisoner's dilemmas involved, or more generally, the publicness of certain goods or services. For such goods the market mechanism is not a good allocation mechanism and there is need for other mechanisms, e.g., political decision-making processes that facilitate expression of these collective action preferences.

3.2.6.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.6.2a The basis of discounting – alternative use – is not always meaningful

Having considered the general theoretical arguments about the proper discount rate, the crucial element needing emphasis in the sustainability context is that the basis for a *specifically determined* discount rate is the *reality* of alternative use. As long as discounting is restricted to market related impacts, the discount rate is a benchmark of performance in alternative use or

⁴² This has become a very popular notion in the current sustainability debate (e.g., Slocum, 2004).

⁴³ Savings or any other collective action. There may be a difference in our individual commitment to driving at low speeds to save energy, in isolation (knowing or guessing that everyone else will not slow down) and as part of a collective action (knowing or guessing that everyone else will slow down too).

uses.⁴⁴ Individuals, households and firms can spend their money – or more generally devote their productive resources – on the project under evaluation or they can spend it differently and then, perfectly realistically, expect a market rate of interest as a return.

Thus, the point to understand is that discounting is strongly related to the opportunity cost approach of CBA. Different projects and different types of costs and benefits have different alternative usages and therewith alternative returns – which implies different discount rates.

However, the familiarity with discounting is different for non-market related impacts; since they may occur in the environmental and poverty domains. Here the realism of alternative returns cannot be used, because for instance, the notion of switching environmental qualities around in various productive applications is very often meaningless. The argument for discounting non-market related effects is then most often defended because of – pure – time preference (Hanley and Spash, 1993). Pure time preference refers to the fact that people prefer to have high environmental quality now, rather than in the future. But how can a precise discount rate be determined from this general notion? There is no non-arbitrary way of finding such a discount rate as there is no real life point of reference for non-market related discounting that gives a clue on how high the discount rate should be. We can therefore conclude that the theoretical as well as intuitive support is quite different between discounting of market related impacts and discounting of other parts of Triple E impacts.

3.2.6.2b The discounting debate is really about limitations to non-market valuation in CBA

If, in a cost-benefit evaluation, long-term impacts should ‘automatically’ be taken into account then – as seen above in figure 3.1 – there is clearly a ‘need’ to use very low discount rates. However, adopting an ‘automatically’ low discount rate means departing from using interest rates that have some link with current real life interest rates based on alternative returns, and the intuitive understanding and experience that stakeholders have with them. In as much as costs are clearly market related this departure is a serious loss. For long-term environmental impacts, the outcomes with low discount rates may be beneficial to ‘sustainability’, or ‘future generations’ for that matter, but it seems to be a rather technical and confused way to settle a dispute,⁴⁵ and using the same low discount rate for economic development impacts may have undesirable consequences as it may favour projects with low returns (Lind, 1995).

The crucial point is that discounting market related costs and benefits has a different rationale than the discounting of non-market related values (Hueting, 1991). The issue of discounting is therefore closely connected to valuation of non-market related effects and cannot be resolved until a clear view is presented of the character and limitations of non-market valuation in CBA. The discounting issue seems to put extra demands on the (comparable) quality of monetary valuation of these effects. Further discussion of this issue of comparability will take place in section 3.3.

⁴⁴ Market related impacts can in principle be part of all three Triple E areas, although of course they will be most prominent in the economic development section.

⁴⁵ Compare Revesz, 1999 about the difference between discounting within a generation and discounting over more than one generation.

3.2.7 Stage six: Apply the Net Present Value test

3.2.7.1 General description

Stage six can be seen as the final result stage of the CBA. Once all impacts in monetary terms are calculated and a suitable discount rate is chosen, the net sum of costs and benefits can be calculated as a present value: that is, as a value as if it were a current cost or benefit. Preferably projects should have a positive Net Present Value (NPV). These projects can be seen as efficient in terms of their use of resources: benefits exceed costs and provide a return above the discount rate. If there is more than one alternative the alternative with the highest NPV is preferred.

The complexity of actually calculating net present values may be unappealing to many stakeholders. However, it should be emphasised that the general idea of projects being efficient or profitable over time is something wide ranges of stakeholders both understand and have practical experience with; this exercise, for instance, is no different from judging whether one can afford a certain mortgage loan for a house, or assessing the profitability of a business activity.

Although the Net Present Value is the most preferred criterion, other tests are possible (Zerbe and Dively, 1994, p. 194) and may be used in combination with the NPV. Internal rates of return can be compared for several project alternatives (the discount rate that would yield a NPV of zero). Furthermore, benefit-cost ratios may be determined (they should preferably be higher than one) but this procedure requires strict – and sometimes counterintuitive – definitions of costs and benefits to lead to comparable ratios.⁴⁶ Sometimes the payback period is proposed (the time needed for total benefits to exceed total costs), but this naturally involves rather limited judgement information.

At this stage CBA theory suggests that one might give different weights to different costs and benefits that accrue to certain groups, for instance to groups in extreme poverty. The marginal utility or value of income naturally differs tremendously between very rich and very poor people. Being willing to pay €1 for something may mean nothing to a rich person, while it may mean a one-week income for a very poor person. The debate about the treatment of these ‘distribution matters’ in CBA has been going on for years⁴⁷ and continues in the sustainability context, especially on the debate of costs and benefits of climate change (Azar, 1999; Johansson-Stenman, 2000; compare also Rietveld et al., 2002).

3.2.7.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.7.2a Calculation of NPVs without difficulty, interpretation requires caution

This stage is a logical conclusion to the other stages. Calculating the results has no further difficulty; the most important aspect now is how to interpret the results. Several yardsticks may be used, and as relative judgement criteria they may all be useful, although one might preferably not use them altogether. Different yardsticks naturally have different pros and cons. Ratios for instance are very sensitive to which elements are called costs and which are benefits: in this respect CBA literature usually prefers identifying negative impacts from a project as a negative benefit rather than a cost. This habit will result in better comparison

⁴⁶ Benefits are then usually defined as impacts from the projects (positive and negative!).

⁴⁷ See for instance Boadway, 1974.

possibilities for benefit/cost ratios but it is difficult to be entirely strict in practice.

Although theoretically the outcomes of a CBA are easily interpreted, in actual political decision-making the aspect of correct interpretation often remains open:⁴⁸ should any project with a positive net benefit be realised? This is best answered after the capabilities of CBA to capture social welfare at large are sorted out in section 3.3.

3.2.7.2b Distributional weights are sometimes considered

Likewise the treatment of how to handle extreme poverty impacts in CBA warrant separate attention. Should separate weights be attached to impacts affecting specific groups? We address this issue in section 3.4.

3.2.8 Stage seven: Perform sensitivity analysis

3.2.8.1 General description

The last stage of CBA is sensitivity analysis. The reason for conducting a sensitivity analysis is during the process of conducting a CBA, several assumptions and estimates of future impacts had to be made, while in the end one may not be entirely sure how appropriate their use or resulting outcomes are.

In a sensitivity analysis certain parameters or assumptions are changed and their impact on the net present value is analysed. In this way one can discover to which parameter changes the outcome is most sensitive and how this sensitivity differs across project alternatives. Sensitivity analysis is therefore closely connected with uncertainty and accountability.

Starting with accountability, the trouble with sensitivity analysis seems to be that from the analyst's perspective more often than not a sensitivity analysis is bound to be somewhat superfluous. An analyst has difficulty making the message of the sensitivity analysis something like: the main results can hardly be trusted. If the analyst thinks something is important⁴⁹ he or she would incorporate it in the analysis, for instance by specifying minimum and maximum likely estimates. Still, in a real life setting with many stakeholders, there may be no guarantee about the pure motives of the analyst, or about the depth of his or her 'paradigmatic fixations'. A sensitivity analysis gives the opportunity for different stakeholders to assess possible outcomes if certain assumptions are relaxed or certain estimates are changed.

Regarding uncertainty, the sensitivity analysis may be seen as the ideal place to explicitly address either soft uncertainty or hard uncertainty (see figure 2.5 in chapter 2). The task of the analyst here seems to be to reveal the areas of uncertainty and ignorance. Although one could argue that it is a difficult task to give much attention to issues about which little can be said, as they are uncertain, the analyst does have substantial responsibility here. The reason for this seems to be that quite often the major result of an evaluation is not in the quantitative conclusions but rather in the increased problem understanding. Because the evaluator is in a privileged position in this respect, he or she should therefore exert effort to communicate his understanding of the areas of uncertainty and their significance (Nas, 1996; Mishan, 1975). Naturally this part of the analysis will be less rigorous than the quantitative calculations and

⁴⁸ Compare for instance the fierce tone of Boardman et al. (2001) on page 15: "The appropriate criterion to use is the NPV rule. Other methods [like the IRR or B/C ratio] may give incorrect answers; the NPV rule does not" with the very modest tone on page 486: "One lesson is that *ex ante* CBA is difficult to do precisely."

⁴⁹ And, it may be added, has no strategic behaviour concerning the outcomes.

will be tailored for any specific evaluation.

3.2.8.2 Assessment of pros and cons in the context of accounting for sustainability

3.2.8.2a Sensitivity analysis may be very rewarding in a conflict setting

In the sustainability context two problems are especially relevant for sensitivity analysis. First, stakeholders involved can have widely different values and conflicting interests. Because of this fact the outcome of the sensitivity analysis may be the most important part of the analysis, as it is here that differences of opinion can be accommodated and the impact of different assumptions, views or estimates highlighted.

3.2.8.2b Uncertainty is not always best treated as footnotes to the main results: Uncertainty may turn out to be a central result

Second, there may be plenty of uncertainty in the analysis due to inherent complexity or simply because of the required international or global scale and long-term view of parts of the analysis. Uncertain and complex impacts may be hard to measure. The obvious danger is that in the debate or political process that uses the results, the quantified results may easily overshadow the less tangible results (Campen, 1986; Connor, 1999; Richardson, 2000).

Summing up, if the analyst finds – along the way in the process of evaluation – that important uncertainties remain, they should be central in the outcomes, and some effort is needed to present these uncertainties clearly. This presentation should naturally not comprise a long list of uncertainties. A long list would be very easy to make, as in ex ante evaluation hardly anything is really certain. An important function of the analysis should be to clear out unimportant uncertainties. What may remain is one or two important uncertain issues, or hard-to-handle conflicting valuations. The analyst can then use his or her increased understanding of these issues and share it with the decision-maker(s) and the various stakeholders.⁵⁰

3.2.9 At this stage what is clear about CBA's performance and what requires further examination?

In this section 3.2 the first step of CBA's performance in the sustainability context has been taken. Several conclusions have been drawn while other points were seen to require a more thorough discussion. Table 3.2 below illustrates the main issues and conclusions discussed above. The table shows which issues or conclusions are now clear and which are still unclear. Furthermore, for the unclear issues the table shows where these are further examined.

⁵⁰ To the judgement-oriented analyst, this state of affairs – at the end of an evaluation effort – is a somewhat disappointing and second best, albeit realistic, option.

Table 3.2 First conclusions about CBA's performance in the sustainability context

	Clear/unclear by now:		Further examined in:
	Clear	Unclear	
Stage one: Define the project (or policy) and impact population			
-CBA silent on incorporation of newly emerging alternatives	X		
-Measurement at global scale is natural to CBA; measurement at lower – explicitly defined scales – often practiced too	X		
Stage two: Identify project impacts			
-The issue of 'social' aggregation		X	3.4
-The issue of reconciling 'broad utility' identification options and 'narrow economic' measurement options ⁵¹		X	3.3
Stage three: Quantify relevant impacts physically			
-Empirical facts about Triple E impacts, their (un)certainly and relations become clear, just like the main areas of uncertainty	X		
Stage four: Value relevant effects monetarily			
-Market values are the basis of CBA valuations of impacts		X	3.3
-The more perfect competition is, the better are CBA's monetary measurement possibilities	X		
-Valuing environmental impacts often requires estimation of artificial market values		X	3.5
Stage five: Discount costs and benefits			
-Possibilities for discounting different Triple E impacts differ; market related impacts have a natural reference point of real alternative opportunities, which other impacts lack	X		
-The discounting debate is really about limitations to non-market valuation in CBA	X		
Stage six: Apply the Net Present Value test			
-Calculation of NPVs without difficulty, practical interpretation requires caution	X		3.4
-Distributional weights sometimes considered		X	
Stage seven: Perform sensitivity analysis			
-Sensitivity analysis may be very rewarding in a conflict setting	X		
-Uncertainty is not always treated best as footnotes to the main results; uncertainty may turn out to remain as a central result	X		

In the following three sections the topics of problematic CBA performance are discussed. Although they seem to cover different issues, in a way they can be seen as treating three different aspects of one central concern: to which extent is monetary measurement of broad welfare notions possible or meaningful.

The sections below will thus have a closer look at CBA's incorporation of broad welfare notions. The main problem of CBA is to reconcile its starting point of broad individual and social welfare notions with its measurement possibilities, which seem to be optimal when limited to market related welfare. This problem has several aspects to be treated subsequently.

First, discussion begins with a general assessment of CBAs of welfare notions and willingness to pay measurements (section 3.3). This general understanding will be deepened in sections 3.4 and 3.5, which both cover specific aspects of broad welfare notions: measuring distributional and

⁵¹ As the text has shown, this issue might also have been classified under stage four.

environmental impacts respectively (compare figure 3.3). These more specific types of impacts – environmental and distributional – have both been discussed at length in the CBA literature and are of particular relevance in the sustainability context. In recent years environmental impact measurement within CBA is used regularly and techniques are being developed further. Distributional impact measurement within CBA has been popular since the 1970s, but despite its technical simplicity, is not practiced much anymore. Discussion of this issue will, however, determine important conclusions.

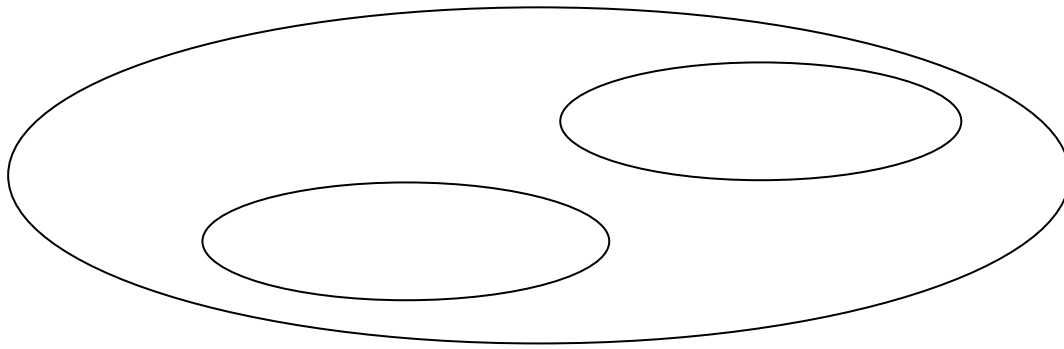


Figure 3.2: The relation of the different sections that follow

3.3 CBA and the incorporation of broad welfare notions

3.3.1 Introduction

Before attending to the general assessment of incorporating broad welfare notions through Willingness To Pay measurements, the different uses of the word ‘economic’ should be clarified. Box 3-1 defines three meanings of the word economic most relevant to the discussion: ‘everyday economics’, ‘economic science’ and ‘welfare economics’.

Box 3-1: Three relevant meanings of the word ‘economic’

In the discussion of CBA three meanings or usages of the word ‘economics’ may be distinguished:

- Everyday economics
Everyday usage economics relates to markets, to production and consumption of goods and services, and to money, budgets and financial aspects: ‘economic growth 1.5%’; ‘good economic performance of multinational’, ‘economic crisis hits ...’
- The science of economics
Economics can further refer to economics as a scientific discipline in a broad way: ‘Nobel prize for economics’; ‘economics different than physics or biology’; ‘economists’ belief that...’
- Welfare economics
Welfare economics is a special branch of economics, which has as its special concern the welfare of society and its individuals at large. It is part of ‘general’ economics as opposed to ‘business’ economics. Despite its highly theoretical nature, it is commonly presented as the basis of CBA*.

** In the concept of Total Economic Value, which, as will be seen in section 3.5, has become popular in the ecological economic literature, the word economic is probably also best understood as referring to welfare economics, although this remains puzzling.*

Welfare economics is often seen as providing the theoretical base for CBA; this basis provides CBA with the task to link broad and general welfare concepts without any operational content to market related, narrow economic value concepts with down-to-earth operational content. The central issue can now be formulated as the problem of relating welfare economics to everyday economics (see box 3-1).

Within the context of this thesis it is neither possible nor useful to give a complete overview of welfare economics, the debates surrounding its main concepts, and all the possible implications of these debates for the practice of CBA.⁵² The discussion here focuses on two crucial elements, the understanding of which is essential for a proper assessment of CBA’s usefulness in the accountable-sustainability context:

- The role of broad welfare notions
- The use of Willingness to Pay (WTP) as a shortcut to welfare assessment

⁵² The debate has a long history: see Porter (1995) and Boardman et al. (2001). The reader may also turn to Mishan (1981b) and Sen (1982) for a collection of rigorous articles related to welfare and utility, which are still remarkably relevant in the sustainability context.

In a nutshell the discussion runs as follows. The first important element is that welfare economics starts from broad notions of welfare. From this starting point, two directions are possible: one can either stick to the broadness – which may be realistic but seems to give little practical guidance for project evaluation – or one can opt for practicality.

Being more practical – the mainstream in CBA – leads to the second crucial element: the willingness of an individual to pay or accept money as a measurement tool for welfare changes. Using this measurement tool has several limitations. The contexts in which this measurement is most fruitful and those in which it seems less so will be specified with the help of the notion of the dual-preference structure: preferences that are held as ‘consumers’ differ from those that are held as ‘citizens’.

3.3.2 *Broad welfare notions*

3.3.2.1 General description, subjective (and mainstream) welfare economists

Discussion starts with ‘utility’, a central concept in welfare economics. Utility is a broad concept at the level of individuals. In welfare economics individuals are expected to maximise their utility. It often has a hedonistic flavour (seeking pleasure), but this need not be so. “The attainment of justice and the appreciation of art, as well as a new automobile, are all activities or goods that give rise to utility because they are things that are chosen” and utility is “a device to help explain what people choose” (Zerbe and Dively 1994, p. 70; Brekke et al., 1996, p. 315).⁵³ So utility, the starting point of CBA foundations, has nothing much everyday economic about it: ‘happiness’ or ‘feeling good’ probably come closest. People try to maximise happiness or feeling good.⁵⁴

Utility in welfare economics is considered to be an ordinal variable to any individual.⁵⁵ One cannot say “I am twice as happy with this as with that,” or “I am half as happy being in this situation as in that.” This would be cardinal utility. Ordinal utility means one can only rank alternatives: “I prefer this over that and that over something else.” Welfare theory assumes one can do that much. The priority of welfare economics is to assert something about welfare of groups of individuals, that is, of society at large. The concept of individual utility as such, however, has no likelihood for aggregation, as interpersonal comparisons of utility are not possible.⁵⁶ If the analyst sticks to ordinal utility the only things that can safely be said when evaluating alternatives is that welfare improves if it improves for every individual, or if at least one individual’s welfare improves without making anyone else worse off. A situation in which it is impossible to improve anyone’s utility without impairing someone else’s is called a Pareto optimum (Ng, 1983).

Within welfare economics two different groups of authors may now be distinguished, which for expositional purposes will be called the ‘subjective’ and the ‘mainstream’ theorists respectively (Blaug, 1985). To the mainstream theorists section 3.3.3 will turn; the subjectivist welfare economists are treated here.

⁵³ Sen (1985) has given an overview of the various interpretations of utility. Compare Islam and Clarke (2002) and Strijker and Deinum (1992) for discussion of some important – aggregate – measurement issues.

⁵⁴ One might also consider terms from psychology or sociology e.g., ‘quality of life’ or ‘well-being’.

⁵⁵ At least since Pareto at the beginning of the 20th century (Blaug, 1985, p. 586).

⁵⁶ In chapter 4 MCA takes a different position considering this impossibility point.

Subjective welfare economists

Pareto refused to take evaluation any further and analyse situations in which some people win and some lose in welfare. Part of the modern subjective welfare theoretical authors (Robbins, Hennisman, and more recently for instance Vollebergh, 1999) stay very close to accepting this ‘impossibility point’. They hold to the fact that “while it is far from easy to discover and evaluate the welfare experience of an individual, this is [even stronger so for analyses] like macro-economics and [...] the welfare of many individuals.” (Vollebergh, 1999, p. 74). Accepting this point is then closely linked to a particular idea of what economic science fundamentally is. “Economics is the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses” (Robbins, 1984).⁵⁷ To the evaluator this means that the judgement criterion should first be clarified externally, as economic science and welfare economics do not generate such a criterion by themselves (Vollebergh, 1999).

3.3.2.2 Assessment of pros and cons in the context of accounting for sustainability

3.3.2.2a Welfare economics, despite its merit of realism, cannot be taken as a ‘basis’ for a practical- oriented CBA

The subjectivist authors in welfare economics take as their starting point a broad, hard to grasp and personal concept of welfare, which many will agree is realistic. Welfare has a personal quality to it that is difficult to comprehend and largely beyond social comparison. However, for a specific project evaluation this starting point does not provide a practical judgement criterion, nor does it provide a specific procedure for deriving it. In this sense welfare economics is highly theoretical and little practical advice follows from it. The notion of the impossibility of interpersonal comparison of ‘utility,’ may be realistic, but not if sticking to such a welfare notion that is broad, personal and ordinal results in little useful things to be said about preferences in a social sense. A criterion for judgement has to be given prior to starting CBA or evaluation in general.

The major lesson therefore seems to be that ‘happiness’ of individuals, or any other broad and realistic welfare concept, is not easily related to measurable, operational welfare.⁵⁸ This lesson seems to be true in any evaluation context and is not specifically associated with sustainability.

A logical consequence of this ‘subjective school’ interpretation of welfare economics seems to be that there is no reason for Cost-Benefit Analysis to be a preferred technique on the basis of ‘economic science’ or ‘welfare economics’.⁵⁹ Any kind of evaluation technique, e.g., Multi-Criteria Analysis, that focuses on both pros and cons of an action to an impact population, may be said to be based on this subjective type of ‘welfare economics’.

⁵⁷ The reader may note that here, welfare economics is somewhat exaggerated: seen not as merely a branch of economic science, but rather as the (core) of the complete economic scientific field.

⁵⁸ As the Beatles insisted: “Money can’t buy you love.”

⁵⁹ Unless of course, as was seen above, in the everyday economic reality the markets would happen to be perfectly competitive, externalities would be negligible, the public-ness of goods would be small, and non-market interdependence would be largely absent.

3.3.3 Mainstream: Willingness to Pay (WTP) as a shortcut to welfare measurement

3.3.3.1 General description

The mainstream in welfare economics has not followed the subjective line of thought. Mainstream welfare economics has tried to establish a link between money values and utility, or put another way, between everyday economics and broad welfare notions.

The central linking device used is ‘willingness to pay’ or ‘willingness to accept payment’ in various forms. As it is impossible⁶⁰ to measure or observe utility directly, one can compare people’s willingness to pay (themselves) for changes that have a certain utility level to them, or accept payment (from others) as a compensation for a change. Welfare economics uses these two concepts: Willingness to Pay (WTP) and Willingness to Accept (WTA).⁶¹

One can appreciate the very straightforward simplicity of the procedure of asking for money values of welfare changes, which has some persuasive power to *clarify intensity of preferences*. But the major question remains: To which extent are WTP and WTA values generally applicable for measuring broad welfare changes?

3.3.3.2 Assessment of pros and cons in the context of accounting for sustainability

3.3.3.2a Monetary valuation of Triple E elements: Hesitation and context sensitivity

Mainstream welfare economists use WTP or WTA as a linking device between money and welfare changes. CBA uses data about preferences from two different kinds of contexts. First and foremost, CBA uses aggregate monetary WTP expressions from the marketplace – which may or may not be adjusted by the CBA analyst. These are revealed preferences: actual behaviour reveals them. Second, CBA may use WTP expressions that are explicitly requested through questionnaires. This latter type of WTP expressions, stated preferences, have become especially popular when assessing environmental impacts (see section 3.5).

That the WTP measuring rod of money has limitations as an expression device for welfare changes is not always acknowledged in CBA literature. WTP and WTA imply a narrowing down of the welfare concept, to changes that individuals are prepared to place money values on, or according to Pigou: that can be brought in relation to the measuring rod of money.

In the sustainability context two points concerning these limitations seem to be especially relevant:

- 1) hesitation to valuation and
- 2) context-sensitivity of valuation.

Ad 1) First, people may hesitate to bring certain things in relation to the measuring rod of money, because they feel it does not have meaning or is ethically unsound. Probably the best example of this is the monetary valuation of human life, but it is also often found in environmental valuation (Merkhofer, 1987; Revesz, 1999; Spash, 2000; Sunstein, 2004).

Ad 2) Second, and perhaps more generally, is that people give WTP expressions in a certain context. Use and usefulness of WTP expressions are limited by the context in which they were given or in which they had to be given.

⁶⁰ See Kapteyn (1985).

⁶¹ See for instance Clinch and Murphy (2001, p. 423) for a brief discussion; WTA should, according to some, theoretically diverge no more than 5%, whereas WTA is sometimes found to be four times greater than WTP for the same amenity.

The discussion below will be more specific about the contexts relevant for sustainability evaluation. The aim is not to make context sensitivity as such problematical. The aim is to clarify from which contexts useful and trustworthy WTP expressions can be derived for the evaluation purposes at hand.

To disentangle this complex discussion, a simple down-to-earth procedure is followed; it explores the objects of valuation in the separate Triple E areas. These objects require different framing aspects, or, one might say, around these objects different contexts automatically ‘emerge’. Three typical objects of valuation from the Triple E areas will be considered:

- a bottle of cola in the USA,
- a tree in the German Schwarzwald and
- a poor man in Liberia.

These objects may be a cause for welfare change, because some project might lead to more cola being available in the USA, or another project might want to protect German woods, while another might want to alleviate poverty in Liberia.

A bottle of cola in the USA

To start with the simple product, the bottle of cola, it may be observed that giving a WTP expression for such an object is very familiar to most people. Most people are usually involved in this type of setting in market related contexts: they determine or express their WTP, make the decision to buy – or not – and have the product for use afterwards – or do without it. So the possibility for individual use of the object of valuation and the familiarity with the mechanism involved gives confidence to the use of WTP values for the bottle of cola as an expression of importance attached to the bottle.

But despite this reason for confidence, it can be emphasised that such a WTP expression is not insensitive to context. Exactly the same bottles of cola – with exactly the same physical characteristics, coming from the same factory – may be very differently valued, depending on the brand given to them, the store selling them, the time they are sold, the advertisements made for them, the popularity of the product to other people, the information available about the method of production, etc.⁶²

A tree in the German Schwarzwald

To proceed to the environmental asset, a tree in the German Schwarzwald, observe that for valuation of this asset the valuation has to be framed, as it is not a typical WTP expression – unless one is active in the timber business. The tree can indeed be regarded as a source of timber, but also as a source of oxygen, as a provider of shade, as a home for birds, as prevention of erosion, as a component of a scenic environment, or as a monument of nature from the past – to mention a few possibilities (De Groot, 1992).

Ambiguity about the context leads to two points. First, before being able to give a WTP expression, framing the environmental valuation is necessary; this implies an analysis of the causality involved in the function of the asset and its interaction with other assets, and how it affects parts of welfare. Second, and in sharp contrast to the cola valuation, there is far more collective valuation relevant, as many of the functions can either only be enjoyed collectively or they hardly have an individual usage component (compare for instance Noël et al., 2000).

⁶² Using revealed WTP expressions in an evaluation may involve a shift in context, which may or may not be warranted, while in a setting of explicitly requesting WTP values, this context has to be ‘controlled’, to make it reflect the evaluation purpose.

A poor man in Liberia

Finally, a poor man in Liberia may be considered. How to value an alleviation of his poverty? Again framing is necessary, as valuation as such is not something most people are often involved in, and individual use is completely absent.

Valuation will differ according to whether this person is the only one in Liberia or whether there are thousands of people living in these circumstances. It may differ according to: distance and relation between respondent and Liberian, respondent's level of wealth, respondent's perception of the Liberian's experience of poverty, the specific option for alleviating his distress, how many other people will contribute money to the option, and finally to respondent's opinion about the chances for success of this option.

Three aspects become clear as one moves from the cola bottle to the tree and then to the poor man concerning the context change of the WTP valuation. These aspects are closely related to the concepts of public goods and externalities or non-market interdependence (see Table 3.3).

- First, the role as an actor changes, and especially the direct use type of welfare that is experienced. The cola can simply be bought and used individually, but the tree is already quite a different thing, unless it is turned to timber, and the poverty of a fellow human being differs even more. One moves from more usage-oriented welfare to more non-usage-oriented welfare.
- Second, the predomination of 'individual welfare' versus more 'collective welfare' changes. Although it may be argued that welfare is – in the end – always an individual affair, it is clear that valuation of the different objects involves collective consideration to a different extent. Individual welfare considerations predominate when considering the valuation of a bottle of cola. Although the environmental functions of a tree may be usage related, their use is mostly of a collective kind: valuation only makes sense when considering function or use in a system that is far larger than the individual. Collective valuation predominates in the case of the poor man in Liberia even more; it seems to involve increasing consideration of collective action, as well as consideration about the collective of which the poor man is a part.
- Third, the familiarity with the WTP mechanism changes. Giving a WTP expression for a bottle of cola is supported by common experience of the market mechanism in which these small scale individual valuations are real and meaningfully accurate; that is, they may lead to getting the product and they are a clear signal in the market place. Realism and meaningful accuracy is quite different in the tree and poor man's case.⁶³ The market mechanism is not a great help here. Modern society has developed a whole range of mechanisms in this area, political processes, intelligent deliberation, lobbying, NGO initiatives, or collective consumer or business actions. Within these mechanisms giving individual WTP valuations is generally unfamiliar, because these mechanisms have a much more consciously collective character.⁶⁴

⁶³ Compare for instance Tiwari, 2000 (p. 284) writing about the lack of incentive by farmers to establish a 'beneficial' change concerning irrigation water.

⁶⁴ These mechanisms, needless to say, are – at this stage of their development – not completely effective; otherwise sustainability would not be a major policy concern.

Table 3.3 Relevance of types of valuation and familiarity with giving WTP estimates for separate Triple E elements

	Use and non-use		Individual or collective welfare considerations predominate		Familiarity of WTP mechanism	
	Use	Non-use	Individual	Collective	Familiar	Unfamiliar
A bottle of cola	+++	+	+++		+++	
A tree	+	++	+	++	+	++
A poor man	-	+++	-	+++		+++

+++ highly relevant nearly always; ++ relevant; + somewhat relevant sometimes; - irrelevant

As long as use values are involved and WTP is part of a familiar mechanism, it may be obvious that analyst and stakeholders alike have a natural understanding about the context sensitivity of the expressions, and therefore also about the limitations and proper use of such values.⁶⁵ The more that non-use values and unfamiliar mechanisms are involved, the more the analyst has to rely on stated preferences, on explicitly requested WTP expressions. The framing of the valuation then becomes crucial for success.

In recent decades an increasing body of knowledge has developed, which illustrates that expressions of preferences and decision-making upon them quite often are *extremely* sensitive to changes in the context (Edwards and Von Winterfeldt, 1986; Kahneman and Tversky, 2000).⁶⁶

First, the literature shows that in explaining and understanding actual WTP choices to people, the simple notion of a use-related utility is not sufficient; transaction utility is also needed. Transaction utility measures the perceived value of the ‘deal’.⁶⁷ Important for assessment of the utility of the deal is that the WTP is given in relation to a reference price – held implicitly or explicitly by the individual in that context – deemed relevant.⁶⁸

Second, the notion of mental accounting may be emphasised. People organise purchases in categories of expenditure and the rules they follow are not neutral to their WTP values. Money

⁶⁵ Zerbe, 1998 argues that preferences should be well formed, or more general, respondents should not find it difficult “to compare two worlds” (p. 423) to warrant the use of CBA.

⁶⁶ This literature treats both analyses of artificial choice examples – like for instance gambling – and real life behaviour analyses – like for instance a taxi driver in New York City determining his working hours. Without trying to give a complete overview of this literature, it may be useful for some readers to grasp important insights this literature has highlighted on the use of WTP estimates. At least the following phenomena violate the viability of taking the standard approach of having a simple WTP as the trade-off expression for intensity of preference (Tversky and Kahneman, 2000, p. 46): 1) Framing effects – Framing of decisions, wording of choices, order of treatment of options and numerous other contextual aspects yield systematically different preferences. 2) Risk seeking – Risk aversion is often assumed in the analysis of decision-making under uncertainty, but in settings where there is a small chance for winning a large prize, and in situations where people must choose between a sure loss and a substantial probability of a larger loss, they seek the risk. 3) Loss aversion – “Losses loom larger than gains.” The observed asymmetry between losses and gains is far greater than can be explained by decreasing risk aversion or by income effects. Compare also Zey, 1992 for various valuable contributions on alternatives to rational choice models.

⁶⁷ The widely analysed and often established loss-aversion lies at the basis for this distinction (Thaler, 2000, p. 248).

⁶⁸ People lying on the beach and willing to buy an ice-cream will give a different WTP for an ice-cream that is bought *for them* (!) in a fancy resort hotel than if it is bought for them in a small run-down grocery store (Thaler, 2000). The existence of transaction utility is of course widely used in the market place when for instance sellers emphasise savings relative to the regular retail price.

in one mental account (say, daily groceries) may not be a perfect substitute for money in another account (say, car travel-related expenditure) (Thaler, 2000).

Finally, the relevance of framing and perceiving choices as gains or losses has proven to be very relevant. Choices involving exactly the same expected outcome may be valued quite differently whether they are framed as losses or as gains.⁶⁹

Two conclusions may now be drawn on the issue of WTP as a shortcut to welfare measurement in the sustainability context.

Valuable options exist for using WTP values for assessments of economic development impacts. However, contrary to what many non-economists tend to think, WTP values for everyday economic objects are as context-sensitive as many other preferences. Using these WTP values in the sustainability context will often mean ‘stretching their use’, as for instance, when they are taken as a basis for long-term or global assessments. Context sensitivity of the WTP values requires caution in these procedures and the outcomes are naturally limited in their interpretation.

Second, using WTP values for other Triple E impacts in the sustainability context i.e., for environmental degradation and extreme poverty, is more complicated due to the context sensitivity of the framing of the choice situation and the subtleties of perception therein. Although this point may be seen as general and not specifically bound to the sustainability context (Vickers, 1965), the problems of framing and context sensitivity seem to increase strongly in the sustainability context, as in this context:

- areas of ‘hesitation’ are approached (e.g., human life)
- non-use values are more relevant
- non-familiar mechanisms for monetary expressions are involved
- and more collective welfare considerations are at hand.

3.3.3.2b *The relevance of the dual-preference structure of consumer vs. citizen*

Several aspects were identified above that make WTP expressions more or less useful or complicated in the sustainability context. To summarise all this from a different angle, the concept of the dual-preference structure of consumer vs. citizen can be used (Campen, 1986,

⁶⁹ A telling example of this framing of gains and losses, which incidentally also reveals some hesitation in valuing human life, is discussed in Kahneman and Tversky, 2000. The choice situation is that the USA is threatened by an outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs A and B have been proposed to combat the disease. The impacts of these programs are as follows:

If program A is adopted, 200 people will be saved;

If program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved.

A large majority of respondents (72%) in this context choose option A, although we note that the expected value of both options is the same. The explanation is that the framing of the reference point is one of a loss of 600 lives, in which case most people are risk averse in this situation.

Now consider a change in the wording of the options. People may now choose between program C and D.

If program C is adopted, 400 people will die;

If program D is adopted there is a one-third probability that nobody will die and a two-thirds probability that 600 will die.

Although analytically or objectively this seems to be the same choice as between A and B, and one might expect the majority to choose option C, most of the more than 150 respondents perceive it to be a different choice and choose option D (78%). Even after being confronted with their conflicting answers, respondents do feel puzzled, but mostly stick to their choice of being risk averse in the ‘lives saved’ version and risk seeking in the ‘lives lost’ version. (Kahneman and Tversky, 2000, p. 5).

Slocum, 2004).⁷⁰ The dual-preference structure underlines that there may be a difference in the preferences (implicitly) expressed in the market and the preferences expressed when explicitly asked for and non-use is more relevant. When in the market place, people are self-interested maximisers, while when explicitly asked, people are more likely to act according to their political capacities and have more public-spirited preferences.⁷¹ This duality is according to some (Campen, 1986) related to Plato's distinction between man as a private individual and man as a citizen.⁷²

As aforementioned, the sustainability context requires both economic realism at the micro level and the adoption of a global perspective and a long-term view at the macro level. The global and long-term perspective is definitely not normally the view of people when buying a bottle of cola. Although people may indeed hold preferences about, say, the global availability of cola or the availability of buying cola for their grandchildren, these preferences are not normally 'expressed' when giving an individual WTP for a bottle of cola (compare section 3.5).

Dual preferences are in many cases in opposition with each other and they may lead to different behaviours (Campen, 1986, p. 161; Common et al., 1997; O'Connor, 2004). To the analyst evaluator the bad news is that there seems to be no simple way of determining which of the two preferences is the 'true' preference: "both are ours, however inconsistent they may be, and which one we invoke depends on the circumstances in which we find ourselves" (Campen, 1986). O'Connor (2004a, p. 175) in this respect speaks of "the permanent co-existence of multiple perspectives."

To the CBA analyst active in the sustainability context this situation should be a cause for substantial methodological concern (Luchini, 2002) as the analyst has to decide in which circumstances respondents find themselves; that is, in which context the WTP estimates are gathered. One cannot: "expect individuals to behave as consumers in regard to hard decisions such as environmental protection [which are seen as] matters for 'social regulation', involving ethical rationality and informed deliberation rather [than matters of] choice on the basis of given, and likely poorly informed, consumer preferences" (Common et al., 1997, p. 227).⁷³

Explicitly referring to the consumer-citizen preference structure the general summary conclusion of section 3.3 may be that the sustainability context of evaluation brings in both consumer and citizen values, while these two types of preferences may be at odds with each

⁷⁰ On the possible resemblance of styles – in terms of participation and power – of consumer and citizen behaviour of individuals, see Hyman and Shingler (1999).

⁷¹ Common et al. (1997, p. 227) quote Sagoff on "the conflict within us," which may be illuminating: "As a citizen, I am concerned with the public interest, rather than my own interest; with the good of the community rather than simply the well-being of my family. As a consumer ... I concern myself with personal or self-regarding wants and interests; I pursue the goals I have as an individual. I put aside the community-regarding values I take seriously as a citizen and I look out for 'Number One' instead" (Sagoff, 1988, p. 8, *The Economy of the Earth*, Cambridge University Press, Cambridge).

⁷² In welfare economic terms this duality is sometimes formulated as preference or utility versus welfare. Preference or utility is based on the outcome to the individual; welfare is based on the outcome to the individual but also on the outcome to others (Ng, 1983, p. 7-13). In special cases they may coincide. Compare also Balbus (1971) for an early in depth treatment of the 'interest' concept in various interpretations. Compare further Granovetter, 1992 for a likewise treatment of embeddedness and treating economic actors as 'over- and under-socialised'.

⁷³ Compare House (2000) who ends his discussion of limits to CBA with a similar position: "Ends should be determined by reasoned discussion, not by *unexamined* preferences [italics inserted], however efficiently they are satisfied."

other and WTP measurements may be rather difficult to use for citizen values.⁷⁴

3.4 CBA and valuing gains for some, losses to others

3.4.1 Introduction

Project evaluation in an accountable-sustainability context is logically a social matter in which many individuals are involved and in which various social values are directly relevant. How does welfare economics address the problem of aggregating individual values?

The starting point for aggregation for welfare economics is individuals' willingness to pay (or WTA) for certain changes. The standard procedure is very simple: add up all the money values of winners and losers and see if overall there is a positive or negative net balance. If there is a positive net balance as a result, the winners could *potentially* compensate the losers: and theory speaks of a Potential Pareto Improvement (PPI) or the Kaldor-Hicks compensation rule. The word potential is used because actual compensation as such is not deemed necessary (or practical).⁷⁵

However, although the discussion thus far may have suggested otherwise, in practice CBA is not much concerned with actually aggregating individual WTP values, especially not when evaluating economic impacts. CBA usually does not start from stated preferences of individuals. It often starts directly from aggregate (competitive) market values: production values or sums of money spent or aggregate estimates of the consumer surplus.⁷⁶ Van der Veen and Otter (2002) emphasize that in this step from the micro to the meso or macro level 'an observation set' is *constructed*, which is a far from simple process.

If, given this practice of working with aggregate market values, the simple rule of adding WTP values should have some bearing on the question whether social welfare in a more general sense has improved, complications arise. In section 3.3 it was already made clear that it is predominantly the individual use-related part of welfare that is commonly expressed through the market. Aggregation of these market related values does not change the situation. Obviously, it leaves the CBA analyst building on market related data with several limitations concerning their broad welfare implications.

One of the most important issues in which this limitation is clearly seen is income distribution, which has been widely discussed in economic literature. Poor people have little to offer when asked for their willingness to pay, or may be willing to accept very low amounts of money in compensation for (even severe) welfare losses. The opposite holds for rich people. So the that is least needed in order to give the Kaldor-Hicks compensation rule some plausibility is that stakeholders and decision-makers involved should judge the existing income distribution as reasonably fair. However, in reality and especially in the accountable-sustainability context, with the urgency of abating extreme poverty, this condition is far from being met (Brekke, 1998; Drèze, 1998; Johansson, 1998).

⁷⁴ The problem is all the more pressing since CBA often uses aggregate market values, which – as seen above – are commonly based on the individual usage part, the consumer part of the preferences. The use and meaning of aggregate market values is examined in section 3.4. It will be noticed in section 3.5 that disregarding this dual preference structure may explain the poor performance of many WTP questionnaires concerning environmental problems.

⁷⁵ A major reason probably being that there will not often be a practical mechanism for the value transfer – a problem that was identified in chapter 2 too (section 2.3.2.4).

⁷⁶ Layard and Walters, 1994, p. 194. As seen in section 3.2.5, shadow prices may also be estimated that better reflect real welfare values. These prices are in general also aggregate values (Bruce and Harris, 1982; Drèze and Stern, 1994) often taken from other, sometimes informal, markets.

Economic literature has a popular way of formulating this somewhat disturbing situation, which is that the PPI or net gain or loss only shows the efficiency impact.⁷⁷ To say whether an efficiency change is really a gain or loss of social welfare requires a judgement on distributive equity, which is beyond the expertise of the economist or CBA analyst. Efficiency is then seen as somewhat value free, while the equity aspect is much more normative. Blaug (among others) however, convincingly argues that both efficiency and equity judgement require ethical considerations (Blaug, 1985, p. 608), so distinguishing both types of judgement only for that reason is hard to defend (Zerbe 1998).⁷⁸ However, distinguishing efficiency from equity on practical grounds has become widely accepted (Brekke, 1998, p. 122).

To somehow tackle the issue of income distribution within CBA it has been suggested to give different weights to costs and benefits that accrue to different groups. In the sustainability context it was argued that especially groups in extreme poverty should warrant attention; so one might consider valuing benefits to these low income groups higher than extra benefits to (already) rich groups.

There should be no doubt about the impact of distributional weights on the final outcome of CBA: giving weights will determine the outcome, and the other way around, virtually any outcome can be 'achieved' by setting the weights 'right'. The procedure is technically very simple, but the crucial question is: how to derive the weights in an acceptable or accountable way?⁷⁹

3.4.2 The difficulty of distributional weights

3.4.2.1 General description

3.4.2.1a The rather limited success of distributional weights

There has been prolonged debate about distributional weights in CBA. The main impetus for the use of distributional weights in CBA came in the 1970s from international organisations like the World Bank and the United Nations working in developing countries, and being actively confronted with the impact of projects on extreme poverty. Although debate continues for instance in the assessment of climate change impacts (Azar, 1999;⁸⁰ Johansson-Stenman, 2000), it was Mishan in the 1980s who made the issues at stake most clear.

To understand Mishan's position, one may start with Squire and Van der Tak (1975), who while working at the World Bank, strongly argued for the incorporation of distributional weights in cost-benefit analyses. Squire and Van der Tak explain the reason why they opt for this approach. The 'traditional' growth-oriented approach to CBA was, according to the authors, often defended on the grounds that it ensured that the available resources yielded

⁷⁷ See for instance Lakshmanan et al., 2001 in the case of costs and benefits of transport.

⁷⁸ More specific in CBA literature the argument has been that disregarding distribution implies an implicit unit weight to all costs and benefits: so one cannot escape distribution matters. Mishan (1982, p. 40) disagrees, mainly as the argument is highly theoretical.

⁷⁹ Compare Miller (2001, p. 30): "This ... illustrates the methodological difficulties of dealing with measures of inequality in a simplified way, and underscores the desirability of using natural measurements wherever possible."

⁸⁰ Interestingly, Azar after discussing the merits of different ways of giving distributional weights in the climate change context, suggests not to put too much research effort into further developing the approach, as it does not adequately solve the ethical issues involved. Instead, Azar suggests more explicit discussion of difficult issues and then aiming for a simple target level of emissions with low risks (Azar, 1999, p. 265).

were maximised and that governments had diverse instruments to redistribute this maximum yield in any direction. The main reason for Squire and Van der Tak to adopt a new approach seemed to be that “the government’s ability to redistribute income in general may be limited” (p. 6).⁸¹

Applying distributional weights requires that equity or a more equal distribution of income or consumption be an aim for a given society. In the Squire and Van der Tak approach the government should give weights: the analyst may interpret government actions and statements and infer weights from them, but this task may be a rather difficult one.⁸² First, for various reasons, government actions and policy statements may contradict. Second, determining whose policy statements or actions within the government should be most decisive is difficult; should they be central or local? Politicians or bureaucrats? Before or after the next elections? This ‘revisionist’ approach to CBA, as Mishan calls approaches to introduce distributive weights, requires that “no difficulty arises in discovering the locus of political authority” (Mishan, 1982, p. 34). In the sustainability context with decision-making aiming to be accountable to many stakeholders and with impacts of single projects occurring in various different countries, this seems even more difficult; which government actions or which government documents should be used as a basis?

Mishan strongly argued against the use of distributional weights in CBA by showing the dramatic implications of introducing distributional weights into CBA. Why is it dramatic? It is dramatic because in the end it leaves CBA without a judgement criterion. The CBA analyst may gather data and check for consistency of policies, but will not be able to render independent judgement of any kind: judgement has to come from the ‘policy maker’ or ‘decision-maker’, whose weights will determine the outcomes.

The cost-benefit analyst in Mishan’s view gives independent judgement on the economic efficiency of allocative changes by adding the monetary valuations of the impact population concerned. If this gives a positive result it can be interpreted as “a change by which – if this gain could be costlessly redistributed among all affected by the change – every person could be made better off” (Mishan, 1982, p. 40). Although – as seen above – it seems very difficult, at least theoretically, to argue that efficiency can be separated from distribution, it seems nevertheless useful to use ‘efficiency’ as a name for this type of judgement (with the concomitant costless redistribution) and see it as separate from judgement on equity.

The use of distributional weights in CBA has apparently largely vanished in World Bank and UN project evaluations, where their main popularity once was (Little and Mirlees, 1994); although now and then the approach is still defended (Johansson-Stenman, 2000).

3.4.2.1b *Ethical consensus*

Mishan sees as the ultimate basis for ‘conventional’ CBA an ‘ethical consensus’ in a society about an economic judgement criterion (Mishan, 1982, p. 41). So having a positive net present value (or social benefit) in monetary terms is an increase in welfare, ‘when, and only when’ *society conceives* it as such. To be more precise, Mishan highlights four *beliefs* that if one or more of them are held widely, produce an ethical consensus:

- 1) Projects that are efficient according to conventional CBA do not usually have detrimental

⁸¹ In the same line of reasoning, they see (in many developing countries) reason to give extra weights to investment over current consumptions, because of several constraints that “limit the government’s ability to increase savings by means of monetary and fiscal policy” (p. 5).

⁸² Squire and Van der Tak have themselves already noticed difficulties here and gave the advice to interpret “with caution” (p. 102).

- impacts on the distribution of income
- 2) The existing tax structure offers adequate safeguards against substantial distributional injustice
 - 3) Over time, and over many projects, efficient changes are unlikely to make the poorest people any worse off
 - 4) Over time, and over many projects, efficient changes have a better chance of raising the living standards of the mass of the population than other changes do.

Two things are noteworthy. First, it is likely that in the sustainability context these beliefs are not generally held. They may hold to an important extent though; and this determination could be assessed within the evaluation. Second, the basis of CBA is formed by beliefs about impacts and ethical consensus about judgement criteria: values which seem to be typically part of the citizen preferences.

3.4.2.1c Description of distributional impacts

Over the years another approach has developed: careful description of distributional effects (Campen, 1986; Lichfield, 1996; Potts, 1999; Eijgenraam, 2000, 2003). There is no 'correction' taking place using distributional weights: the impacts (costs and benefits) are specified for different groups in the relevant impact population. The analyst should, after he has presented his outcomes carefully, describe the impacts it has on specific groups in society that warrant special attention. This careful description can also take the form of presenting end-results of CBA in tables that show impacts on various stakeholder groups.⁸³ This approach has probably been developed most thoroughly by Lichfield (1996). In the Planning Balance Sheet of Lichfield, different stakeholder groups are identified and costs and benefits that are elements of CBA are shown for all groups.⁸⁴

3.4.2.2 Assessment of pros and cons in the context of accounting for sustainability

3.4.2.2a Judgement consensus among stakeholders is the final basis of CBA

A very important point has arisen from the discussion of the distributional impacts about the foundation of CBA. The point is that the final basis for conducting CBA and using its judgement criterion is not "as popularly supposed (even within the profession)" (Mishan, 1982, p. 41) an objective way of measuring change in societal welfare. This judgement cannot "be derived from individual welfare functions, no matter how elaborate" (Mishan, 1981, p. 257). The basis for judgement is '*beliefs*' that should be '*widely held*'. These beliefs are about the real life relation between increased efficiency and the impact on living standards in general (over many projects and over time). This final basis, as should be noticed, fits the demands that the sustainability context places on evaluation: accountability towards broad groups of stakeholders (Farrow and Toman, 1999).

We can observe further that these beliefs are not necessarily widely held. Mishan has tentatively argued that in today's pluralistic, non-traditional and fast-changing society this

⁸³ E.g., Eijgenraam et al., 2000 (and 2003); Kingma, 1996.

⁸⁴ Lichfield (1996) has further developed this approach in the Community Impact Analysis (CIA) – see also chapter 5.

consensus may be dissolving.⁸⁵ The more divided people are about the value and impact of products, services and production processes, the less of a basis there is for CBA's prescriptive statements (Mishan, 1981, p. 261-262).⁸⁶

3.4.2.2b Careful description of distributional impacts the best option, with minimum demand of standardised attention to the extremely poor

Although some authors suggest that careful description of distribution impacts can occur as an afterthought to the CBA analysis, it should be emphasised that the CBA analyst usually does not know the impacts on different stakeholders (but may have a rough idea). As seen above, despite the theoretical attention CBA pays to individual based valuations, the analyst more often than not uses aggregate values. Careful description can therefore only be done if the analysis in an early stage has been designed to investigate these impacts.⁸⁷

To guide such an analysis one has to identify stakeholder groups. Identification requires judgement, as these groups are not 'naturally given'. CBA theory is simple and straightforward in defining the total impact population: if one sees an impact somewhere, the people whom it affects should count as part of the impact population. However, identifying stakeholder groups may be difficult, as choosing groups is a normative matter (Hellendoorn, 2001). This choice would have to be made by the decision-maker. Although preferably the analysis should be as objective as possible, this does not seem to be a serious objection. CBA may gain in relevance, provided of course the normative choice is made explicit.

A major difficulty here is that many groups can be identified. To mention just a few possible group-pairs (of which further divisions are possible): producers and consumers; high income groups and low income groups; local population and others; women and men. Working with many divisions may in itself generate problems and problems may also occur because groups overlap.⁸⁸

Furthermore, an important implication of breaking down CBA results to stakeholder groups and carefully describing them somehow, is the requirement of conceptual clarity of these disaggregated results to the identified stakeholder groups. One way to try and achieve this may be to strive for consistency between financial analysis (for various stakeholders) and the somewhat broader 'economic' analysis of CBA (Potts, 1999;⁸⁹ Fabré, 1997). This consistency may imply that effects should relate as much as possible to the units in which incomes are received (Potts, 1999, p. 583) and showing impacts on value added may be very useful.⁹⁰

The recent debate on climate change has shown the difficulties of including stakeholders in

⁸⁵ In which (e.g.) many resources are used for activities that part of society sees as unworthy or degrading and others see as without harm or liberating or fashionable.

⁸⁶ The fact that throughout the years apparently no consensus was formed about the use of distributional weights, even though the procedure is quite transparent, and although useful data exist on which they may be based, may serve to illustrate both the importance and the difficulty in achieving consensus – even if it is consensus about a 'revisionist' element of CBA.

⁸⁷ E.g., a poverty analysis, or analysis of gender; compare Little and Mirlees, 1994; see also Lichfield, 1996 for combining CBA outcomes with describing impacts to stakeholders.

⁸⁸ One person may for instance be part of the women group, the low-income group and the consumer group. Staying close to the non-double counting strength of CBA seems to imply that preferably stakeholder groups are chosen such that they lead to non-overlapping groups.

⁸⁹ The domestic price numeraire approach of UNIDO 1978 was developed for this (Potts, 1999). French development project analysts used it (in their Effects Method), but it never became popular in the non-Francophone world.

⁹⁰ As a result outcomes may then also be easier to relate to the national income accounting framework and concepts (Potts, 1999, p. 584; Eigenraam, 2000).

evaluation in a sustainability context, especially stakeholders who have widely differing views and contrasting interests (Tóth, 1997, 2000; Lempert, 1999).

In the sustainability context of evaluation it is very easy to identify multiple and diverse stakeholder groups (see chapter 2). The presentation and interpretation of CBA outcomes may become rather complex as a result (Van den Bergh, 2004). There seems to be an irresolvable tension between the simple single number efficiency judgement of CBA and the possibly complex and open descriptive information about distributional impacts (Campen, 1986). Therefore, it seems natural to try to limit the number of stakeholder groups in CBA, for which the explicit break down of results is shown.

In the sustainability context stakeholders who live in extreme poverty require special attention; it seems logical then that analysing impacts on this particular stakeholder group should become standard practice in CBA.

3.4.2.2c Everyday aggregate market values reflect parts of welfare changes connected with efficiency and individual use; they do not reflect fundamental values like distributional fairness

The reader may recall the discussion in 3.1.2 about the market mechanism and what lies beyond it; welfare theory asserts that the more reality meets the ideal market mechanism, the more aggregate market values can be trusted as indicators of welfare. Market failure in reality prevents complete trustworthiness of market values and implies that using real life aggregate values by the CBA analyst will usually mean “making the best of a bad situation.” The same section also showed that public-ness of goods (national defence), externalities of production (smoke), and externalities of consumption (road congestion), and more generally non-market interdependence of welfare (“keeping up with the Joneses”) usually lie beyond values of the market mechanism.

Now that the issue of income distribution has been discussed, these points can be understood far better. Aggregate market values are a good starting point for measuring efficiency impacts, but as this section has clearly shown they are not a good starting point for broader welfare notions such as fairness of distribution. Now fairness of distribution may relate to many different aspects of broad welfare, like the distribution of power, distribution of personal education or development options and also to the more everyday economic issue of distribution of income, discussed above. Still even this later issue is ‘too fundamental’: it lies beyond the preference values expressed through the market mechanism.

In section 3.5 we will observe that a clear understanding of what lies within reach of monetary estimates and what is beyond reach is crucial in the multidisciplinary sustainability debate. Perhaps running ahead a little, a link between these fundamental environmental valuation issues and the distribution issues here may be shown. When in section 3.5 and the appendix to this chapter we will look at the often confusing debate around the use of WTP values in the ecological economic literature, it will be emphasised that efficiency does not refer to consumers’ WTP only; efficiency requires considering the supply curve as well. In the marketplace demand is confronted with supply and especially the supply costs of extra units. Regardless of total utility – and thus the willingness of ‘consumers’ to pay – only things that are *limited in supply* receive a positive price.⁹¹ Consequently, there is no way to know how much market prices as such do reflect an ‘absolute or broad’ welfare measure – or a change therein for that matter.⁹² As the saying goes: “the best things in life are free.”⁹³ People act in

⁹¹ See Aldred (2002) on ‘commodification’.

⁹² Compare the appendix on Costanza et al., 1997.

the market place on the basis of the ‘best things’ already being there⁹⁴. Clean air, clean water, open space, raw materials have all been – and some still are – without a price since they were once abundantly available. It is only when these ‘best things’ become short in supply that their value can be expressed in the market place, provided of course that there is some way to supply it. If this happens one will see *Lauderdale’s paradox* occurring: public wealth may decrease as, for instance, clean water becomes more scarce, while private wealth may increase, as for instance, private companies emerge to start selling bottles of scarce clean water (Daly, 1998). Thus when environmental assets become scarce in supply then distribution issues – gains to some, losses to others - may come in too.

At this point two further topics concerning the merits of CBA may be reconsidered; the deepened understanding which section 3.4 has brought to bear may make a reconsideration rewarding. The reason for treating these topics at this point in our discussion – although they are not directly linked to distribution as such - is that both in section 3.3 and in the upcoming section 3.5 the emphasis on WTP – somehow - leads to a consumer and demand curve ‘bias’. The discussion in the current section is more balanced in that it refers to economic or market *system* outcomes: both supply and demand, both producers and consumers are relevant.

3.4.2.2d *Supply curves show info-richness of market values*

Now that the point that market values can only be understood as a combination of supply and demand has been stressed, it may once more be useful to address explicitly the power of CBA in the sustainability context. Understanding the supply curve is perhaps the easiest way to appreciate this power. The supply curve reflects production costs, which, incidentally include – normal rate of return to capital (i.e., profits). The crucial aspect to understand is that the curve reflects the costs of various inputs needed for the production of a certain good or service and it therefore reflects the outcomes of various other markets. If the supply curve for a table, for example, is considered, it reflects the price paid for, say, the tropical-tree timber that had to be bought to make it. In this price the wage of the wood-cutter is a major element. Furthermore, the wage of factory workers who combine the wood and a metal frame to construct a table, are reflected in the price paid, just as is the capital invested in the plant, and the transport costs to ship it to the shop. If changes occur in the supply chain because tropical trees are becoming scarce, or factory wages for workers are lowered, these will change the position of the aggregate supply curve. Let us return to the examples of section 3.3.3.2:

- the German woods can no longer be used for timber production, which puts extra pressure on other timber sources, and
- the poverty in Liberia is successfully tackled, which may lead to more expensive sea transport under Liberian flags,
- then the suppliers of wooden tables used in U.S.-cafés may decide to raise the prices of their products – that is, the supply curve shifts somewhat upward,
- which in turn may lead to a small upward shift in the supply curve of cola supplied in

⁹³ A point that may raised here as it remains a source of confusion in the ecological-economic debate (see Costanza debate in par. 3.4.7 and the linked functions of nature approach of De Groot, 1992), is that goods very important to people’s welfare, like for instance, water, do not always have a high price in the market. Other things with very limited use, such as diamonds, may have a very high price in the market. The diamond/water paradox was already noticed by the founding father of economic science Adam Smith (Smith, 1976, chapter 4). So it is not necessarily absolute value, or welfare that is expressed in market prices, but rather something else. One needs ‘marginal analysis’ to explain that which is expressed.

⁹⁴ Just like they act in the market place on the basis of the distribution of income already being there.

cafés.

For the purpose of evaluation, the reader may appreciate the tremendous amount of information reflected in market values, and the natural way of causally valuing related steps in the production.

3.4.2.2e Aggregate market values are important and consistent everyday economic realities without broad welfare connotations

CBA mostly limits itself to objective aggregate market data and uses individual valuations modestly. The textbook defence often is that these aggregate data can be used to measure (changes in) social welfare, but it has been shown that this claim has to – and can be – qualified⁹⁵. It seems to be valid only in ‘ordinary market conditions’ and generally speaking, only in as much as these values relate to individual usage in the dual preference structure.

By now we can formulate the connection between CBA and welfare. The CBA practice of using mostly objective market data *and* excluding transfers *and* for instance focussing on aggregate production value may remain. The analyst can choose to stick with these parts of CBA practice – not because they can be used in any sense as ‘true’ measures of ‘broad notions’ of welfare, but because they may generate important information about efficiency impacts – in other words, about everyday economics.⁹⁶ Everyday economic values may be very incomplete; they may focus too much on individual utility but they are real and are deemed important. These CBA-generated values provide judgement information on economic impacts: on viability or efficiency in an everyday economic sense. They may thus provide an economic judgement criterion which many stakeholders understand naturally, since they are involved on a daily basis in matters of efficiency, budgets and income possibilities.

⁹⁵ Formulated technically, CBA textbooks commonly rely on Willig (1976) for establishing the relation between the demand curve’s consumer surplus and welfare. The maximum theoretical support for using consumer surplus as a monetary measure for welfare changes, and thus as theoretical foundation for CBA, is that: ‘under ordinary conditions’ the consumer surplus is a good approximation of the theoretically more directly utility-related concepts of Equivalent Variation or Compensating Variation and that “the error of approximation will often be overshadowed by the errors involved in estimating the demand curves” (Zerbe and Dively, 1994, p. 90). Compare also Bruce and Harris (1982, p. 756) who derive criteria for compensation for ‘small’ projects that “depend only on local information in the neighbourhood of the initial equilibrium.” See also Alston and Larson (1993) for a somewhat broader discussion. See Stennek (1999) for a discussion of the (negative) impact on the usefulness of the consumer surplus as a measure of welfare change in the case where consumers are not risk-neutral.

⁹⁶ If one thinks about the everyday economic base of CBA, this may explain a puzzle in the history of theory and practice of CBA. Various authors looking at CBA history have noted that its welfare economic foundation is very problematical: there is a large gap between theory and practice. Krutilla (1981) speaks of the “unaccountable growth of applied welfare economics [CBA] in the face of severe theoretical criticism.” Pearce and Nash show that practice of CBA grew tremendously when the theoretical base (welfare economics) was seen by many as ‘truly dead’ (Pearce and Nash, 1981, p. 2; cited in Campen, 1986, p. 70; compare Zerbe, 1998, p. 420: “[...] practitioners have generally ignored the critics [...].”) If it is acknowledged that the basis of CBA is largely rooted in everyday economics, it is not hard to explain why the practice has grown. Many projects have very clear-cut everyday economic impacts: costs and benefits. Much of everyday economic costs and benefits will therefore be deemed relevant to any evaluator and stakeholder whether he or she is a trained welfare economist or not. Interestingly, Sugden and Williams seem to acknowledge this view clearly in their book *The Principles of Practical Cost-Benefit Analysis* (1978). Contrary to most CBA textbooks, they do not begin with utility and welfare economics. Their starting point is the everyday economics of ‘financial appraisal,’ be it in the setting of a commercial firm or in a public agency. They see CBA as an extension of financial appraisal and they in fact use welfare economics as a basis for this extension, but not as their starting point. Compare also Polak and Heertje (2000, p. 152) on practice and theory of CBA for infrastructure assessment: “[...] [CBA] could be operationalised in a rather ‘practical’ way, which was not obvious from its theoretical, primary formulation.”

One final aspect of using (revealed preference) objective market data rather than (stated preference) willingness to pay can be mentioned: it relates to the fact that revealed preferences are consistent as a systematic (dynamic) whole over the entire economy. Willingness to pay (or WTA) values are largely partial and static. People may say they value a thing very highly, but if they do not actually have to pay for it, then one never knows precisely what the value means. Formulated more strictly: it is unclear how to interpret these results, as there is no clear point of reference as to whether it is great or small.⁹⁷

Real market values do not have that possibility of inconsistency.⁹⁸ What people receive as income is spent or saved. Money spent by consumers is received by producers; and then spent again. Money saved is then invested by others, with an interest or dividend payment in return. People have real budget constraints; they cannot spend more than they have without borrowing or using saved resources.

3.4 Conclusions on handling distributional impacts

An overall conclusion on distributional impacts seems to be that ‘carefully describing’ them is the best alternative to using distributional weights – as no consensus has been achieved on the latter. For clear interpretation and presentation of CBA results, the number of stakeholder groups for which impacts are analysed and shown should be limited. Assessing impacts on the extremely poor could be standard in a sustainability context.

The most important lesson from the debate on distributional matters is that it clearly shows that a CBA that does not address these matters does not refer to ‘broad’ welfare notions, since broad welfare notions would include notions about fairness of distribution: taken as a given in a standard CBA approach. CBA results give judgement information on efficiency only. In practice this judgement information is predominantly based on the everyday economics of the individual use part of the dual preference structure and may be (further) limited by various forms of ‘market failure’. The ultimate reason for accepting these CBA results as relevant and important is the consensus among stakeholders that they are indeed important and relevant. This consensus in the sustainability context cannot be taken for granted, but its existence has nevertheless to be assessed.

3.5 CBA and valuing environmental impacts

3.5.1 Introduction

This section will analyse how well CBA performs in assessing environmental degradation impacts. For many environmental assets, or functions of assets, no market values exist. CBA requires monetary valuation of impacts and the analyst will therefore have to calculate prices or monetary values for environmental impacts. Although CBA has a long tradition of incorporating

⁹⁷ One does not know their stated preference for other things (unless this is requested too; this would inevitably become a very time consuming practice). Furthermore if people were actually required to pay the value, they would spend less elsewhere (or save less). One cannot predict the direction they would take; these repercussions cannot be traced easily.

⁹⁸ Compare for instance Clarke (2002) who finds higher stated preferences than revealed preferences for a health service.

non-market effects on, for instance, consumers (Mishan, 1975), applying cost-benefit analysis to projects with important environmental effects is still complicated (Hanley, 1992; Van den Bergh, 2004).

For over a decade, discussion on valuation of environmental values and functions has been intense (Barbier et al., 1990; Barde and Pearce, 1991; De Groot, 1992; Dubgaard et al., 1994; Pearce and Moran, 1994; Pethig, 1994; Costanza et al., 1997; OECD, 1997; Faucheax and O'Connor 1998; Pires, 1998; Farrow and Toman, 1999; O'Connor and Spash, 1999; Dabbert et al., 2001; Shi, 2004; Veisten, 2004).⁹⁹ Based on the debate regarding the Costanza et al. (1997) attempt to value the world's ecosystems and services, this discussion is surrounded by much conceptual confusion (see appendix to this chapter). This confusion revolves primarily around the attempt to reach *comparable* valuation results in the environmental and economic domain.

The central question is: How well equipped is CBA – with its reliance on monetary valuation – to evaluate environmental effects? It will be argued that in the accountable-sustainability context, the possibilities for monetary valuation of environmental impacts is useful only in limited cases.

The discussion has two main parts:

- the concept of value as it relates to environmental assets and services (section 3.5.2)
- the possibilities for different valuation techniques to grasp environmental values (sections 3.5.3 – 3.5.7)

The discussion will conclude in section 3.5.8.

3.5.2 The concept of environmental value

3.5.2.1 General description

3.5.2.1a The concept of Total Economic Value

Before treating valuation techniques the concept of value requires explanation. This procedure is essentially standard practice in the ecological economics (Edwards-Jones et al., 2000; Nunes et al., 2000; Bateman, 1994; Garrod and Willis, 1999).¹⁰⁰ When discussing the value of environmental assets and changes therein, the aspects of value should be captured by the valuation techniques. The ecological-economic literature uses the concept of Total Economic Value as a starting point (Fromm, 2000).¹⁰¹ By adding all the separate value components from an environmental asset we derive total economic value. The Total Economic Value is mostly split between a use value and a non-use value part. Non-use value is often subdivided into existence and bequest values.¹⁰²

Use value is mostly subdivided in direct use value, indirect use, and option value. A forest can be taken to illustrate the value components of an ecological asset. People can derive existence value from knowing that tropical rainforests in Africa exist, although they may never see them or 'use' them. Bequest (or altruistic) values can be derived from knowing that others, often

⁹⁹ The volume of readings by Willis et al. (1999) also provides a useful introduction to the debate.

¹⁰⁰ See Van den Bergh (2000) for a useful discussion of the ecological economic field.

¹⁰¹ See Box 3.1 for the use of the word economic. One might argue that 'Total value' as such (without adding 'economic'), could do equally well, as all values seem to be somehow included. Compare De Groot (1992).

¹⁰² Kopp (1992; see below) argues that non-use value is really a better word for existence value as such, which would then include bequest value. See Fromm (2000) for the terminology of different value components.

specified as future generations, will have the possibility of enjoying the environmental asset. Direct use value would (for instance) be the recreational value of a forest. The indirect use value is derived from ecosystem functional benefits, such as regulation of water quality by forests. Option value is derived because of having the forest ‘stand-by’ as an option for visiting, or some other direct or indirect future use.

3.5.2.2 Assessment of pros and cons in the context of accounting for sustainability

3.5.2.2a *Valuation is about anthropocentric value*

It may be noticed firstly that the value components of the Total Economic Value concept are all anthropocentric. Focussing merely on human values attached in one way or another to ecological assets is not above critique, especially in a sustainability context (Rogers, 2000).

For many non-environmentalists, this point, which is based on a rejection of the usual anthropocentric framework,¹⁰³ is difficult to understand, let alone accept.¹⁰⁴ A rejection of the anthropocentric focus in the valuation is often explicated in ethical and moral theories or principles for action (compare Edwards-Jones, et al., 2000, p. 79-83 for an introductory overview).

There is no need to delve deeply into these theories for our purposes. Suffice to say for now that this ‘deep-respect’ attitude (or deep ecological view) towards the non-human environment is non negligible in modern industrialised society¹⁰⁵ and perhaps even more so in indigenous cultures such as that of native Americans or Aborigines. It may then be useful to distinguish: 1) human use value 2) human non-use value and 3) ecological value in its own right. Different valuation techniques and their possibilities and limitations will be discussed below. These techniques will relate only to 1) and 2) above.

3.5.2.2b *Non-direct use values are central, but they complicate comparability among Triple E aspects*

By looking at the Total Economic Value components it may be clear that in the accountable-sustainability context all these value components are deemed important for an environmental asset. However, in many cases there are no market prices available for these components. In the table below this situation is illustrated for a forest¹⁰⁶.

¹⁰³ That is: the framework whereby things have value only to the extent that one or more persons value them – a framework explicitly adopted by welfare economists but also by many others.

¹⁰⁴ Referring specifically to the anthropocentric ‘bias’ in using WTP estimates, Ackerman et al. eloquently expressed this point (quoted from Campen, 1986): “One looks in vain to the economist to provide a value for the benefit *accruing to fish*, which have no money to sacrifice.”

¹⁰⁵ Compare for instance Shaw (2001) quoting Leopold (*not* even a deep ecologist): “It is inconceivable to me that an ethical relation to land can exist without love, respect and admiration for land, and a high regard for its value. By value, I of course mean something far broader than mere economic value.”

¹⁰⁶ Compare Jepma and Munasinghe (1998, p.130) for a tropical rainforest example.

Table 3.4 Total economic value of an environmental asset and the relevance of its several components in a sustainability context

		Environmental asset (a forest)	Relevance in sustainability context	Market price available?
Use value	Direct use value	Visiting the forest	High	No* – only sometimes
	Indirect use value	Profiting from the water regulation by the forest	High	No* – not generally
	Option value	Having the option for visiting or water regulation	High	No*
Non-use value	Existence value	People – far away – knowing (and appreciating) that the forest exists	High	No
	Bequest	Knowing that others (future generations) may profit from the forest	High	No

* This availability may depend on ownership. If the forest is privately owned and access can be denied, there may be a market price for visits.

The major issue in the ecological-economic debate is about achieving *comparable* assessment and measurement on both economic and environmental impacts. On the basis of the earlier discussion (in section 3.3), the reader can examine table 3.5 below, which resembles 3.4 above on the total economic value of a forest, but here a car is taken as an example. Table 3.4 lists comparable value components for an economic asset; it also shows which of the different types of use values one can reasonably expect to be reflected in the aggregate market values most often used in CBA practice.

Table 3.5 Total economic value of a car and the reflection of its several components in an aggregate demand curve

		Economic asset (a car)	Relevance in sustainability context	Reflected in its price?
Use value	Direct use value	Driving a car	High	Yes
	Indirect use value	Watching the car pass on a Sunday (positive), or (negative) disliking the noise it makes	Unclear	No
	Option value	Having the option for driving or for watching it pass or disliking the noise	Unclear	Yes and no*
Non-use value	Existence value	People – far away – knowing (and appreciating) that the car exists	Unclear	No
	Bequest	Knowing that others (future generations or future buyers) may profit from the car	Unclear	No

- Notice that a car is typically an asset type of good; a 'capital' good that is used over longer periods of time. This is why an option value is reflected in its price. If the same table were to be made for a chocolate bar, one would not expect an option value in its price.

The first major point to learn from table 3.5 is that also for everyday economic assets the various value components comprising Total Economic Value do exist. People may be very happy knowing that their grandchildren will someday drive the car (bequest value) or be happy to know that a specific (fancy?) car exists (existence value).¹⁰⁷

¹⁰⁷ There is some mismatch in the forest/car comparison; the long life that may be expected from a forest cannot be expected for a car, which makes the bequest value for 'future generations' somewhat artificial for an individual

The second important point in the table is that these values are only partly reflected in the demand curve type of preferences relevant to market transactions and therefore are not reflected in its price. Prices for economic assets commonly do not reflect much of the non-direct use type of values deemed important for environmental assets. Table 3.5 also shows that it is unclear how relevant these values are in the sustainability context.

The inclusion of non-use values in environmental valuation is a central issue in the valuation debate, a debate that – as stated above – has been very intense in recent years. However, arguments for and against were already clarified in the discussion of Kopp (1992) versus Rosenthal and Nelson (1992). Rosenthal and Nelson strongly object to including non-use values in CBA, while Kopp favours their inclusion.

Rosenthal and Nelson object for three essential reasons:

- First, non-use values are not properly analysed in an economic analysis to resolve matters of ‘cultural symbolism and social ideology’; that is, matters concerning broad welfare issues such as “fundamental value disagreements on the proper relationship between humans and nature” (Rosenthal and Nelson, 1992, p. 117).
- Second, other tangible or intangible goods have a limitless range of non-use values and because there is no practical way to limit this range, CBA analysts will have an infeasible amount of work.
- And third, even if one could somehow limit the non-use values to be assessed, there is no accurate and reliable way to measure them, as outcomes depend too much on the context in which they are requested.

The first argument relates strongly with the welfare theoretical base already discussed – and to the main conclusions drawn about it. Does the welfare criterion in a practical CBA analysis really include any broad value or utility, or is it restricted to more everyday economic issues and economic efficiency? Rosenthal and Nelson seem to agree with the discussion of the distribution earlier. “Historically economists have usually regarded the basic values of society – such as the distribution of income (which could also be said to have existence value) – as outside the scope of the economic efficiency analysis” (Rosenthal and Nelson, p. 117). Kopp argues for the opposite: “If, as economists, we can treat pure public goods within the welfare theoretic structure, we can treat existence value within the same structure” (Kopp, 1992, p. 123).

The second argument is conceptual and directly linked to tables 3.4 and 3.5. Anything and everything can have non-use value: the family farmer, the local neighbourhood, Ford Mustangs, and there is no limit to existence values that can be identified. People may be very happy to know that family farmers exist, even if they do not buy their products, they may appreciate their neighbour’s garden while not paying for it, and they may be pleased to know that Ford Mustangs are produced even if they do not drive one. Rosenthal and Nelson argue that the attempt to include these elements in CBA is beyond the workload options of a CBA analyst.¹⁰⁸ Kopp sees the workload argument as “a weak reed upon which to base a rejection of non-use values” (Kopp, 1992, p. 127); if these non-use values exist, the economist should not deny his duty to inform decision-makers about the size and distribution of these values.

The third argument is more practical: there is no reliable measurement technique for existence

car. Perhaps at this point of the comparison it would be more fruitful to compare the bequest value of a group of cars (e.g., a specific type of car) with a forest (a group of trees).

¹⁰⁸ Compare Pierrehumbert (2003, p. 263): “Keep your pencils sharp because your answer to this and *a few thousand other questions* {italics inserted} ... will determine whether it’s worthwhile to halt global warming.”

values; this measurement issue is certainly relevant to the other arguments too. CBA literature acknowledges the importance of distribution matters, but has not found a satisfactory way to measure it. The options for including the seemingly limitless range of existence values in CBA is at least partly dependent on measurement options as well. Therefore, before proceeding, it may be useful to step back and examine the entire range of possibilities for measuring environmental values: the toolkit of the CBA analyst.

3.5.3 Valuation techniques: Classification

For the valuation of environmental goods, several approaches have been developed. These techniques have no relation to deep ecology (see above). As Bateman stated, they are about human preferences and the aim is to “produce measures that are *directly comparable* with the values of marketed goods” (Bateman, 1994, p. 48-49; italics inserted).

Different classifications of valuation approaches are possible (Hanley et al., 1995, Hoevenagel, 1994b; Bateman, 1994; Howarth et al., 2001)¹⁰⁹ and different authors classify the available methods differently. Here Garrod and Willis (1999) are followed in giving an overview of valuation methods; they distinguish between revealed preference and expressed (or stated) preference techniques.

Revealed preference techniques (3.5.4) start from existing market data and existing market behaviour and try to facilitate these data to reveal intensity of preferences for different environmental situations. Within this group Garrod and Willis identify several techniques: market price and cost measures of value techniques (3.5.4.1 to 3.5.4.4), the travel-cost method and the hedonic price method (3.5.4.5).

Expressed (or stated) preference techniques ask individuals explicitly how much they value an environmental good. The techniques can be captured under one name: Contingent Valuation (3.5.5).

The discussion of valuation techniques for environmental impacts concludes in section 3.5.6.

3.5.4 Revealed preference techniques

3.5.4.1 Dose response and human capital

3.5.4.1a General description

The method tries to value the environment through the effect that changes in the amount of pollution (the ‘dose’) have on the level of an economic activity or consumer utility (the response). Imagine a tonne of SO₂ emissions per year and the specified amount of damage to crops, buildings or individual health it causes. If one multiplies the damage by the market price a monetary value can be derived. The dose response has potential for wide application, but its main difficulty is complexity or lack of clarity of many dose response relationships (Bateman, 1994).

The human capital approach can be seen as a special case of the dose response approach; it concentrates on the impact of adverse environmental conditions on human health and on the costs this involves. These costs are measured by income lost through illness and expenses for medical treatment.

¹⁰⁹ See De Boer et al., 1997 for a Dutch overview of techniques.

3.5.4.1b Assessment of pros and cons in the context of accounting for sustainability

The dose response and human capital technique directly measures the impact on aggregate market values central to any CBA. So the comparability with other elements of the CBA seems to be optimal. However, there are two major limitations; first, it is often difficult to establish a relation between the dose and the response; and second, non-use values are not valued.

*3.5.4.2 Effect on production or opportunity cost**3.5.4.2a General description*

Environmental regulation may put constraints on the production process and hence influence production costs or output. The 'effect on production' approach measures the change of the value of the output and can estimate both positive and negative effects. It is often used to buyout production rights (e.g., agricultural or mineral extraction) when they suffer from environmental regulation. The method can and has been applied to a wide variety of valuations of environmental goods. Its main limitation (like above) is the complexity of the relation between environmental conditions and production. Garrod and Willis (1993) mention examples of this complexity: Emissions of sulphur dioxide (SO₂), nitrogen oxides (NO_x) and ammonia lead to acid rain and subsequently to economic losses in timber production in Europe. Low-level ozone (O₃) concentrations from transport emissions that lead to reduced crop yields is another example. Estimating characteristics of these types of relationships can be complex both regarding the physical level and the economic valuation level.

The opportunity cost approach resembles and is sometimes equivalent to the 'effect on production' approach. Opportunity costs are benefits foregone of different alternatives. The opportunity costs of a nature conservation scheme may be the losses to agricultural production; this approach is very useful when the environmental benefits of a change are hard to enumerate: one turns to the benefits foregone. The benefit foregone is often valued at market prices. The opportunity cost approach has been used to value a wide range of environmental protection measures.

3.5.4.2b Assessment of pros and cons in the context of accounting for sustainability

The conclusion on the use of these techniques does not depart much from the dose response approach discussed above; there is once again the positive of comparability and the two negatives of complex relations and non-estimation of non-use values.

*3.5.4.3 Mitigatory and preventive expenditure and averting behaviour approaches**3.5.4.3a General description*

These approaches value non-marketed environmental goods through the amount that people spend on market goods and services: that 1) mitigates environmental circumstances and impacts (e.g., water filtration systems that mitigate water quality); that 2) prevents them from occurring (buying of bottled water to prevent drinking polluted public water); or that 3) averts exposure from them (people may move elsewhere and avert the impact of traffic noise, for example).

3.5.4.3b Assessment of pros and cons in the context of accounting for sustainability

These approaches seem to have the same benefit as those discussed earlier, in the sense that they relate directly to the aggregate market values used in the economic part of CBA, and are

therefore comparable. From the examples above we see that the costs of the actions taken do not necessarily reflect the size of the welfare losses suffered, as the palette of *possible* actions may be rather poor. These techniques are only relevant when there is a direct and clear link between the environmentally unfavourable situation and the options for behaviour. Furthermore – again – one cannot estimate non-use values.

3.5.4.4 Replacement cost

3.5.4.4a General description

‘The replacement cost approach assesses the value of a natural resource by how much it costs to replace or restore it *after it has been damaged*’ (Garrod and Willis, 1999, p. 39; italics inserted). Above it was noted that different authors classify the available methods differently and a confusing amount of synonyms exist for – nearly – the same approaches under this heading.¹¹⁰ It seems that elimination cost methods, cost-effectiveness or shadowcost approaches (to achieving a given level of environmental improvement) should fall under this heading. The replacement cost approach in several forms has found application in both the US and Europe. Garrod and Willis mention it as the basis of the CERCLA act in the US, which requires full restoration or replacement of natural resources caused by pollution incidents, but also as having been used in Europe to value the biodiversity targets under the Rio Convention and the drainage function of canals. In the Netherlands, the nature compensation instrument (introduced in LNV/VROM, 1993) to repair for nature losses due to infrastructure development, can also be seen as closely related to such an approach.¹¹¹

3.5.4.4b Assessment of pros and cons in the context of accounting for sustainability

The replacement cost method of valuation is a rather practical and down-to-earth approach through its cost focus. These costs are obviously market-related, which could be a positive situation for comparability, but one should not jump to conclusions.

Replacement costs as such are not clearly related to value. Generally speaking costs will rise the more of an environmental asset is restored; but the valuation of extra restored units will tend to fall, the more we restore. This is why the classification of Garrod and Willis of replacement costs under the heading of revealed preferences may be somewhat awkward: replacement costs do not reveal any obvious preference value (compare Bateman, who places this approach under a different category of non-demand curve approaches).

If complete restoration is possible clearly both use and non-use values are covered/restored. However, we do not yet know whether full restoration is more costly or cheaper than the values attached to the asset. Only when a socially optimal amount of restoration is chosen, only then are the cost of restoration a good proxy of the value of the damage. However, another methodology would be needed to assess this.

If incomplete restoration or some form of compensation is the only alternative,¹¹² then the

¹¹⁰ Compare De Boer et al. (1997); see further Vögtlander (2001) for an innovative approach in this field starting at the business level.

¹¹¹ For instance, concerning the construction of the High Speed (train) Line Amsterdam-Antwerp/Brussel, in case of inevitable damage to nature, ‘[t]he disappearance of these natural elements will be compensated by creating new natural areas elsewhere: for each tree that is cut down, a new tree will be planted elsewhere, or a public garden will be laid out.’

(http://www.hslzuid.nl/hsl/uk/lijn/Harmonizaton_with_the_surroundings/Environmental_compensation/index.jsp)

¹¹² One must keep in mind that environmental assets can have many functions that may require compensation (see de Groot, 1992; and Costanza, 1997).

uniqueness of the environmental asset or environmental quality is especially relevant; the more unique the asset or quality, the less feasible compensation becomes¹¹³.

Finally, it seems that valuation through compensation or forms of replacement often implies the introduction of compensation projects. Therefore, given the fact that no true valuation can be given, the analyst may also choose to introduce new specifications of project alternatives, which include these compensation plans.

3.5.4.5 Travel cost method and hedonic pricing

3.5.4.5a General description

With the travel cost method, valuation is inferred from travel costs made. This method is used mainly to estimate the demand curve for recreation sites. Entry into many recreation sites is free of charge but people do incur expenses by travelling to the site; they incur both monetary expenses (e.g. gasoline) and time and effort 'costs'. Both individual demand curves and demand curves for regions or zones at different distances from the site can be calculated, although this calculation is a far from simple task. Among other things, account has to be taken of the quality of the site, of alternative sites that could be visited and of household and income characteristics. Account also has to be taken of whether a visitor is a pure visitor, a transit visitor also going elsewhere, or a wanderer who simply enjoys the journey. The valuation is an estimation of use value. The technique is best applied to well-defined recreation sites, or to well-perceived attributes within such a site.

Valuation with the hedonic price method is inferred from market prices for goods for which environmental quality is one of the attributes comprising the value of the good (e.g., housing). Houses are valued because they provide living space, because of their architecture, and their distance to workplace, etc. Some of these attributes may be related to environmental qualities: e.g., situated close to nature areas. In analysing price differences between (comparable) houses with and without the environmental quality, one may derive a value for that quality.¹¹⁴

The hedonic pricing technique can also be applied to wage rate differences. Choosing jobs and a wage that goes with it may be influenced by the recreational or leisure amenities in the neighbourhood; but most often this technique is used for labour with health risks. The higher wages in high risk labour situations are taken as a basis for the valuation of health by employees.

3.5.4.5b Assessment of pros and cons in the context of accounting for sustainability

The travel cost valuation techniques have a circumscribed use. Both travel cost and hedonic pricing have problems identifying environmental values from other (for example, economic) values. Travel cost valuation uses travel costs made by visitors as an estimate for the environmental value of a recreation site; estimation is not straightforward, as there may be numerous reasons for visiting, and the value only relates to direct use value. Hedonic pricing has some scope for wider applicability but it also suffers from empirical difficulties in separating the environmentally-related causes for price differences from other causes, and also

¹¹³ Compensation in one form or another will often require so called 'weak' sustainability – allowing trade offs between different kinds of natural and man-made capital.

¹¹⁴ See Elhorst et al., 1999 for an example using real estate brokers as intermediaries for estimating consumer values.

only relates to use values.

One problem that may complicate the empirical estimation of representative indirect use values with these techniques is selective sorting, e.g., people strongly disliking travel noise probably do not live near highways.

In summary, the travel cost method and hedonic pricing technique have limited applicability, although the latter's applicability is somewhat broader. Neither technique covers non-direct use values.

3.5.4.6 Conclusions on revealed preference

The revealed preference valuation techniques above, both those that Garrod and Willis group under 'market price and cost measures of value' and those of the travel cost and hedonic pricing, share many of the same characteristics. They can produce value measures directly comparable to the economic part of CBA. However, their overall applicability in the sustainability context is limited, mainly due to the complexity of establishing a clear relation between environmental circumstances and economic market variables. Furthermore, even if this relation is *sometimes* clear, the highly relevant non-use values are not estimated.¹¹⁵

3.5.5 *Expressed (or stated) preference: Contingent Valuation Method (CVM)*

3.5.5.1 General description

In the discussion on valuation of the environment, the CVM methods are probably the most popular.¹¹⁶ CVM asks individuals to value or to act, *contingent* on the occurrence of a hypothetical situation. Individuals are asked directly, for instance, about their willingness to pay for the creation of a national park or the improvement of an environmental situation. The reason for its popularity seems to be that CVM is in principle capable of capturing non-use values so relevant to the sustainability context.

CVM is usually performed with questionnaires, either through personal interviews or through mailings. At first, CVM interviews were mostly open; they asked people for their maximum amount to pay or to accept.¹¹⁷ Since the mid-1980s however, CVM studies have used discrete choice models, which allowed for the acceptance or rejection of specified options.

Recently CVM is also performed in the choice experiment form: valuation through asking respondents in an experimental setting to value attributes in relation to price differences. (Hanley et al., 1998). The most common approach is to present several alternatives to individuals and ask them to choose which one, if any, they most prefer. Single or repeated choices from various (sets of) alternatives reveal the trade-offs an individual will make between attributes. This latter conjoint analysis method is, if it is compared to the 'single-shot'

¹¹⁵ The revealed preference methods cannot value all environmental goods, as there is, "by definition, no related market good for the mere existence, as distinct from use" of environmental assets (Garrod and Willis, 1999, p. 126).

¹¹⁶ Compare Smith and Pattanayak, 2002 who speak of a "flood of [CVM] estimates" (p. 291). Use has been greatly stimulated in the United States since the 1993 advice of the NOAA panel on contingent valuation of which the economists Arrow and Solow were part; see Clinch and Murphy (2001); Poe et al. (2002); Niewijk (1994); Ruigrok (1999).

¹¹⁷ It is often overlooked in CV surveys that projects have winners and losers, and that 'negative bids' (positive bids on the situation not occurring) should often be a possibility too. See Clinch and Murphy (2001).

approach of CVM, essentially a structured method of data generation. The carefully designed choice tasks help reveal the *factors influencing choice*. Interestingly, in this choice experiment form, CBA shifts towards MCA.¹¹⁸

CVM in either form has wide applicability, as in principle many (if not all) environmental assets can be valued in this way.

3.5.5.2 Assessment of pros and cons in the context of accounting for sustainability

3.5.5.2a *Despite high hopes, CVM does not seem to provide reliable non-use values*

Thus it seems that a technique has not yet been found that is able to measure non-use values. In the debate Kopp versus Rosenthal and Nelson mentioned above, the latter raise four problems related to CVM that seem to cover the essentials of the prolonged debate on CVM:¹¹⁹

- 1) knowledge about the good in question – respondents can only value what they know exists
- 2) the context in which the good is described – valuations differ strongly depending on the context in which they are described
- 3) embedding – valuation of elements, e.g., one or two separate streams, of a whole river delta, is tricky¹²⁰ and
- 4) symbolism – in valuing one river it may stand as a symbol for all rivers, or for the entire natural environment.

Kopp asserts that the first critique is unsound. Why should one have to know the name of the fish before suffering a loss when made aware of its death? The second and third only reflect natural situations. Of course, Kopp asserts, valuation is context-dependent, and of course embedding does take place: one should be alarmed if it did not. The fourth critique can, according to Kopp, be taken into account by a proper design of the valuation.

In line with Kopp's position, recent analyses have shown increasing use of CVM in the sustainability context.¹²¹ Although many authors apparently think that CVM is a useful tool, mounting evidence is emerging that is critical of the most promising feature of CVM: its potential to capture non-use values; this turns out not to be as reliable in practice.¹²²

Greater evidence shows that when relatively complex information is required, and when mainly non-use values are involved, CVM does not yield valid results.¹²³ In his thesis on the scope and

¹¹⁸ See Urli and Nadeau (1999) for a perspective on conjoint analysis in the MCA field. Hanley et al. (1998, p. 426) mention the tension between the stated preference outcomes from this type of method and the revealed preference approach.

¹¹⁹ Compare for instance Vatn and Bromley (1994); Clinch and Murphy (2001) on problem 1; Brouwer et al. (1999); Bulte and Van Kooten (1999); Common et al. (2000); Smith and Pattanayak (2002) on problem 2; Niewijk (1994); Poe et al. (2002); Holmes et al. (2004) on problem 3; Ashby (1980) and De Groot (1992) on problem 4.

¹²⁰ Compare Holmes et al., 2004 for an example.

¹²¹ See Poe et al., 2002 for a short historical introduction; compare Smith and Pattanayak 2002; Brouwer et al., 1999.

¹²² Again here, like in CBA development in general (see above) a divergence between the popularity of practical application and the scientific profession can be seen (Clinch and Murphy, 2001).

¹²³ See Diamond (1996) for an attempt to check *consistency* for these inherently difficult to validate estimates. Compare Wierstra et al. (2001) who show for different environmental amenities that the more non-use values are relevant and the more complex they are, the less suitable is CVM.

validity of CVM, Hoevenagel concludes that the method will produce its most valid outcomes if “familiar, small, short-term, reversible environmental goods are involved” (Hoevenagel, 1994c, p. 219).¹²⁴ In the accountable-sustainability context the standardisation and measurement, which is easily understood by different stakeholders is crucial¹²⁵. Furthermore, it may be emphasised that transferring estimates from one evaluation to another is often a necessity because of the high costs of setting up CVM studies (Boardman et al., 2001, p. 382; compare also Costanza et al., 1997).¹²⁶

One thing that conflicts most with these demands is the context specificity of calculated values. This context specificity proves to be a greater problem than many analysts (like Kopp above) had thought. The values derived for environmental goods seem to be valid in a specific context, but it has become clear that the values cannot be easily transferred from one site to another, or from one policy context to another (Garrod and Willis, 1999, p. 12; O’Connor, 2000, p. 172).¹²⁷

This general lack of transferability strikes at the heart of the individual estimates. According to Brouwer (2000, p. 137) “the problem [of environmental values transfer] is much more fundamental than previously acknowledged” (Brouwer, 2000, p. 137). “How can environmental values be reliably predicted across sites and people if currently much, if not most, of the variability of the values in original studies cannot be explained?” (Ibid, p. 143; Spash, 2000) Brouwer stresses the point that simply taking preferences as given, and expressed valuations with them, is not good practice.¹²⁸

The point here seems to go quite a bit further than mere ‘context sensitivity’; the meaning of the expressed values is unclear, as is how they should be interpreted, and how they can be compared to market based values of other impacts.¹²⁹ Some authors argue that CVM responses are better understood as expressions of attitude than as indications of economic preference (Moss et al., 1996; Kahneman et al., 2000; Niewijk, 1994). Seeing them as attitude expressions helps to explain the mentioned anomalies or extreme context-sensitivity in CVM responses (Spash, 2000). The sensitivity is easily understood once it is acknowledged that preferences in many of the complex and unfamiliar choice situations at hand are imprecise and open; only a more basic ‘attitude’ remains.

The dual consumer/citizen preference structure may be a useful way of clarifying the problem, because it may be regarded as the crux of the four problems identified above by Rosenthal and Nelson (Blamey et al., 1995). The heart of the matter seems to be the lack of comparability of inadequately monetized citizen related values with the other: consumer related monetary values

¹²⁴ See also Veisten (2004) and Hoevenagel (1994a).

¹²⁵ Compare Jepma and Munasinghe (1998, p. 129) in their book on climate change: ‘Furthermore, the valuation [of environmental effects] should be based on a reasonably well founded methodology, since speculative assumptions could confuse decision makers’.

¹²⁶ Interestingly, this fact as such resembles CBA practice (regarding economic impacts) of using mainly aggregate data, while its theoretical base often suggests a more individual approach (see above).

¹²⁷ On the (limited) possibilities of so-called benefit transfer see also Bergland et al., 2002; Engel, 2002; and Bateman et al., 2002. For the more general picture on meta-analysis and value transfer, see the complete volume: Florax et al., 2002.

¹²⁸ Interestingly, from the perspective of the coming chapters 4 and 5, Brouwer expects much from more stakeholder involvement and a multi-criteria decision-making process.

¹²⁹ Compare Ritov and Kahneman (1997, p. 47-48): “...we must understand how people value public issues in general and environmental issues in particular before we can provide useful answers to the question of how values should be measured.” See also the complete volume of Bazerman et al., 1997. Murray et al., 2000 see a somewhat comparable situation in large numbers of low level cost-effectiveness analyses on health interventions (p. 236).

in a CBA.¹³⁰ The reader may recall that the challenge of environmental valuation was formulated as: “producing measures *directly comparable* with the values of marketed goods.” If the best explanations of occurring anomalies point to the dual preference structure, then in this respect, CVM techniques can hardly be seen as successful.

However, in the light of the dual preference structure, recent developments in CVM can be clearly understood and appreciated. The choice experiment form for CVM is logical, as it aims *to discover* the structure of the preferences. Developments around provision point mechanisms¹³¹ show that in CVM, reflection about the possible overall outcomes (such a reflection is of course a citizen concern) may be fruitful in trying to bridge the gap between both types of preferences.¹³²

3.5.6 *Would MCA not provide a better starting point than CVM?*

As a general conclusion in the sustainability context unclear measurement still largely hampers useful interpretation of CVM results. Although CVM is the only valuation technique that can potentially measure non-use values, this proves to be far more problematical in practice than many think. Quite like the result of the measurement of distributional issues, clear (read: comparable) measurement of non-use environmental issues proves to be difficult, if not impossible, in many cases.¹³³ Outcomes can often only be interpreted and used in the context they were made, and even then many unexplained anomalies still exist.

Measurement techniques are continuously improving.¹³⁴ The choice experiment forms and provision point mechanism are useful developments that may help solve some problems as they move CBA in the direction of ‘discovering’ preferences and assessing individual preferences under different social context alternatives. But lack of clarity among other things remains about how to limit the range of existence values for future assessment – on the everyday economic side too. Comparable measurement will thus continue to be problematical in CBA.

Because of these difficulties with CBA in the sustainability context, the question now arises as to whether or not a more explicit value-disclosure technique such as MCA will provide a better starting point. Chapter 4 will provide more insight into this.

¹³⁰ As Blamey et al. (1995, p. 286) notice, CV studies may be interpreted as surrogate referendums, but this makes the question of why one would seek this type of response through this type of means even more urgent.

¹³¹ See Poe et al., 2002. A provision point – in this CVM context – is a minimum level of aggregate contributions below which the public good is not provided. In CVM a money-back guarantee can be included if the provision point is not reached. If there is an ‘overshoot,’ part of the contribution can be refunded or more similar public goods/services can be provided.

¹³² See for instance Clinch and Murphy (2001, p. 429) for bad CVM results due to substantial disagreement (29%) with a proposed subsidy scheme.

¹³³ Compare Connor (1999) about non-monetizable impacts and sustainability.

¹³⁴ In their discussion of the future of all the valuation techniques, Garrod and Willis see two major trends. The first trend is towards increasing complexity, towards which they are somewhat critical. They see that the complexity of the estimation techniques and experimental designs will continue to increase in the coming years. This is only partly a response to the problems encountered in environmental goods valuation but also partly “a reflection of the market for papers in distinguished academic journals” (Garrod and Willis, 1999, p. 216). Apart from this increasing complexity, they see more attention being paid to the decision theoretic framework underlying expressed preferences, a trend they are convinced will be more fruitful and more likely to enhance the credibility of these techniques (outside academic and expert communities). The current thesis may be seen as part of this trend.

3.6 Conclusions: CBA's performance in a sustainability context

At the end of section 3.2 conclusions on CBA performance in the sustainability context were grouped and divided into some conclusions that are clear and others that are unresolved. By now the whole chapter can be reviewed and all conclusions about CBA performance in the context of accounting for sustainability can be shown (see Table 3.6 below).

Generally speaking, the new conclusions concerning the issues further discussed in sections 3.3 to 3.5 show that CBA is primarily about efficiency in a market related context and that an ethical consensus (read: judgement) among stakeholders is the final basis for using CBA (section 3.3). Fundamental issues including the distribution of income and power in society are taken as given in CBA, but cannot be explicitly valued (section 3.4). Technically simple attempts to work with weighting of impacts to different groups have suffered from a telling lack of consensus. Valuation of environmental issues is often problematical since non-use environmental impacts continuously prove difficult to catch in CBA, let alone that they could be comparably measured to economic development impacts.

These conclusions will be used in chapter 5, but for now two points are noteworthy. First, the discussion has clearly shown where CBA strengths and weaknesses lie when it has to function in the accounting for sustainability context. Second, a major question arising from CBA attempts to address environmental non-use values in the sustainability context is: Would a more explicit value-disclosure technique such as MCA not provide a better starting point than CVM?¹³⁵ This question seems to be a good reason for turning to chapter 4.

¹³⁵ See for instance Joubert et al., 1997.

Table 3-6 Conclusions on CBA's performance in the sustainability context (per stage)

Conclusions CBA
Stage one: Define the project (or policy) and impact population
(1) CBA silent on incorporation of newly emerging alternatives (2) Measurement at global scale is natural to CBA; at lower – explicitly defined scales – often practiced too, although not always easy
Stage two: Identify project impacts
(3) Aggregate market values are important and consistent everyday economic realities (4) Consensus among stakeholders is the final basis of CBA (5) Broad (subjective) welfare economics is as much at the basis of CBA as of MCA
Stage three: Quantify relevant impacts physically
(6) Empirical facts about Triple E impacts and their relations become clear, just like the main areas of uncertainty
Stage four: Value relevant effects monetarily
(7) CBA is mostly about efficiency (8) Framing WTP questions is a subtle task (9) WTP as a shortcut in the sustainability context meets areas of hesitation and suffers from too high context sensitivity (10) People hold a dual consumer-citizen preference structure; WTP estimates are most valid in the single preference structure of the market related context (11) WTP values for everyday economic objects are as context sensitive as many other preferences. Using these WTP values in long-term or global assessment values requires caution as their use is 'stretched' towards 'citizen' preferences (12) The causal relations between markets makes the combination of economic understanding and monetary valuation a powerful evaluation tool (13) In the absence of market imperfections, externalities and public goods the causal – supply chain – relations on markets simplifies evaluative measurement as direct impacts shown on markets can be measured instead of a great many indirect effects (14) Although without 'market imperfections' CBA is largely unnecessary, the more perfect competition is, the better is CBA's monetary measurement capability (15) Environmental valuation is about anthropocentric value (16) In environmental valuation non-direct use values are central, but they complicate comparability among Triple E aspects (17) Certain environmental valuation techniques can give comparable monetary estimates if complexity is not a problem; but non-direct use values are not covered. Even CVM, despite the high hopes of some, does not seem to provide reliable non-use values (18) A more explicit value-disclosure technique (MCA) might provide a better starting point than CVM
Stage five: Discount costs and benefits
(19) Possibilities for discounting different Triple E impacts differ – market related impacts have a natural reference point that other impacts lack (20) The discounting debate is really about limitations to non-market valuation in CBA
Stage six: Apply the Net Present Value test
(21) Calculation of NPVs without difficulty, interpretation requires caution (22) For distributional impacts <i>distributional weights are sometimes considered</i> , but careful description seems to be the best option, with standardised attention to the extremely poor
Stage seven: Perform sensitivity analysis
(23) Sensitivity analysis may be very rewarding in a conflict setting context (24) Uncertainty is not always best treated as footnotes to the main results; uncertainty may turn out to remain a central result

Appendix to chapter 3 Illustration of pitfalls of environmental valuation - The Costanza attempt

Introduction

This section illustrates how problematical environmental valuation may be in the accountable-sustainability context and which pitfalls should be avoided. It analyses the heroic attempt by Costanza et al. (1997) to value the world's environmental goods and services monetarily, and the subsequent critique it has aroused. The authors thought that valuation was necessary to acquire an adequate weight for environmental values in decision-making, and they also thought their results could be used by other analysts without the necessity to conduct separate valuation studies.

The attempt is very interesting as it acts as a brief but powerful illustration of the issues in monetary valuation in the sustainability context discussed above. Three reasons are noteworthy:

- First, the valuations of Costanza et al. explicitly aim at selection for project evaluations where loss of ecological services must be weighed against other benefits (Costanza et al., 1997, p. 259).
- Second, the attempt aims at establishing values at a global level. The importance of this point should not be underestimated. It was said in section 3.3 that aggregate market values are a consistent whole, which greatly facilitates the interpretation of small-size changes. For environmental valuation to be made manifest at the small-scale project level, this same situation – having a global consistent ‘whole’ to interpret size – seems indispensable (Brouwer et al., 1999).
- Third, the interdisciplinary character of the attempt closely resembles the setting of accountability to broad groups of stakeholders. In this setting conceptual clarity is crucial but, as the debate shows, it proves very difficult to establish.

The Costanza study was published in *Nature* and aroused worldwide attention and debate. The scientifically most comprehensive debate on the merits of the Costanza attempt was published in a special issue of *Ecological Economics*, in which the most important aspects for our purposes are discussed below.

The method of the Costanza attempt

The Costanza et al. study attempted to estimate the monetary value of the services of the world's ecological services and the natural capital stocks that produce them (Costanza, D'Arge et al., 1997). Following the logic of the assumed welfare theoretic base of CBA, these services “contribute to human welfare” and represent part of “the total economic value of the planet.” The authors used mainly secondary sources: over 100 CVM and other valuation studies – mostly at lower scale levels. They estimated the current economic value of the various ecosystems services as a flow expressed in a US-dollar value per hectare per year, and explicitly aimed at estimating ‘marginal values’, directly comparable to market values.

A total number of 17 ecosystem services were identified: gas regulation, climate regulation, nutrient recycling, food production, raw materials, recreation, and so on. Services were estimated for 16 biomes including oceans, coastal zones, forests, wetlands, cropland, and urban areas. The values were mostly calculated by using previously published valuation studies. Various valuation methods were used “to estimate both the market and non-market components of the value of ecosystem services.” Estimates focussing on damage costs, or repair costs were used, but many of the methods were based “either directly or indirectly, on attempts to estimate ‘the willingness-

to-pay’ of individuals.” (Costanza et al., 1997, p. 255). The importance of the total economic value – linked to the consumer surplus notion – was underlined in the article by discussing supply and demand curves for ordinary goods and environmental goods.

The results of the Costanza attempt

The monetary valuation results of the Costanza attempt are as follows. The value of the ecosystems services was found to be mostly outside the market. The total yearly value of all services was estimated, as a minimum, to be in the range of US\$ 16-54 trillion (10^{12}), or on average \$ 33 trillion per year. Costanza et al. directly compared this value to the gross world GDP of US\$ 18 trillion per year.¹³⁶ Roughly half of the ecosystem services value came from nutrient recycling. Second in the row are cultural services with some 10% of total economic value. Looking at the biomes it could be seen that roughly two-thirds of the value comes from marine biomes (e.g., open oceans) and one-third from terrestrial biomes (e.g., wetlands and forests). Values per hectare per year are about US\$ 600 for marine areas and about US\$ 800 for terrestrial areas.

The attempt was thoroughly discussed in a special issue of *Ecological Economics*. From all the topics raised in the 12 articles published in that special issue, two elements are the most important for the purposes of this thesis:

- CBA basics: comparability (economics-ecology) and aggregation
- Stakeholder understanding in the (policy) debate: monetary information and other information

CBA basics: Comparability and aggregation

The Costanza paper tried to develop monetary estimates for environmental services that would be comparable to market related data; it compared its outcome to global GDP. An important critique of the Costanza paper is that it showed that the authors misunderstood and incorrectly interpreted the limits to WTP estimates and the meaning of consumer and producer surpluses in the context of environmental goods and services.

One obvious and intriguing puzzle for many commentators was that the apparent Willingness to Pay estimate far exceeds the globally available budget of GDP. Costanza et al. emphasise that the “real value is almost certainly much higher, even at the current margin” (Costanza et al., 1997, p. 259). This result of the WTP based estimate exceeding global GDP underlines a point stressed in section 3.3: that WTP estimates may become difficult to interpret if they depart from everyday economics, or, as the standard CBA textbook states, if the use of consumer surplus estimates is not restricted to ‘ordinary conditions’. Everyday economic aggregates are a consistent whole: “money income is either spent on current needs or saved (i.e., invested). Either way, income equals outgo” (Ayres, 1998, p. 18). “People might be willing to pay more ...but they are unable to do so” (Hueting et al., 1998, p. 34). Adding up partially derived WTP estimates will easily lead to inconsistent estimates, as their size is hard to interpret.¹³⁷ The size of the monetary value may be large, but if one turns to table 3.4 above,

¹³⁶ Later corrected to be around US\$ 25 trillion per year in their base year 1994.

¹³⁷ It is especially tricky when there is an optimum level of provision.

how large would such WTP estimates be for non-use elements of ordinary market goods? What of the WTP estimates of use values for the first slice of bread and the first sip of water? Interestingly, Ayres estimates the value that seems reasonable if it had been part of the consistent whole: that is, if the economic system had a sector with the ability to price environmental services now often offered free of charge.¹³⁸ In such an integrated system the cost of providing the ecosystem services (the supply curve) would also be relevant. In a steady-state situation – and after an initial clean up of perhaps a few hundred billion dollars – Ayres estimated the consistent value then to be a few percent of GDP (Ayres, 1998, p. 19). This is not the place to check the estimates, but the point is clear: “What use are monetary values if the link with everyday economic money is cut?”

Stakeholder understanding in the (policy) debate: Monetary information and other information

In the sustainability context of evaluation accountability to stakeholders ranks high. Daly comments on the difficulty and pitfalls of interpretation of these aggregate Costanza totals based on margins. A reporter asking Daly about the significance of the Costanza study himself commented: “we are a lot richer than we thought, and must think carefully about what we should do with the extra wealth” (Daly, 1998, p. 22). This comment is of course completely opposite to what the Costanza study intended to show, but it clarifies the difficulties one may encounter when explicitly referring to dollar values, which are part of the everyday economic reality where money values can be spent and where having things of high dollar value indicates wealth. The reporter was naturally confused by what the economic literature calls the *Lauderdale paradox*; it shows (as was seen in section 3.3) that only scarce things demand a price: the best things in life *are* free, but not without value. If they increase in scarcity, which is a welfare loss, they may start demanding a price in the market: a public welfare loss may lead to more private wealth. According to Daly, in this interpretation the Costanza estimate “screams at us” how scarce natural capital is becoming (Daly, 1998, p. 22). Daly points to the fact that there are “evident physical consequences of excessive human expansion that scream the same message without need of explicit valuation. But for those who only hear dollars, let us scream now and then in dollars!” Note that according to Daly, the physical consequences are evident, and it seems then that the consequences can be clearly understood by many stakeholders. Norgaard et al. (1998, p. 38) touch on this same point: of whether expressing value in monetary terms is a more effective way enrich people’s understanding of impacts on the environment than, for instance, through narratives. El Serafy holds that “estimation in physical units may be sufficient to bring [the decline of environmental services] to the attention of citizens and politicians, and even economists” (El Serafy, 1998, p. 27).

Toman focusses solely on this crucial point of the benefits of monetisation versus other physical information. In his view, the Costanza paper should be read as expressing the view of the authors that the value of ecosystem services is not only large but underappreciated. However, can this message, wrapped up as it is in aggregate dollar terms, be correctly interpreted by different stakeholders? Toman argues that it cannot. In his view little can usefully be done with a “serious underestimate of infinity.” Such a figure is on equal footing with the ‘default value of zero’ that is sometimes used in CBA for a “difficult-to-measure ecological value” (Toman, 1998, p. 58). These estimates are not linked to “any particular values experienced by particular people in a particular place at a particular time.” Because of this lack of linkage it gives no insight into the directions of current changes in ecosystems and

¹³⁸ Within national boundaries in the economically-developed countries governments or private agencies already play that role to an extent.

the relative urgency of different changes.

This type of monetary valuation of environmental aspects is, according to Toman, not information-rich enough to determine policy choices. Apart from the fact that there is limited understanding of the physical world, Toman points to the difficulties of reliably ascertaining people's preferences concerning trade-offs between economics and environment and changes in these preferences when changes occur in the information available or the social context (Toman, 1998, p. 59). Toman argues that using CBA for the assessment of opportunity costs of environmental protection and even environmental benefits – to the extent that one is really and consistently able to do so – may provide important information and discipline in project-evaluation contexts.¹³⁹ But it will usually not give a complete picture in an evaluation. Where CBA evaluation is lacking, unreliable or incomplete, then at that point non-monetary information should be combined with it. For reconciling competing values, they should be combined in such a way as to facilitate both informed decision-making and accountability. Research should be directed to finding such an information-rich and economically grounded framework (Toman, 1998, p. 59). This thesis can be seen as following this last suggestion.

¹³⁹ An interesting point that draws on several of the points made above is the importance of economic realism and finding new mechanisms of value transfers. Even if our valuations are more correct than Costanza in a technical sense, the calculation of a dollar value does not facilitate a value transfer better than a physical estimate would. The mechanism – whether through a political process or through adoption of market processes – has to be integrated in the action of particular people at a particular place, however. Rees (1998, p. 51) quoted Ophuls and Boyan (1992) for underlining the everyday economic link to environmental decline and the little usefulness that economic valuation of public goods as such may have for it. Regardless of the value of nature services – be it from a global or a non-use perspective – ecological and social disaster may happen “not so much by the evil acts of selfish people as by the everyday acts of ordinary people whose behaviour is dominated, usually unconsciously, by the remorseless self-destructive logic of the commons.” This logic, it should be noticed, is the lack of relation between preferences as an individual/consumer and preferences as a citizen.

CHAPTER 4: MCA IN A SUSTAINABILITY CONTEXT

4.1 Introduction

4.1.1 General

Multicriteria analysis (MCA) is a range of evaluation techniques that tries to help decision-makers in situations when judgement depends on more than one criterion.¹ MCA can be seen as the main judgement-oriented alternative to CBA (Howarth et al., 2001; Merkhofer, 1987). In MCA, the performance of project alternatives is confronted by the objective(s) of the decision-maker(s), stakeholders, or social welfare at large. Put simply, multiple criteria analysis takes a set of alternatives, defines a set of criteria² for decision-making and ways of measuring³ them, and derives weights for the criteria. Alternatives have different scores on the different criteria. Performance scores of alternatives on the different (sub) criteria are aggregated into a total score.

An MCA problem can be expressed in a matrix format: with alternatives as columns and criteria as rows and the scores in the cells (Hwang and Yoon, 1995). This performance matrix is often seen as the core of MCA (DTLR, 2001). “The basic conviction underlying every [MCA] approach is that the explicit introduction of several criteria, each representing a particular dimension of the problem to be taken into account, is a better path for robust decision-making, than optimizing a single dimensional objective function (such as in cost-benefit analysis).” (Bana E Costa et al., 1997, p. 30). MCA belongs to the class of non-monetary evaluation methods (Clímaco, 1994).

The classical example of the MCA approach is Benjamin Franklin’s ‘moral or prudential algebra,’ which is more than 200 years old. Franklin, having to decide upon a difficult measure, wrote down pros and cons in two columns – over the course of three or four days. And “when I have thus got them all together in one view, I endeavor to estimate their respective weights; and where I find two, one on each side, that seem equal, I strike them both out. If I find a reason pro equal to some two reasons con, I strike out the three;and thus proceeding I find at length where the balance lies; and if, after a day or two of further consideration, nothing new that is of importance occurs on either side, I come to a determination accordingly” (Yoon and Hwang, 1995, p. 4).

We can compare CBA and MCA by looking at their historic development. As chapter 3 has shown, CBA comes both from a tradition of trying to determine objectively the welfare impact of projects and from the everyday economics of cost and finance. MCA’s historic development is characterised most by its aim of helping an individual decision-maker (or a well-defined group) in a complex problem setting (Bana E Costa et al., 1997; Nijkamp et al., 1990; Voogd, 1983).

¹ See Bana E Costa et al. (1997) for a useful ‘constructive’ overview of the development of the field.

² Also called attributes or goals.

³ Measurement is usually in different units.

4.1.2 Different approaches to MCA

MCA is explained in several textbooks on MCA⁴ (e.g., Pomerol and Barba-Romero, 2000; Olson, 1996; Yoon and Hwang, 1995; Janssen, 1994; De Vries, 1993; Yu, 1985; Edwards and Newman, 1983; Voogd, 1983; Keeney and Raiffa, 1976). However, MCA seems to be less of a unified technique than CBA. Consequently, there seem to be more options available to the MCA approach as a starting point for the assessment of its performance in the sustainability context. Despite this variety and sophistication, Beinat (2001, p. 51) notices: “Nonetheless, most practitioners still rely on simplified and frequently rough MCA implementations, which are a far cry from the sophistication of the largest part of MCA literature.”

Starting from the core of the performance matrix many different approaches have been developed.⁵ The development of different MCA approaches has depended mostly on the type of information available and the preference structure of the decision-maker. The type of information can be qualitative, quantitative or mixed; it can be certain or uncertain, and so on. The preferences of the decision-maker for instance may be highly structured in some problem areas but hardly structured in others, which may sometimes permit cardinal ratings and other times only ordinal ranking.

Three broad groups of MCA techniques can be identified (following Stewart, 1992 and Stewart and Losa, 2003):

- Goal programming and reference point approaches
- Descriptive methods and outranking methods
- Value function approaches.

Goal programming and reference point approaches are, generally, most suited for situations having close interaction between the MCA analyst and a single decision-maker and have a very large number of alternatives (Stewart, 1992, p. 578 and 584).

Descriptive methods and outranking methods are often used when compensation between performance areas is difficult and when it is doubtful whether a ‘true’ order of alternatives exists which corresponds with the preferences of the decision-maker(s).

Value function approaches are very much judgement oriented. They present a full ranking of alternatives. Furthermore, they probably represent the standard (and most popular) MCA approach. The most widespread value function approach is the additive scoring. Additive scoring means that the performance of an alternative is determined by adding the – standardised and weighted – scores on different criteria. The alternative with the highest overall score is preferred. Additive scoring approaches are generally regarded as very effective and easy to understand (Stewart, 1992, p. 573).

4.1.3 Presenting the essentials of MCA

In choosing what to present as the essentials of MCA, this chapter considers both the results of chapters 2 and 3, and the requirements of a clear presentation. More specifically, given the context of accounting for sustainability and CBA’s performance in this context, the main interest is to find an alternative to CBA that seriously addresses non-use value measurement,

⁴ Or sometimes in books on decision-making like for instance Beroggi, 1999; Hammond et al., 1999. See Kleindorfer et al. (1993) for an overview of different streams in decision sciences.

⁵ Some see this positively as ‘a melting pot,’ others see it negatively as a ‘Babel-like confusion of tongues’ (Clímaco, 1997, p. 3; DTLR, 2001). Urli and Nadeau (1999) argue that the dispersion of MCA into many different disciplines and application domains may imply a ‘loss of unity and identity for the field itself’ (p. 41).

stakeholder involvement and distribution analysis, but remains judgement-oriented.

Goal programming therefore seems to be the least useful in the sustainability context as the goals or performance levels for different sustainability aspects may not easily be determined in advance; the number of alternatives is usually not that large; and accountability seems difficult to achieve in the more interactive forms. Descriptive and outranking approaches may be very useful in the sustainability context, as they can function well in a context in which trade-offs are particularly difficult to make. However the aim is to look for a judgement-oriented approach, these methods are clearly 'second-best'. One might have to settle for second-best in the end, but one has to see first whether a stronger judgement oriented approach could work within the given context.

This chapter will therefore present a value function approach of MCA. It will follow the so-called MultiAttribute Utility Technology (MAUT). MAUT is a central technique within the additive score value function approaches, and seems very suitable to clarify the essentials of the MCA approach. Below the discussion follows the stages of a MAUT-MCA as presented by Edwards and Newman (1982), because they, more than others, address the issue of how to involve stakeholders in an MCA.⁶ Edwards and Newman identify seven stages in an MCA.

- Stage one: Identify objects and function(s) of the evaluation
- Stage two: Identify stakeholders
- Stage three: Elicit and organise value dimensions/ attributes
- Stage four: Assess the relative importance of value dimensions/attributes
- Stage five: Estimate scores of alternatives on low level dimensions
- Stage six: Aggregate scores with relative weights
- Stage seven: Perform sensitivity analyses.

4.1.4 Two-step approach

Below in section 4.2 is a brief general description of every stage on the basis of the existing literature. The description is followed by an assessment of the pros and cons of the relevant stages, with special attention to the demands of the sustainability context.

The assessment of MCA in the context of accounting for sustainability follows the same two-step procedure as in chapter 3: Step 1 – section 4.2 – discusses the stages of MCA and assesses MCA's performance in the sustainability context along the way, provided the assessment is rather clear-cut, and can be handled fairly succinctly. In step 2 – sections 4.3 to 4.5 – the topics of MCA performance that warrant lengthy treatment are discussed. Although the difference between MCA and CBA seems profound, this chapter will meet with some of the same issues discussed in chapter 3. The discussion here can therefore be kept to a minimum.

⁶ See also Niculae and French (2003, p. 196) who note that the need for public participation has become increasingly recognised in the MCA field.

4.2 The stages of an MCA

4.2.1 *Stage one: Identify objects and function(s) of the evaluation*

4.2.1.1 General description

The first stage in an MCA is to determine the objects and the function of the evaluation. In project evaluation the objects of the evaluation will be the project alternatives. MCA does not have the same emphasis on the ‘nul-alternative’, or zero-option that CBA has, but there is nothing preventing the use of it. MCA will normally have several alternatives which are to be compared.

The first stage also addresses the function(s) the evaluation is intended to perform. This reflective stage is unusual in CBA, as it is an ex ante reflection on the evaluation process. It introduces the opportunity for the analyst to make explicit methodological and practical choices that shape the evaluation. The choice of an MCA method is increasingly seen as crucial to successful decision support (see for instance Ozernoy, 1997). This selection process is in itself a difficult multi-criteria problem, which is not easily structured. The development of computer-aided MCA expert systems may help to assure that the choice is made from a large enough set of possible methods (Janssen et al., 2000) compare also Eom and Min, 1999). One disturbing observation is that experiments about the impact of choosing different MCA methods indicates that the choice of the method often determines decision outcomes (Ozernoy, 1997).

4.2.1.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.1.2a *MCA does not stress the need for a ‘nul-alternative’ or zero-option.*

We noticed above that MCA does not stress the need for a ‘nul-alternative’ or zero-option,⁷ but there is nothing against using it. Is there a real or only apparent difference with CBA here? In the project evaluation context as sketched in chapter 2, the difference is not real. One should remember that any evaluation technique strives for clarity about the ‘nul-alternative’, or the determination of the net effect (Rossie et al., 1999, p. 235 onwards). The reason for its extra attention in CBA is probably that because market-values based evaluation implies tracing a causal chain, it often cannot manage without a rather elaborated nul-alternative. Furthermore, MCA is used in more diversified and small-scale settings than CBA. Within the sustainability context, a nul-alternative will also be logical in MCA.

4.2.1.2b *MCA reflects on the functional context and tailors the evaluation to the project*

In the context of accounting for sustainability, some of the important functional and object choices have already been made (see chapter 2).⁸ Beyond that, this stage incorporates an important and conscious effort to justify functional limitations of the evaluation as well as choose a specific evaluation technique.

⁷ Unlike CBA. Voogd (2004, p.223) argues that ‘...the ‘zero’ alternative was always badly defined and never seriously elaborated’ in evaluations in which the goals are not the subject of discussion

⁸ This thesis could even be seen as a *general* pre-selection process within this stage.

4.2.2 Stage two: Identify stakeholders

4.2.2.1 General description

In stage 2 stakeholders are identified. Stakeholders are defined as individuals or groups that have an interest or stake in the evaluated project alternatives, and who are important enough that their interests be considered. The stakeholders' main role is to identify and structure the value attributes of the evaluation; that is, stakeholders should provide and structure the judgement criteria for the evaluation. Stage 3 will explore this further, but first we address the question of how to identify stakeholders.

We observed in chapter 2 that it is common practice in any evaluation setting to identify potential stakeholder groups. In the context of accounting for sustainability, the number and size of the stakeholder groups will be higher, compared to other contexts.

Edwards and Newman try to limit the potentially long list of stakeholders by distinguishing between stakeholders with some decision power and those who are only affected by the decision. Although some interaction may be necessary with the second group to at least represent their values well, they claim that the evaluator's main concern should be on the first group; stakeholders that are "likely to use evaluation in making decisions."⁹ This group could consist of people who have some (indirect) influence on the project, be they primary decision-makers on the project, sponsoring agents, higher hierarchical levels in government or private decision-making groups, the office that is funding the evaluation, etc. The evaluator should pay careful attention to the first group and should be in extensive interaction with them. Both identification and interaction with stakeholders may require extensive knowledge and political skills of the evaluator. Although Edwards and Newman try to limit the potentially very long list of stakeholders, their commitment to involve as many relevant groups as possible is strong: "whatever limits money, time and co-operativeness make necessary, it is always better to include too many stakeholders than too few" (p. 35). Even if the analyst omits several stakeholders, the list will nevertheless be unwieldy and difficult choices will have to be made about who to include and who to exclude.

4.2.2.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.2.2a *Involvement of stakeholders makes the analyst's political and communication skills a critical factor for successful evaluation*

Contrary to CBA, MCA stresses the importance of active interaction between the evaluator and the most relevant stakeholders, for the specification of the judgement criteria. Organising fruitful interaction with the selected stakeholders depends largely on the evaluator's communication and analytic skills. If evaluators do not have adequate knowledge and skills to identify stakeholders and interact with them, the quality of the evaluation will suffer.

4.2.2.2b *Involvement of all relevant stakeholders may be nearly impossible*

Identification and limitation of the relevant stakeholders and their representatives is not a simple task if the list of (potential) stakeholders is very long; therefore in the evaluation context of accounting for sustainability, choosing which stakeholders should be involved is rather complicated. Any affected population should be represented, but the 'population' could

⁹ This claim is made more generally, for any type of evaluation, by Patton (1997). Patton argues for only involving people who are really interested, instead of trying to serve anonymous audiences.

very well be spread around the world. If the project is a topic of debate, then both sides of the debate should also preferably be represented; this will undoubtedly give stark differences in values from the start. Compared to CBA's stance of choosing a – high spatial level – impact population, the MCA approach may be seen as 'biased' towards nearby and 'involvable' stakeholders to the neglect of faraway impact populations.

The problem of stakeholder involvement in MCA will be addressed in greater depth in section 4.3.

4.2.3 Stage three: Elicit and organise criteria

4.2.3.1 General description

The next step in the analysis is probably the most important and simultaneously the most complex step in MCA (Edwards and Newman, 1983, p. 51): eliciting value structures from stakeholders. This step departs strongly from CBA theory and practice.

In an MCA the judgement criterion is not given in advance,¹⁰ but it is determined as an integral part of the analysis. The analyst has to define criteria, on the basis of which judgement is passed.

What does the MCA literature say about the characteristics of criteria? Bouyssou (1990) calls a single criterion a tool, which allows the comparison of alternatives according to a viewpoint or, from a more operational standpoint, a particular significance axis. In any specific project evaluation there will be several criteria (Saaty, 1980). With regard to group characteristics of the criteria in an MCA, Keeney and Raiffa (1976) are often cited (see for instance Edwards-Jones, et al., 2000, p. 156). They argue that the set of criteria in a multi-criteria analysis should have the following characteristics (Keeney and Raiffa, 1976, p. 50-53):

- Complete – There should be no further criteria which can be used to judge between the project alternatives
- Operational – Each criterion must be capable of being measured in some significant way
- Decomposable – A complex decision requires both value judgements and empirical estimates about many things. In an MCA the 'many aspects' must be decomposed (and often ordered hierarchically) to limit the complexity (Olson, 1997). If the criteria are decomposable, value judgements and measurements can be made in small steps of limited complexity
- Non-redundant – No aspect of the problem should be accounted for more than once
- Minimal – No smaller set of criteria that satisfy the conditions above should be available

These requirements as a set should not be thought of as easy to realise. They represent the theoretical ideals, but in any real life setting they will conflict with each other. Operational specification often requires further decomposition, thus increasing the number of sub-criteria, while completeness may lead to redundancy as the chances of causal interrelatedness and preferential interdependence increase (Winterfeldt and Edwards, 1986, p. 43).

Normally, defining the criteria in an MCA is to be settled between analyst and decision-maker. The process involves both empirical matters about the causal relation between

¹⁰ There are also multiple judgement criteria involved. As we will see later, when these criteria get (common) weights attached to them, they combine into one judgement criterion.

different aspects, and preference matters about the values the decision-maker holds concerning the different criteria, and their interdependence.

The process of 'value elicitation' from decision-makers (or stakeholders) requires the analyst to ask many 'why?' questions. So analyst and decision-maker are 'delving deep' into values (Keeney, 1992, p. 150) and the analyst is necessarily biased in this structuring process.¹¹ A major purpose of this process is to sort out inconsistencies in values. If decision-making remains informal, or if analyst and stakeholder do not dig deep enough, inconsistencies will remain. Therefore, a significant intrinsic quality of MCA is its requirement that actors involved in the process of MCA clarify relevant values and their relations.

Understanding the essentials of trying to meet the theoretical ideals of a criteria-set is crucial for a full appreciation of MCA's performance in the sustainability context. A simple example from Keeney and Raiffa (1976, p. 53-55) illustrates these essentials nicely.¹² The example runs as follows: A doctor has to perform a critical operation and may have the overall objective to 'do the best for the patient'; the highest order criterion in the value tree. In the process of deliberation with the MCA analyst, this highest order criterion may at the beginning be subdivided into 'minimise costs' and 'avoid death'. The higher order criterion has been decomposed, the sub-criteria are clearly operational and non-redundant, and may be minimal. But the question arises as to whether the criteria set is complete. Doctor and analyst may think it is not and decide that the amount of pain and suffering to the patient is an important criterion too, and so a third criterion should be added: 'minimise pain'.

However, this important criterion is not very operational, as pain may not be easily measurable or meaningfully measured to the patient. After some consideration, doctor and analyst may find that using a proxy criterion 'number of days which the patient must stay in bed' might be the most useful third criterion. It seems to be related to the amount of pain suffered because of the medical surgery and it is clearly more operational. Next, as a further step in the process, in this specific situation doctor and analyst may find that the cost of the operation is dominated by the single cost-component of the number of days the patient stays in bed. The criterion 'minimise costs' has now become redundant and may be eliminated from the structure. After this process, doctor and analyst may find that two criteria 'the number of days in bed' and 'the probability of death' is a (decomposed) set that is complete and operational, has no redundancies, and seems to have a minimal size.

In many evaluation contexts it is logical to involve stakeholders in the process of eliciting and organising criteria. In the example above the patient might be involved as a stakeholder, as may be the patient's family, as well as hospital management. Involving stakeholders directly affects the characteristics of the hierarchical criteria structure, as the demand for more objective operational measurement increases. To see the different layers in a hierarchical tree as means/end relations is often useful: the lower order criteria are means towards the higher order criterion or objective. Generally, in any hierarchical structure, one moves from, say, one broad non-operational criterion (e.g., 'the good life') to a few less broad criteria (e.g., 'convenience', 'safety,' etc.) down to a larger number of narrower operational criteria. (e.g., 'minimise door-to-door travel time,' etc.). Whereas a single decision-maker may easily restrict him or herself to subjective assessments of a few broad criteria, in a setting of accountability to stakeholders, the decision-maker should continue as far as possible with 'jointly held

¹¹ Bana E Costa et al. (1997, p. 34) remark: "...structuring is a mixture of art and science... with analysts as 'scientific artists'."

¹² See Bana E Costa, 2001 for a more recent practical case.

objective conceptions' (Keeney and Raiffa, 1976, p. 46).

The MCA analyst first has to ask different stakeholders/decision-makers about "what values *from their perspective* [the project] may affect" (Edwards and Newman, 1982, p. 36). These values are structured into criteria used to assess the performance of the different project alternatives.

When eliciting the criteria in interaction with stakeholders, one may begin with statements on the project objectives. The statements may already differ for different stakeholders, but values should not be confined merely to objectives, as other values may be (positively and negatively) affected by the project.

MCA tries to develop a single, internally consistent value structure from the responses of stakeholders.¹³ Such a structure, as mentioned above, usually involves a value 'tree' with a limited number of higher order values broken down into branches and twigs reflecting lower order values. In the setting with multiple stakeholders "nothing guarantees that the values elicited [from them] will be similar enough so that they can be arrayed in the same value tree" (Edwards and Newman, 1983, p. 41). Trying to accomplish only that is one of the most demanding and very useful tasks of the MCA analyst. If no common structure exists, that is, if several or all stakeholders have a different criteria structure, there is no obvious way to relate different judgements.¹⁴ Therefore, for a powerful use of MCA, after having asked various stakeholders about their values, the analyst has to *persuade* all stakeholders to accept a common value structure.¹⁵

In order to achieve agreement among stakeholders about the common criteria structure, the analyst can try at least three important things: First, the analyst should separate value attributes (criteria) from aspects that stakeholders are merely curious about. Value attributes should make a difference when judging the project alternatives. If the attribute varies, the judgement should change, if not then the attribute is not useful. Interestingly, assessing in advance the possible measurement of the value attribute may help here too. Trying to operationalise the value attribute may help sharpen its actual meaning, and thus clarify whether it can or should be included.¹⁶ So the judgement concept and the measurement are explicitly intertwined in MCA: concepts that cannot be measured are not used.

Second, the analyst should be keen on standardising terminology. Values can be expressed in myriad ways. If many slightly different definitions of essentially the same value lead to many slightly different criteria in an MCA value tree, a serious form of double counting – redundancy – is introduced. To assist in clearing out these differences, one can look at specific measurement possibilities to find out whether noticeably different values are measured.¹⁷

Third, stakeholders may be persuaded to accept criteria, which they feel are not important, into a common structure by allowing, later on, in stage 4, different stakeholders to use different weights. They may even give criteria a zero weight, reflecting no importance at all. Allowing

¹³ Moss et al., 1996 (p. 212) speak of "constructing a consensus."

¹⁴ Stockman et al. (2000, p. 138) note that during the decision-making process positions of the stakeholders tend to freeze, which is detrimental to ideals of any evaluation.

¹⁵ Edwards and Newman suggest a last resort option (if all else fails) for the analyst. He or she may simply list values of different stakeholders as different branches in a value tree. It seems that MCA then has become a descriptive tool instead of a judgement tool (compare section 4.1.2).

¹⁶ Hammond et al., 1999, p. 80: when "... discussing how they would measure [...] would clarify their concerns."

¹⁷ In cases of strong stakeholder conflict even measurement issues can be very heavily debated (e.g., accident rates of nuclear power plants). In these cases Edwards and Newman suggest including different measurements as complementary: simply use them all and allow different weights to be attached..

different weights of different stakeholders is of course a serious departure from a common judgement criterion. Nevertheless, conflicts about weights seem to be easier to interpret, discuss (and perhaps resolve) than conflicts about structure. The reason for this is because differences in weights “lend themselves so easily to compromises.”¹⁸

4.2.3.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.3.2a Triple E impacts can easily and equally be incorporated into MCA

MCA defines a judgement criterion involving different aspects as part of the analysis. Contrary to CBA, MCA can treat all Triple E impact areas comparably from the outset. The first noteworthy aspect is that MCA can evaluate Triple E impacts by having separate criteria for the three dimensions.¹⁹ This capacity is probably the major advantage of MCA. It can handle a set of different criteria, even when they are qualitative, or ‘intangible’, as may be the case with certain environmental projects. The different Triple E impacts of a project can be analysed within their own dimensions (Nijkamp and Van Delft, 1977; Zimmerman and Gutsche, 1991, p. 20). Provided of course that the analyst can decide on a set of criteria that meet the Keeney and Raiffa requirements: to be complete, operational, decomposable, non-redundant and minimal, all three dimensions seem to be on an ‘equal footing’; they are treated as conceptually comparable from the outset.²⁰

4.2.3.2b Values are structured in the process of MCA

Building a common value structure in MCA in which a set of criteria is acceptable to both decision-maker and analyst, is crucial.²¹ If more stakeholders are involved, then establishing a commonly accepted value tree is perhaps the most important achievement. The analyst is truly an analyst when eliciting and organising values from stakeholders: defining, modifying, eliminating and sharpening, then organising the criteria to gain stakeholder approval for the next step. But before it can be judged how workable the ‘digging into values’ is in the context of accounting for sustainability, fundamental problems and possibilities have to be addressed, such as the possibility of organising stakeholder involvement in the context (section 4.3) and the problems associated with large value trees (section 4.4).

4.2.3.2c MCA depends on measurement options for clarifying and structuring judgement criteria

In an MCA several methods can be used to persuade stakeholders to adopt a common judgement structure of criteria. One important way of clarifying judgement values and real or unreal differences herein proves to be looking at the possible measurement or operationalisation of the values. This measurement should aim at ‘jointly held objective conceptions’. Therefore, in MCA there is a close link between judgement concept and measurement.

We can observe that CBA and MCA are equal in this respect, as CBA is also limited to

¹⁸ See Espinasse et al., 1997, for weights in the context of negotiation among different actors.

¹⁹ See Munda, 1993.

²⁰ Using the Triple E dimensions in a sustainability context, one might argue that it is the characteristic of being complete that will often be problematic, as many more aspects may seem relevant.

²¹ This practice contrasts sharply with CBA’s (ideal) ambition to evaluate broad welfare impacts independent of stakeholders. Chapter 3 has shown that this ambition is not realistic in the sustainability context.

measurable impacts.²² However, they differ in the ‘distance’ to be bridged between overall judgement concept(s) and measurement, and more specifically, in the way to bridge the distance.

4.2.3.2d In a complex decision-making context the issues of interdependence of preferences and double counting by unclear causality are important

Being complete is obviously hardly possible in most real life cases, and striving for completeness will hamper the usefulness for decision-making (Keeney and Raiffa, 1976, p. 38). The problems of trying to be complete are even more problematic in the sustainability context; such an attempt in a complex sustainability setting will easily lead to far too many criteria. In turn, a large number of criteria will lead to many interdependent elements, either in terms of factual interdependence or in terms of preferential interdependence. Double counting of impacts then becomes a major problem. It seems therefore that the demand for completeness is very difficult in our context, and thus requires further analysis. The problems associated with a large number of criteria and the difficulties of establishing trade-offs in such a setting will be dealt with in greater depth in sections 4.4 and 4.5.

4.2.4 Stage four: Assess the relative importance of criteria

4.2.4.1 General description

The next stage in the MCA is assigning weights to the different criteria. Attributes identified in the value tree are usually not considered to be of equal importance. To assign weights to attributes is essential to MCA, and the reader may observe that it is judgement in optima forma. Weights will differ from stakeholder to stakeholder,²³ so once again the role of the stakeholders is very important. Stakeholders, or their representatives, have to be consulted again to give weights on the, hopefully, common value tree made in the previous stage.²⁴ In a multi-stakeholder setting the analyst cannot achieve the two tasks simultaneously; first agreement on a common value tree has to be reached, and then weights are given.

MCA practice shows that giving weights is sensitive to how the criteria are hierarchically structured (e.g., Bana E Costa et al., 1997, p. 34). This is known in MCA literature as the splitting bias. Bell et al. (2001) illustrate this point in a controlled experiment with a total of six criteria. They report that the three criteria forming the lowest level twigs in the hierarchical structuring had half the weight they had in the non-hierarchical structuring. This result underlines once again the importance of the structuring phase of building criteria, as well as the need for accountability in building that structure to prevent arbitrary outcomes and manipulation.

There are several procedures for giving weights²⁵ of which the most important is ratio

²² Being measurable or not may not always be absolute or easy to determine: it may be a process of discovery to determine the extent to which certain impacts can be measured in an acceptable way (Wiedeman, 1978).

²³ As discussed at length in chapter 3, even one stakeholder may have different and conflicting preferences depending on the context in which they are observed or requested. This issue will be further discussed later.

²⁴ See Oskam (1989) for an interesting CBA-like revealed preference approach of ‘inferring weights’ from a sufficiently large number of past decisions.

²⁵ The ‘default’ option is giving equal or unit weights. If there are four attributes or criteria they all get a 0.25 weight. This approach may be applied: 1) because for some reason it may be too difficult to elicit the weights from the stakeholders 2) because the weights from different stakeholders may differ strongly and this may be one

weighing. In its general form it is preferred by many in the MCA field (Stewart, 1992), and for good reason. The procedure may be demanding to stakeholders and decision-makers, but it does give direct and precise weights. To fully grasp the most significant aspects of giving weights in its preferred form, a specific form of ratio weighing – as explained in Edwards and Newman (1983) – is discussed here briefly.

Ratio weighing starts with the ranking of attributes from most important to least important. The lowest valued attribute may then be given a weight of 10 points. Attributes deemed utterly unimportant (zero weight) should not be used. Next, the stakeholder judges how the relative value of the other attributes sizes up: e.g., 4 times as much as the lowest, or 2.5, or 17 times as much, depending on the stakeholder's feelings. Edwards and Newman stress that stakeholders be shown how this ratio weighing works and its affects on the outcomes. An understanding of what they are doing will help stakeholders assign clear weights. As a consistency check this procedure is repeated by taking the next lowest values attribute as a basis. To this basis again a value of 10 is assigned and the former lowest attribute is disregarded. Repetitive rounds are performed until it has been done for only the last two attributes, which are the highest in rank. The analyst then checks to see if the weights given are consistent. The ratio between attributes should be the same, regardless of the base attribute. If not, the judgements should be reconsidered until they are consistent. To save time, Edwards and Newman suggest that one not follow the complete round of 'shifting base attributes', but to conduct this at least once. After the inconsistencies are cleared the weights are finally normalised, that is, they are rescaled to values between 0 and 1.

Another popular technique to derive weights which resembles the ratio weighing procedure of Edwards and Newman is Saaty's Analytical Hierarchy Process (AHP) (Saaty, 1980). The main apparent difference between the two is that AHP does not take the lowest valued attribute as a base, and therefore has less of an absolute anchor.²⁶ AHP repeatedly takes two attributes and asks the stakeholder to express whether the one is deemed as: much more important than the other, somewhat more important, of equal importance, of lower importance, of much lower importance, and so on; usually nine options are given. When all pairwise comparisons have been made, relative weights are calculated. These weights can be checked for consistency. The questions posed for eliciting values from stakeholders in AHP are very accessible, but this accessibility may be somewhat deceptive.²⁷ The qualitative answers are interpreted mechanically as specific quantitative ratios, and the checks for consistency are complicated and difficult to understand (Stewart, 1992, p. 574).²⁸

Two significant aspects have become clear from this discussion. First, the weight-giving in its standard form is not a 'one-shot' try. Decision-makers and stakeholders have to repetitively judge the relative value of different criteria against each other. Second, value judgements are

way to reach agreement and 3) because in a sensitivity analysis it may be seen how much the specific weights influence outcomes, by confronting it with unitary weights.

A second option is to derive weights from ranks. Ranking attributes is a relatively easy procedure for stakeholders to express the importance of attributes: they list the most important attribute first, the next in line second, etc. The analyst can then calculate a weight from this order using the rank numbers by either normalising inverse rank numbers or normalising the reciprocals of the normal rank numbers. Normalisation means dividing the raw weights by their total sum, and assures we have weights between 0 and 1 adding up to 1.

²⁶ Although this can be remedied (Stewart, 1992, p. 574).

²⁷ See Lahdelma et al., 2003, for a (theoretical) comparison of two pairwise comparison methods.

²⁸ Another important problem, to be discussed later, is that the context of giving weights is isolated from the specific ranges of options available (Stewart, 1992, p. 574). The value of pairwise comparisons is well-established however; with these separate comparisons it may be easier to learn about one's preferences and the dominance structure of the alternatives (see for instance Hammond et al., 1999, p. 81).

made consistent along the way. Preferences are usually *not* consistent and the process of involvement in MCA helps those involved to structure their thinking about the problem situation (Keeney, 1992), and thus make preferences more consistent. Many problem situations are ‘messes’ and MCA in these instances does not necessarily ‘solve problems’, but often provides a formal way of ‘mess reduction’ (Stewart, 1992, p. 569; French, 1998).²⁹

4.2.4.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.4.2a Assigning weights may be difficult, especially with many stakeholders

The second important part in MCA where stakeholders give input for the judgement criterion is by giving weights to the criteria, or the attributes. This process can require much time and effort from stakeholders, especially when many criteria are involved.

The process of assigning weights often requires clearing out inconsistencies in preferences or values, which means that MCA accommodates for preferences that change during – and as a result of – the evaluation. This sharpening of preferences is generally seen in MCA as a major benefit of the formal analysis: it reduces the ‘mess’ in the problem situation. In the sustainability context this procedure may be difficult primarily due to the many stakeholders who are (potentially) involved. The difficulties associated with having many stakeholders will be elaborated in sections 4.3 to 4.5.

Another aspect that may be specific to the sustainability context is a lack of information about impacts and thus (near) absence of preferences about these impacts. To counteract this possibility, decision-makers and stakeholders are often informed first in order to arouse emotion and activate thinking about values.³⁰ This situation resembles the case of contingent valuation studies in chapter 3.

Finally, at the border of stages four and five, we notice that there is little standardisation of measurement in MCA: every project evaluation will have another set of stakeholders, and in every setting new criteria and weights may emerge. This procedure contrasts sharply with CBA theory, which more strongly emphasises the idea that the welfare impact of projects can be determined objectively, by looking at data about actual behaviour. However, as seen above, in the practice of evaluation in the sustainability context CBA has comparable difficulties in establishing correct weights (and impact scores) as MCA.

4.2.5 Stage five: Estimate scores of alternatives on low level criteria

4.2.5.1 General description

The next stage in MCA is to establish scores on the low-level dimensions; that is on the lowest level twigs of the criteria tree. To understand the occasionally subtle difficulties here, one may simply think of a one-level criteria tree. There is no fundamental difference with a hierarchical value tree, while it clarifies that this stage is about comparable measurement on several criteria.

In MCA it is crucial that scores be comparable. It is self-evident that different criteria are

²⁹ The phrasing ‘messes’, according to Stewart, is from Ackoff, R.L., 1979: The future of operations research is past. J. Opl. Res. Soc. 30, 93-104.

³⁰ This may take the form of text, slideshows (Le Maitre et al., 1997) or videos (Wenstøp and Seip, 2001).

initially measured differently. One may be in square meters, another in value judgements from ‘excellent’ or to ‘very poor,’ and still another in Euros. Because of these differences all scores have to be transformed. A common procedure is to transform all scores to a score between 0 and 100, where 100 is the most positively judged and 0 is the least valued. One should observe that these 0-100 scores are judgement scores, so all scores are to be transformed to a judgement scale. The transformation or scaling facilitates aggregation but it naturally involves some level of serious information loss.

Depending on the specific criterion a suitable transformation procedure or curve can be chosen; this curve is also called a value function (Bana E Costa, 2001; Stewart, 1992). Simple linear transformation or any other transformation curve can be used, although one should realise that understanding may suffer in the case of very complicated functions. Although there can be logical and detailed arguments in favour of complex transformation curves, notice that if, in practice, the outcomes depend on a finely-tuned transformation curve, then the alternatives may actually be too close together to warrant strong conclusions about ordering (Edwards and Newman, 1983, p. 70).

A very important aspect of the scaling is that the maximum and minimum scores should be *plausible* maxima and minima, that is, the minima and maxima are actually attainable in the specific evaluation context (Bana E Costa, 2001) as there is a close relation between weights and measurement. Consider two criteria, each transformed to a 0-100 scale. If in one criterion the scores of all the alternatives range from 45 to 55 because the (transformation) maxima are chosen very widely, while on the other criteria the 0 and 100 scores are regularly reached, we can easily see that the too widely dimensioned criterion (with 45-55 scores) will be given too low a weight.³¹ The scores on this criterion will never matter much, unless an extremely high weight is given to the criterion from the outset. Therefore, it is very important to define maxima that, according to ex ante judgement, are plausible enough to be reached.³²

It was seen above that weights and the plausible measurement scale are closely interrelated. To disregard the relation between weights and a plausible measurement scale is what Keeney calls “the most common critical mistake” in giving weights in MCA. “The importance of an objective must depend on how much achievement of that objective we are talking about” (Keeney, 1992, p. 147). For instance, saying that air pollution is three times as important as abatement costs, is ‘completely useless’ for building a value structure (Keeney, 1992, p. 147-

³¹ We observe that the alternatives taken into consideration also have an important role in the outcomes of the scaling. If, for instance, three different variants of a new motorway between two cities are considered and to these alternatives a fourth one is added which invests in a railway connection, the scaling may be completely different, and the differences between the motorway variants, which might first have looked impressive, may be far less impressive after the railway variant is introduced. Although this change should be deemed real as the decision space is truly altered with the introduction of a new alternative, it shows the sensitivity to the possibly strategic introduction of alternatives. This sensitivity may diminish the actual weight of a criterion by changing plausible maxima and minima.

³² Consider the swing weights in Keeney et al., 1990. Two more scaling issues are noteworthy as they reveal the subtleties involved in correct scaling. Some of the scores to be transformed may already be judgement scores; for instance, scores ranging from ‘excellent’ to ‘very poor’. In this case the analyst should consider whether such scores are actually plausible. In practice stakeholders may be reluctant to consider the extremes. In such a case it may be wise to let ‘good’ (instead of excellent) be a score of 100 and let ‘poor’ (instead of very poor) be the lowest (0) score. Another issue is the transformation of a percentage variable such as the share of people in absolute poverty in a certain region. One might be tempted to directly use the variable dimension itself. We should by now not be surprised that this is not the correct way to proceed. The analyst should make a judgement about plausible percentage scores (say, between 4% and 40%). Then 4% should be ‘0’ and 40% should be ‘100’.

148) as this might warrant spending enormous amounts on nearly negligible diminishing of pollution.

The interrelation of weights and measurement sheds light on the notion that one cannot assign weights in the absence of context; in other words weights cannot be assigned until there is some idea of the relevant size of the changes (Stewart, 1992, p. 574; compare Bulte and Van Kooten [1999] who make the same point about CBA related Contingent Valuation studies).³³ Interdependency is a somewhat complicating factor in MCA, especially when (many) stakeholders are involved.

Eliciting a transformation curve (in combination with weights) from stakeholders is generally a demanding task for both analyst and stakeholders. Edwards and Newman are rather sceptical about mail and return procedures for this element: personal interaction is much preferred (Edwards and Newman, 1983, p. 71); this seems to be in line with most MCA literature, which stresses the interactive process of value elicitation (Bana E Costa, 2001; Keeney, 1992). One important reason is that many value elicitation studies show inconsistencies in the first round; these may be solved in an interactive process. Since the analyst cannot closely interact with many stakeholders, the only practical way is to work with a limited number of people.

4.2.5.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.5.2a Achieving comparable measurement implies information loss and affects judgement

MCA requires the different scores on the criteria to be transformed to a common judgement range. This scaling is necessary to achieve a common measurement and judgement scale but it has consequences: it leads to information loss, since measurement is no longer in natural dimensions. Stakeholders' understanding may suffer from this fact.³⁴ Second, the transformation of different measurements to a common scale involves several subtle impacts on the judgement outcomes, as a different way of scaling is essentially the same as giving a different weight to a criterion. This procedure may be seen as contrasting sharply with CBA's practice, at least in as much as they relate to market related impacts. Compared to the citizen-preferences related impacts that CBA tries to capture with (for instance) CVM techniques, the transformation process of MCA seems to be much more explicit, open, and well-structured.

4.2.5.2b Interdependency between knowing the relevant range of impacts and giving weights is not ideal for accountability

The complicating interdependency in MCA of the need to have some idea of the relevant size of the impact prior to the weights being assigned seems especially vexing in the sustainability context; it basically amounts to having the impact measurement already reasonably clear before finalising the judgement criterion. However, from the accountability standpoint, this order of treatment is not ideal, and the situation may worsen if there is a high level of uncertainty about the size of the impacts, and even about the causality of impacts. There may be recourse to experts specifying the possible range of the impacts in some cases (Bana E Costa, 2001), but in the context of accounting for sustainability it can be difficult to find experts who are knowledgeable and acceptable to all stakeholders.

³³ In CBA theory the diminishing marginal utility ranks high: lower value (weights) to extra units. Only in a setting in which small changes are relevant a constant weight is thought to be defensible.

³⁴ The incentive to achieve consensus about sophisticated measurement in natural dimensions may even diminish since *any* underlying measurement can be transformed.

4.2.5.2c The preferred personal interactivity of stakeholders with MCA analyst is a complicating factor

Personal interaction between analyst and stakeholders is much preferred for value elicitation, and necessitates the utmost quality of, and trust in, the process of representing a potentially large number of stakeholders, since not too many people can be involved in personal interaction³⁵. Those not actively involved will not have experienced the process of ‘structuring the mess’.

4.2.5.2d The ‘open’ treatment of impacts spread over time within MCA contrasts sharply with CBA’s discounting practice

The issue of evaluating impacts spread over time is worthy of comment. Treatment of these impacts in MCA is open to deliberation between analyst and stakeholders about weights and scaling (and measurement). If one thinks of the practice in CBA of mechanically discounting monetized impacts over, for example, a 30-year period against a fixed discount rate without any stakeholder consultation or further deliberation, the contrast with MCA is very large.

4.2.6 Stage six: Aggregate scores with relative weights

4.2.6.1 General description

Stage 6 is the stage where all is put together, and final results are calculated. The project alternatives and their scores on the lowest level twigs of the judgement criteria between 0-100 have been determined. Weights between 0 and 1 have been given for the highest level criterion and for every branch and sub-branch. These weights indicate the relative importance of the criteria. Although several aggregation rules exist, the most practical and useful is multiplication ‘through the tree’ and summing all scores to obtain a final ‘utility’ score for the project alternatives. The higher the score, the better the project alternative.

Although the final outcome is now easy to calculate, Edwards and Newman emphasise the use and value of sub-aggregation. Apart from the overall score, the scores per criterion may be analysed the same as the outcomes for different stakeholders if they have given different weights. The utility of looking at these sub-aggregated outcomes depends on the function of the evaluation and on the number of alternatives and criteria. With many criteria and many alternatives, sub-aggregation may become less useful. In a situation in which there are difficulties in establishing higher order trade-offs (compare stage 7 below), further attention to sub-aggregate scores may prove to be very valuable.

4.2.6.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.6.2a Difficulty of establishing trade-off may lead to second-best MCA with only partial judgement

The sustainability context is typically one where there may be difficulty in establishing higher order trade-offs. In such a case, attention to sub-aggregate scores may be useful, which essentially means moving from full judgement to partial judgement. We will address this more fully in section 4.5.

³⁵ The number of people involved depends on evaluation ‘size’; i.e., man-power and budget. The number of involved stakeholders will not often surpass, say, 30 people, while 5-15 would be more common.

4.2.7 Stage 7: Perform sensitivity analyses

4.2.7.1 General description

In general, it is important to be aware that a sensitivity analysis is of little value if the weights or impact measurements are varied randomly. Although computational power is increasing continuously, it seems fruitless to produce graph after graph, or table after table, to show sensitivity to this or that variable or parameter. In any MCA there are many numbers that can be varied, and varied simultaneously. A somewhat random process may easily be confusing and frustrating (Edwards and Newman, 1983, p. 93). Any sensitivity analysis should be guided by a careful inspection of the outcomes.

The most important sensitivity analysis is probably the impact of weights. Since the weights are the essence of value judgements and are purely subjective, people may disagree strongly about them. The impact measurements – before scaling – are usually more objective, or more dependent on expert judgement. Because the higher-level weights naturally have the highest impact, varying these systematically is the logical first step, provided of course that there are arguments for variation – such as a lack of consensus among stakeholders or substantial uncertainty among experts.

Other aspects that may be considered are the specification of parameter ranges that leave the final order of alternatives unchanged. We may also ask what values different parameters must have in order to achieve a given ranking of alternatives (Pomerol and Barba-Romero, 2000, p. 265). Another possibility is to conduct a conflict analysis (Bana E Costa, 2001): to analyse which stakeholder groups win and lose through various alternatives.

Edwards and Newman raise an interesting and important point at this stage. They recommend keeping costs separate from the rest of the MCA criteria in earlier stages of the MCA. Their justification is that they find that valuing costs versus other criteria is particularly problematical and is likely to be more controversial than other relative valuations. In the sensitivity analysis they advise the calculation of costs per utility point, for non-dominated project alternatives (Edwards and Newman, 1983, p. 92). This procedure may provide important information on which project alternative should be preferred, without costs being fully integrated within the MCA.

4.2.7.2 Assessment of pros and cons in the context of accounting for sustainability

4.2.7.2a *Additional analysis may give some extra judgement possibility in case of trade-off difficulties*

Although Edwards and Newman only separate costs from all other criteria, the point they make can be generalised. The generalisation refers to situations when assigning weights between criteria is particularly difficult or controversial, then branches of the MCA tree can be kept separate. Although ‘full judgement’ can then no longer be passed, the analyst still has several options for further interpretation and, what may be called, semi-judgement. For instance, the analyst may use the calculation of ratio scores to compare alternatives, or may look for reference ratios from other projects.³⁶ The suggestion to keep costs separate from the other criteria seems to give room for the combination of CBA outcomes with MCA results. However, one should be aware that CBA outcomes involve more criteria than just costs; they

³⁶ The calculation of Benefit/Cost ratios within CBA is somewhat comparable.

combine various (sub)criteria to an overall judgement measurement on efficiency.

4.2.7.2b A first MCA can be followed by another and yet another, including newly emergent alternatives

Another relatively prominent aspect in the MCA literature from the sensitivity analysis stage is that MCA is a (first) step in a decision-making process. The first result may be followed by another MCA; the analysis can be extended to include more alternatives that are generated through the MCA process (compare Foxon, et al., 2002). The MCA can be reiterated around the best (non-dominating) alternatives, with additional attributes. MCA literature acknowledges that even informal decision-making may be the best way to proceed.³⁷

4.2.8 At this stage: what is clear about MCA's performance and what requires more attention?

We have seen in this chapter that the sustainability context, which implies looking at Triple E consequences at global scales and lower, over long periods of time, provides possibilities as well as problems for MCA. Table 4.1 gives an overview of our conclusions thus far. It distinguishes between conclusions that have by now become clear, and those that require further discussion. For MCA evaluation three main areas were identified that warrant in-depth discussion. In the remainder of this chapter these areas are addressed:

- The organisation of stakeholder involvement (section 4.3)
- Interdependence and double counting (section 4.4)
- The judgement orientation of MCA (section 4.5).

These issues may be seen as treating three different aspects of the same central concern: to which extent is the essential MCA process of assigning weights meaningful or useful in a complex multi-stakeholder setting?

³⁷ Stockman points out that convincing information is generally more easily accepted in earlier stages of decision-making (Stockman, 2000). This experience might be a reason in itself for stressing the option of reiterating MCAs. Strategic behaviour however may increase (compare Wenstøp and Seip [2001]) once resulting outcomes become clearer.

Table 4.1 First conclusions about MCA's performance in the sustainability context

	Clear/unclear by now		Further examined in:
	Clear	Unclear	
Stage one: Identify objects and functions of the evaluation			
-MCA does not stress the need for a 'nul' alternative	X		
-MCA reflects on the functional context and tailors the evaluation	X		
Stage two: Identify stakeholders			
-Involvement of stakeholders makes the political and communication skills of the analyst a critical factor for successful evaluation	X		
-Involvement of all relevant stakeholders may be nearly impossible		X	4.3
Stage three: Elicit and organise criteria			
-Triple E impacts can be easily incorporated in MCA on an equal basis	X		
-Values are structured in the process of MCA, but their workability in the sustainability context is unclear		X	4.3 & 4.4
-MCA depends on measurement (options) for clarifying and structuring judgement criteria	X		
-In a complex decision-making context the issues of interdependence of preferences and double counting by unclear causality are important		X	4.4 & 4.5
Stage four: Assess the relative importance of criteria			
-Giving weights may be difficult especially with many stakeholders		X	4.3, 4.4 & 4.5
Stage five: Estimate scores of alternatives on low level criteria dimensions			
-Achieving comparable measurement implies information loss and affects judgement	X		
-Interdependency between knowing the relevant range of impacts and giving values is not ideal for accountability	X		
-The preferred personal interaction between stakeholders and MCA analyst is a complicating factor	X		
-The 'open' treatment of impacts spread over time within MCA contrasts sharply with CBA discounting practice	X		
Stage six: Aggregate scores with relative weights			
-Difficulty of establishing trade-offs may lead to a (second-best) MCA with only partial judgement and disaggregated results		X	4.5
Stage seven: Perform sensitivity analyses			
-Additional analysis may give some extra judgement possibility in case of trade-off difficulty	X		
-A first MCA can be followed by another and yet another, including newly emergent alternatives	X		

4.3 MCA and organising stakeholder involvement

4.3.1 Introduction

Although important conclusions have already been drawn about the involvement of stakeholders, this section will look more closely at this subject. The participation of stakeholders has become increasingly standardised in the evaluation literature in recent years (Gregory, 2000; p. 180). It mentions several arguments in favour of participation: it can empower people; the expertise of various stakeholders may be more fruitfully involved; and commitment to implementation may increase (Gregory, 2000; Patton, 1997; Renn 1993; Guba and Lincoln, 1989). However, in practice stakeholder involvement is far from easy, and especially in the sustainability context it may be extra difficult.³⁸

In the sustainability context there are two specific problems that hinder the fruitful involvement of stakeholders:

- First, it will be impossible to involve all stakeholders because there are too many individuals and they will be physically or geographically distant from each other.
- Second, stakeholders will often have opposing interests with little room for compromise, which leads to the fundamental problem of ambiguity of the perspective taken by stakeholders.

4.3.2 Organising many distant stakeholders

To tackle the first problem the analyst may try to *represent* stakeholders in order to arrive at a workable setting. In the general evaluation literature this would be called a problem of procedural justice. If an evaluation aims at procedural justice emphasis is given more to the correct organisation of the evaluation than to the content. Probably the ideal situation for MCA in a sustainability context would be to work with elected representatives, that is, a situation in which members warrant involvement of their representatives. Although this procedure would ensure that everyone has some formal share in the decision-making process, it is not very practical, as more often than not these representatives will not exist. However, if one cannot work with formal representation, how can one be sure that the non-involved stakeholders will accept the critical choices made by their ‘representatives’?

The global context of sustainability has shown that in several day-to-day decisions many distant stakeholders are directly affected, and -- in line with evaluation theory -- these should be actively involved. However, there does not seem to be a practical and convincing procedure,³⁹ for the involvement of this ‘worldwide stakeholder group’. To illustrate, one may consider a concrete decision: on the tea brands available in restaurants of Dutch universities. Such a choice will directly influence the economic future of some low-income tea-growers in Kenya and India.

³⁸ See Gregory (2000) for an in-depth discussion of sometimes ‘rather naive assumptions’ in participatory methodologies. Gregory argues for a simple analysis of power aspects among potential participating stakeholders, which may give more insight into the specific barriers to participation (at different levels, from personal to global) and may suggest ways to circumvent them.

³⁹ In evaluating the relation among 48 world problems, Hwang and Lin (1987) used stakeholders who were in a position in society that enabled them 1) to synthesise a large amount of information on world issues and 2) to exert influence on actions related to those issues (32 participants, 5 from less developed countries).

Theoretically, it might be ideal to involve these stakeholders for information about the real life impacts, for structuring the value tree and possibly for finding new alternatives. However, practically speaking:

- a) a much larger budget than is regularly available is required, or extra budget that is deemed unnecessary to many of those who are financing an evaluation
- b) the stakeholders might not have the time to be involved
- c) one might not oversee at the beginning of the evaluation who should be involved (the analyst may have overseen Chinese tea-growers)
- d) there will be large differences in cultural background, knowledge, language, etc. and these differences are problematical for fruitful involvement.

Solving all these problems may be beyond the skills and knowledge of most evaluators.

Furthermore, one should keep in mind that decision-making with a small group of representative stakeholders is often a subtle social process, involving non-evaluation (or non-task) activities that influence the evaluation process (Poole and Hirokawa, 1996, p. 11). There is abundant literature about the problems involved in small group decision-making: power differences, knowledge differences and crucial aspects of communication, and finding common ground (Teisman, 1998; Hirokawa and Poole, 1996; Scheper, 1991; Kacprzyk and Fedrizzi, 1990; Hwang and Lin, 1987). In accordance with this literature, section 4.2 has shown that interaction between analyst and stakeholder is crucial and intense, implying that after the evaluation there will be a marked difference between involved stakeholders and non-involved stakeholders. In the accountability context the public at large or various other potential stakeholder groups may prove to be very relevant. Therefore, if an evaluation method focuses too much on the stakeholders involved, without trying to be informative, enlightening and/or convincing to other stakeholders, this procedure has little guarantee for success.

The analyst therefore encounters a predicament. Complex problems seem to require complex decision-making tools. These tools necessitate the close involvement with a limited set of stakeholders to interactively structure the complexity. However, many non-involved stakeholders will assess complex problems and the outcomes of decision-making too. This fact requires that methods be simple and transparent (Gregory, 2000, p. 195). Although this dilemma is not easily resolved, we will argue below that standardised measurement of impacts – together with a less judgement focussed approach to MCA – may help to somewhat escape the dilemma. In section 4.5 and in chapter 5 this problem-escape-route will be expanded upon.

4.3.3 Reconciliation of diametrically opposing interests

Tackling the second problem of opposing interests may be at least as difficult as the first. In the sustainability context stakeholder involvement will often not be particularly fruitful if members are stuck with diametrically opposing interests. Some stakeholders will lose and some will gain. This situation may lead to opposing value trees or weights, and in the sustainability context of extreme poverty, this condition is especially relevant. The procedure of stakeholder involvement presupposes the possibility for compromise and negotiation; but the extreme poverty situation will limit these possibilities and therefore a successful involvement is not assured.

Several MCA authors argue that stakeholders can be asked to adopt a different perspective outside of their own personal interest. Keeney (1992, p. 150), for instance, argues that if stakeholders are involved in the MCA process of structuring criteria and giving weights, the

analyst should first discuss which perspective the stakeholder should take: self interest or public interest?⁴⁰ If the analyst can persuade stakeholders to adopt the non-individual viewpoint, there is obviously greater room for compromise, but the process may end up entangled by more problems than it started with.

Thus far it was implicitly assumed that stakeholders represent a clear judgement perspective, and that the value of stakeholder involvement seemed to be the embedding of various clear-cut, real life perspectives into the evaluation. Isolated and all too distant decision-making could therefore be prevented and accountability could be enhanced. If we now acknowledge that stakeholders can choose to adopt different perspectives this complicates accountability, as account also has to be given for the choice of perspective. Furthermore, the analyst will have to ensure that stakeholders do not switch perspectives along the way; they could for instance answer familiar evaluation questions from an individual perspective and others from a general perspective. O'Connor (2004a, p. 175) emphasised precisely this point as "the permanent possibility of shifts in perspective." Finally, the realism of their non-individual perspective could be questioned – depending of course on the context. Many product developers and marketing experts fully comprehend the gap between stated opinions and actual behaviour.

This gap has already been encountered with the dual consumer-citizen preference structure. CBA had a strong focus on actual behaviour and seemed to have more trouble with incorporating public preferences. MCA now seems to do the opposite and may more easily disregard behaviour-based realism.

4.3.4 Conclusions on organising stakeholder involvement

4.3.4a Practical impossibilities limit representative stakeholder involvement

Although the involvement of stakeholders may theoretically clearly be helpful, there is little guarantee that it will succeed in practice. Finding and productively involving representative stakeholders for MCA in a sustainability context seems to exceed practical possibilities in many cases, and a gap exists between involved and non-involved stakeholders. There is apparently little guarantee that the value tree and the weights given are acceptable to non-involved stakeholders. Given the need for acceptability by non-involved stakeholders, the dilemma is to find simple and transparent methods that assist in complex problem situations.

4.3.4b Opposing interests of stakeholders discourages the search for a viable common perspective; broadness of perspectives can be achieved and can be useful

Furthermore, opposing interests will limit fruitful interaction unless stakeholders adopt a perspective not limited to self-interest, which also uncovers the dependency of an MCA approach on the adoption of a consistent perspective by stakeholders, and simultaneously casts doubt on the feasibility of finding such a perspective. We encountered this problem in chapter 3, where the dual preference structure was discussed. CBA proved to be more on one side of the spectrum, and MCA can now be seen as being more on the other side. Although one notices that both are incomplete, there is no way to know which is the better approach.

From a practical viewpoint our major conclusion thus far is that stakeholder involvement will

⁴⁰ Wenstøp and Seip (2001) involved civil servants because of their unique knowledge of some MCAs: "they were told to participate in the valuation sessions as normal civilians" (p. 59).

not naturally lead to a consensus. The best practical function of stakeholder involvement seems to be the assurance of a wide array of perspectives; that is, the involvement of a non-representative but broadly mixed set of stakeholders may prevent the analyst and decision-maker(s) from too easily adopting their own versions of reality. Moreover, such an involvement may be valuable for finding ‘jointly held objective conceptions’ and measurements; as towards this aim it is not the number of stakeholders but rather their different viewpoints that seem to be decisive.⁴¹

4.4 MCA and interdependence of criteria and double counting

4.4.1 Large value trees

MCA in a sustainability context will easily lead to large value trees. The reason for the emergence of large value trees is twofold. First, there is the complexity of the issues as such; including Triple E aspects, global level and long-term impacts (Azar et al., 1996).⁴² Second, the problem is supposedly caused by serious stakeholder involvement in MCA: practice shows that in such a complex setting, stakeholder involvement processes easily lead to large value trees. Although MCA literature emphasises keeping the number of value attributes low (“eight would be about right and fifteen would already be excessive” (Edwards and Newman, 1983, p. 51), in practice this proves to be rather difficult (Adler and Posner, 1999). The sustainability context is also complex in its interrelations, which are not always well understood. Stakeholders can therefore be expected to bring in many aspects that are not easily ordered and reduced to a small number of attributes. The context of accounting for sustainability seems to be conducive to large value trees (Foxon et al., 2002). A few examples may serve to illustrate this point.

Janssen (2001) presents an overview of Dutch experience of MCA use in Environmental Impact Assessments. These experiences are good examples of appraisal in a sustainability context. In the Netherlands public participation in decision-making is deeply institutionalised; all relevant parties must be heard and there is a strong motivation for a consensus-seeking approach. Janssen clearly shows that the number of criteria is large in all 21 examples considered. The minimum is 14 criteria with no hierarchical groups and the maximum is 100 with a hierarchy of 10 groups. According to Janssen especially evaluations with a high political profile finish with large numbers of criteria.

Edwards and Newman (1983) present an example of an evaluation of school desegregation in Los Angeles that ended with a value tree having 144 different sub-criteria (in seven groups). Although this was one of the largest trees they had ever seen, they do find that a “far too large” value tree is characteristic of most value trees in important public policy areas.

Keeney et al. (1990) in a German evaluation of energy scenarios developed a common value tree for all stakeholders with 106 criteria – in eight groups (see below in BOX 4.1).

These examples illustrate that the process of making a common value tree in the sustainability

⁴¹ Heap (1995) finds that the ‘consensual methodology’ of evaluation is the most useful part of the Fourth generation evaluation of Guba and Lincoln.

⁴² Hanley (2000) discusses among others, the 134 indicators proposed by the UN Commission on Sustainable Development in 1995.

context has a so-called ‘upward bias’. Stakeholders in the sustainability context have different interests and understandings; they are in an international setting and (at the least) will be somewhat interested in long-term impacts. If one then starts by asking stakeholders for value-elements in the value tree:

- one will have a manifold of slightly different definitions and thus criteria; *unless* the analyst can persuade stakeholders to accept common definitions⁴³;
- one will have many things about which stakeholders are curious and which may have some relevance to the value of the projects; all these items or concerns will lead to criteria *unless* the analyst can prove them (in advance) not to be relevant.
- one will have strongly opposed interests, which will lead to different criteria for different stakeholders, and even different ideas about correct measurement. *Unless* the analyst can persuade stakeholders to combine or drop criteria, this will lead to simply adding them to the list (and allowing zero weights later on).
- one will introduce more and more causal chains and interdependence between criteria, as the lists become longer; *unless* the analyst can uncover the links and interdependencies and persuade stakeholders to modify or drop criteria for the otherwise detrimental consequences to the analysis.

Unless the analyst is extremely skilful, and unless stakeholders are very co-operative, the analysis will easily end up with large value trees. If one does have a large value tree, there will be many criteria with very small weights, which strongly reduces the significance of the weights.

4.4.2 *Interdependency of criteria and double counting*

Large value trees are to some extent simply a reflection of the myriad concerns of different stakeholders. So in the context of complex decision problems they are unavoidable (Edwards and Newman, 1983, p. 51).⁴⁴ The reader may recall the five characteristics a set of criteria should have according to Keeney and Raiffa (section 4.2.3). In the complex sustainability context the demands of attempting completeness dominates over the demands of being non-redundant. Completeness means the absence of further criteria, which can be used to judge between the project alternatives, while non-redundancy (or no double counting) means that no aspect of the problem should be accounted for more than once. We can now observe that long lists of criteria meet the demand for completeness as a result, but the demand for non-redundancy suffers most.

Despite the fact that the emergence of large value trees is easily understood, large value trees undoubtedly reflect a loss of perspective. Different criteria scores in an MCA are compiled and aggregated into one index, but this procedure is only warranted if the causal relations are clear. As Yoon and Hwang (1995, p. 31) express it: “When we thoroughly understand the functional relationship within a system, we may formulate an index to represent its effectiveness.” If these functional relationships are not understood the problem situation may

⁴³ Stockman (2000, p. 139) points to the fact that people tend to be myopic in two respects: they give higher weights to short-term costs and benefits than to longer term items; and they tend to see small differences around their own positions more sharply than differences among remote positions.

⁴⁴ These problems seem to be among the type of problem groups that are messy, ill structured, strategic, soft, wicked, ambiguous, fuzzy (Scheper, 1991).

seem to be structured, but with a large value tree it is not structured clearly.⁴⁵ As Janssen (2001) has shown, the value tree becomes simply an addition of requests from all stakeholders, instead of being the result of a systematic analysis of the type explained in the example of paragraph 4.2.3. There will typically be some confusion between means and ends in the criteria, dependencies among criteria, double counting, and inconsistencies in spatial and temporal scales (Janssen, 2001, p. 104).⁴⁶ In other words, too many criteria will result in an unclear judgement concept..⁴⁷

Box 4-1: Illustration of the problems of interdependence and double counting

To make tangible the problem of interdependencies and inconsistencies in scale, the case of ex ante evaluation of new energy policies for Germany as published by Keeney et al. (1990) is illustrative. The case as such can be seen as state-of-the-art and relevant; well-qualified and experienced analysts were present, serious stakeholder involvement was organised and the problem of different energy provision scenarios for Germany is highly relevant in the sustainability context. Keeney et al. developed a common value tree for all stakeholders with 106 criteria (in eight groups). Without presenting a thorough analysis of the complete tree it may be illustrated how judgement becomes complex when the value trees enlarge. Under the group of criteria of 'financial technical and material requirements' the reader may find at the lowest level of sub criteria all the cost items: investment costs, operation costs, R&D and more. At the same time under the criteria group 'national economic impacts' one may find at the lowest level the criterion 'employment' – which is clearly a case of severe interdependence (or double counting) as employment costs – i.e., wages – form a large part of the cost items mentioned earlier. Furthermore, there is a group 'political' criteria. One of the lowest level sub-criteria is support by majorities, while two others under the same group are protection of minorities and protection of the majority from disruptions by minorities. This combination seems somewhat awkward and more like a list of concerns from different stakeholders rather than an analysis of criteria. Finally, stakeholders had concern about the danger of terrorist attack in some energy scenarios, as there is such a lowest level sub-criterion under the group 'International impacts'. Under the heading 'Health and safety,' one may find threats to life, which is closely related to terrorist dangers, while it is also clearly a disrupting minority (see above). Already from this brief illustration we notice that causal relations in such a large tree jump up and down the tree and do not stay within the sub-branches as is preferred. Furthermore, the requirement that all value aspects should be accounted for only once is easily violated. In sum, the clarity of the judgement provided by the MCA suffers severely from a situation of large value trees.

When applying MCA in a sustainability context the problem of interdependence of criteria is crucial because many impacts will be uncertain and several feedbacks in evidence will be only partly known and partly understood;⁴⁸ this remains one of the most important unsettled problems in MCA (Fishburn and Lavalley, 1999).⁴⁹

We can make a distinction here between empirical interdependence and preference

⁴⁵ Nijkamp and Van den Bergh (1997, p. 181) in a discussion of economic analysis of environmental issues find that: "A proper use of multicriteria analysis presupposes inter alia the existence of an adequate quantitative environmental-economic model."

⁴⁶ Brouwer et al., 1999, in a meta-analysis of CV studies notice that because of the interrelatedness of ecosystem functions "double counting is a real problem and attempts to aggregate up to system-level values are fraught with difficulties" (p. 50). See Florax (2001) on some interesting problems of meta-analysis as such. Tell (1978) has suggested factor analysis as a method for reduction of the number of criteria.

⁴⁷ Wenstøp and Seip (2001, p. 60) notice that (in a specific setting of rating power plant impacts) valuations produced arbitrary results because the panels giving the weights could not relate emotionally to the index scales – because they had "too many different practical implications, and therefore became too theoretical."

⁴⁸ Comprehension of the criteria to the decision-maker is very important in the acquiring of valid weights (Wenstøp et al., 1997).

⁴⁹ See Carlsson and Fuller (1997) for an attempt to actively use measures of interdependence.

interdependence⁵⁰. The classical example of preference interdependency is the choice of wine (red or white) with the type of food (meat or fish). To most people there is no independent preference structure here, as most people prefer red wine over white with meat, but white wine over red with fish (Vansnick, 1990, p. 96). If the analyst observes preferential interdependency, then this fact can indicate an – as yet – badly composed criteria structure, since it reveals other (deeper or higher order) preferences which have not yet been specified (compare Olson, 1997, p. 12). Although preference interdependency is best kept absent, it is possible to work with in an MCA, provided that throughout the possible impact range of a single criterion more is preferred over less (Olson, 1997, p. 36). Empirical interdependency is regarded as the more serious problem of the two.

However, in the sustainability context the amount of complexity and lack of familiarity with the future choices seem to be too large to fruitfully apply the distinction between preference and empirical interdependency, as a mix of both is most common. In any real decision-making problem in a sustainability context it is common to have criteria with several interdependencies, especially if the criteria are a mix from both sides of the dual preference structure.

The outcome of this situation will in general complicate the productive use of MCA; it may even lead to doubts among (involved) stakeholders about the utility of MCA for making judgements. The complex situation in MCA is structured into a neatly organised formal decision space, and a tension is introduced between modelling and reality⁵¹. This tension is probably most explicit in the weight-giving phase of MCA. Several people resist the decomposition of value judgements, for instance, making pair-wise comparisons to give weights, because they are uneasy about the presumed independence.⁵² In a recently performed test of several MCA techniques in a decision-making setting with four to six criteria, most participants preferred holistic direct assessment of alternatives over (any kind of) MCA structuring requiring giving weights (Bell et al., 2001). The apparently simple judgements of giving weights in an MCA evaluation may – in a complex setting – be experienced as more burdensome than the original decision. Stakeholders may ultimately prefer their own unaided judgement (Ozernoy, 1997).

The problem of interdependence of criteria resulting from (far) too large value trees can be tackled by reconsidering the demands for a clear criteria structure. The requirement of attempting completeness is an impractical goal in the sustainability context, because the danger of irrelevance of the whole MCA evaluation looms larger as the criteria lists become longer. This thesis argues that the main practical way of circumventing the conflicting demands of trying to be complete, while also avoiding double counting and interdependency among the criteria, is to focus especially on higher order criteria of the ‘highest degree of

⁵⁰ Also named respectively, environmental and value independence.

⁵¹ Yates (1990, p. 351) sees the following major generalisations about ‘personal representations’ of the decision situation: “(1) they are often limited in scope; (2) dominated by concrete explicitly presented information; (3) their elements tend to be course and often affective, not merely cognitive; and (4) those elements typically are richly interconnected with one another in several ways.”

⁵² According to some authors on MCA, to tackle the issue of interdependence, stakeholders should artificially separate value or preference independence from physical dependency: “[applying MCA in such a setting] requires a willingness and ability on the part of the user to think of the system in ways that may not be possible *in reality* (e.g., increase in temperature with no change in sea-level rise)” (Bell et al., 2001, p. 248). However, it seems a bit odd, to say the least, to claim that such a highly artificial separation can be a real solution. The issue of choosing a more individual or collective perspective – the dual preference structure – was already discussed above; but adding the extra issue of asking stakeholders to adopt a fantasy perspective seems to be of little value.

importance' (Yoon and Hwang, 1995, p. 9 – referring to Pardee, 1969).⁵³

4.4.3 Trade-off difficulties

As we now know, MCA requires the weighing of criteria, which is one of its most crucial and powerful aspects. Giving weights implies that the different criteria are completely compensatory (Martinez-Alier et al., 1998). After the weights have been established there is some specific trade-off relation among all criteria: a loss on one criterion can be compensated with a specified amount of gain in another. How realistic is this precision in trade-off relations in the sustainability context?⁵⁴

One should be aware that, as a matter of principle, accountability is best served if the weights are determined in a separate step, in isolation from the particular choices at hand (Stewart, 1992, p. 571). Only then can there be some objective justification for decisions, because MCA in which the weights are determined in the process of evaluation has numerous, sometimes subtle, possibilities for manipulation. However, in section 4.2 we saw that many authors doubt whether weights can be sensibly given without knowing the approximate size of the impacts of the specific alternatives under consideration. This position seems to make the establishment of weights in advance impractical.

If weights are best determined in close connection with the relevant measurement range, then the aspect of scaling comes into play here too. In MCA there is the necessity of scaling the measurement of the criteria to a fixed range. MCA is flexible in accommodating any type of input, from qualitative to quantitative, and transforming them into a ratio scale one way or another. This flexibility is a major strength of the technique; but this strength also has an important weakness that seems to be especially relevant in the accountability context. Through the transformation of the impact measurement to a judgement scale the measurement becomes very specific to the decision at hand. The most important drawback of this transformation is that the evaluation becomes less transparent to non-involved stakeholders. Measurement is no longer 'natural' in say, million Euro cost, or number of people affected. This scaling procedure makes interpretation of the outcomes and their crucial determinants, the trade-off relations, more difficult, since the natural reference points, both in size and dimension, are lost. This problem increases the distance between involved and non-involved stakeholders.

An essential point is that trade-off relations may be especially arduous to establish between higher order criteria. As Janssen notices in the assessment of Dutch MCA experiences with large value trees, the lower order weights usually do not create much controversy. They are often based on expert opinion or theoretical knowledge, thus for these weights accountability is less of a problem. The higher order weights however tend to be increasingly political and much more likely to provoke conflict (Janssen, 2001, p. 104).

⁵³ See also Özelkan and Duckstein (1996, p. 48) who focus on the most important criteria "in terms of discriminating power on the alternatives" and Olson (1997, p. 18) who speaks of focussing on the most important criteria to prevent 'analysis paralysis' coming from trying to include "everything that could be of importance to the decision-maker."

⁵⁴ Compare Voogd (2004, p.214): '...in public decision-making situations the use of [...] 'true weights' derived from questionnaires is entirely meaningless!'

4.4.4 Conclusions on interdependence of criteria and double counting

4.4.4a Interdependency and double counting is hard to avoid in a complex decision-making setting

MCA identifies judgement criteria relevant for a decision. In the sustainability context this generally leads to very long lists of criteria originating from the various concerns of different stakeholders. Completeness is well served by this long list but non-redundancy – the demand that the list contain only mutually-exclusive items – suffers from it. To fruitfully apply MCA the analyst must understand causality between aspects and thus the interrelatedness between criteria. However, the sustainability context often has feedbacks that can only be understood when looked upon from higher spatial levels, and there is usually uncertainty there about the relation between Triple E impacts.

A clear judgement structure may be too difficult to achieve in this complex decision-making context with many interdependencies. In such a setting stakeholders may prefer unaided decision making over the allegedly simple but actually complex MCA procedures of giving weights.

4.4.4b Large value trees make MCA less effective – Focus on the most important criteria may help

Practice shows that in evaluation processes that actively involve stakeholders there is great danger that the value tree becomes unwieldy. The upshot of this situation is that the judgement concept in the end is not crystal-clear. One way around the conflicting demands of completeness and non-redundancy seems to be a stronger focus on the higher-order criteria of the highest importance.

4.4.4c Partial judgement will often be the maximum result of an MCA

As the problems of giving weights are far fewer for lower order criteria than for higher order criteria, MCA in the sustainability context may be seen as a process of discovering the areas where stakeholder consensus about compensation is easy to achieve and those where such a consensus is hard to accomplish. In the situation of lack of consensus on higher order trade-offs, but greater consensus on lower order criteria-weights, only partial judgement is possible.⁵⁵ This situation seems to be quite common in evaluation in the sustainability context.

4.5 MCA and its judgement orientation

4.5.1 Introduction

Applying MCA in the sustainability context has difficulties mainly because of the complexities of the problem situation which hinders the establishing of trade-off relations, and because the benefits of stakeholder involvement seem to be limited. Although this chapter focussed on finding a judgement-oriented MCA approach, we can now conclude that this type

⁵⁵ Alexander, 2001, p. 320-321 remarks: “The inescapable conclusion ... is that the stage of identifying and prioritizing goals and objectives, which is intrinsic to all multi-objective decision-making methods, is one we could do without.”

of MCA approach does not function well in the context at hand. Stakeholders will experience difficulty in reducing different values or objectives to a single measure –whether monetary or not – and there will be incommensurability of values. (Martinez-Allier, et al., 1998; Adler and Posner, 1999, p. 247; Aldred, 2002). Partial judgement may therefore be the most that one can achieve in the sustainability context.

The question remains as to whether there is a less judgement-oriented MCA alternative that is more suitable for this context than the MAUT-MCA approach. We identified two types of MCA as being relevant at the beginning of chapter 4:

- Descriptive methods and outranking methods
- Value function approaches.

Our focus thus far has been on the second approach. We ask next if the use of descriptive and outranking approaches is possible.

4.5.2 Descriptive and outranking approaches

First we can explore the descriptive approach in MCA; this approach is regarded as the most basic form of MCA (DTLR, 2001, p. 4.3.2) in that it is an MCA that ends with a performance matrix: each row describes an alternative and each column describes the performance of the alternative against each criterion. Since the measurement in a performance matrix is natural for every criterion, as no scaling has occurred, this matrix has maximum measurement clarity to every stakeholder. It may already be a very important result of an MCA analysis as we observed above. However, the performance matrix will generally be far too large to warrant easy interpretation, and a further attempt is necessary to reduce the number of the criteria.

The outranking approaches try to reduce the set of non-dominated alternatives by formalising the degree of dominance. It may in certain cases lead to a full ranking of alternatives. Various outranking approaches exist. Most of them can be traced to one of the most popular decision aid approaches: ELECTRE (Roy, 1990b; Brans, 2002).

ELECTRE uses a partially compensatory decision rule. Like the MAUT-MCA above, it does require weights and criteria. Basically, ELECTRE counts how often an alternative outranks another on the available criteria (concordance). Judgements are made per criterion. The more an alternative outranks other alternatives the higher is its concordance score. This result is modified by the discordance score, which defines unacceptably low levels of performance on criteria. If an alternative scores below one of these (veto) levels, it is either eliminated or its concordance score is reduced. (Stewart, 2003). Essentially alternatives are compared pair-wise for every criterion.⁵⁶ The concordance score gives the *power* of an alternative. The discordance score gives the *weakness* of an alternative (Brans, 2002). Although it was stated above that the outranking counts the number of times an alternatives outranks another on the separate criteria – which implies an implicit unitary weight for all criteria – in reality most of the outranking approaches use the non-unitary weight of the criteria for calculating the strength and weakness

⁵⁶ See for instance Vreeker and Nijkamp (2001) for an outranking application based on the so-called Regime analysis.

of an alternative.⁵⁷ This procedure obviously resembles the MAUT-MCA approach discussed above. It is not impossible to further introduce elements of the outranking approach in a value function approach like MAUT; especially if one accepts that one can reiterate and do MCA in rounds, the differences may not be that large. However, accountability could suffer from using such a procedure (Stewart, 1992).

4.5.3 *Decision aid versus decision-making approaches in MCA*

The crucial issue at stake here can best be understood by making a distinction between approaches that focus on *decision-making* and approaches that focus on *decision aid* (Roy, 1990): MCDM and MCDA, respectively. The decision aid approach (MCDA) is closely connected with the outranking and descriptive approaches in MCA. The decision-making approach (MCDM) is identified with the value function approach that was central to the discussion in this chapter. Several authors argue that in the sustainability context more attention should be given to decision aid (Vincke, 1992; Bana E Costa and Pirlot, 1997, Rauschmayer, 1999 and 2001; Martinez-Allier, et al., 1998; Faucheaux et al., 1998).

Multi-criteria decision-making (MCDM) often requires strong restrictions on the structure of the decision-makers' preferences (weights and criteria) and the possibilities for handling the inconsistencies in these preferences. The behavioural realism of these restrictions is not often addressed (Korhonen and Wallenius, 1997; Svenson, 1998, p. 352-354). Approaches more focussed on decision aid claim to provide a useful alternative, as they better allow for incommensurabilities of values, ignorance and uncertainties, and fuzziness of the set of feasible alternatives. These approaches will obviously have a somewhat weakened judgement orientation; no precise numbers indicate what should be done (Rauschmayer, 1999, p. 22).

The debate between MSDM and MCDA is not always clear. Is there an essential difference in approach or is it merely a matter of degree, and a matter of the specific context of the evaluation? For our purposes here, it is necessary to settle this issue. Therefore, below the main distinguishing features of a decision-making approach versus a decision aid approach are shown. According to Roy, one of the 'founding fathers' of the decision aid approach, the two approaches differ in many respects. Table 4.2 shows the major differences of MCDM and MCDA as seen by Roy (1990a). The differences relate to their 'general framework' and to their 'crucial preoccupations' within these frameworks.

⁵⁷ It should be noticed that outranking approaches are less sensitive to introduction of new alternatives - in the sense that new alternatives may change all earlier rankings (Arrow and Reynaud, 1986).

Table 4.2 The differences between decision-making and decision aid approaches

Multi-criteria decision-making (MCDM)	Multi-criteria decision aid (MCDA)
DIFFERENCES AS TO THEIR GENERAL FRAMEWORK	
<i>I Alternatives</i>	
A well-defined set of feasible alternatives	A not necessarily stable set of potential actions/alternatives (frequently modified through the decision process itself)
<i>II Preferences</i>	
A model of preferences, well shaped in a decision-maker's mind, rationally structured from a set of attributes. The decision-maker refers to those and only those attributes.	Criteria (and pseudo-criteria, or threshold criteria) reflect preferences of one or several actors; acknowledging that among areas of firm conviction lie hazy zones of uncertainty, half-held beliefs or, indeed, conflicts and contradictions. There should be as wide a consensus as possible about the definition and measurement of criteria, but there is no need for consensus about weights.
<i>III Mathematical formulation</i>	
A well-formulated mathematical problem aimed at finding the highest 'utility' implicitly present in a decision-maker's mind	An ill-defined mathematical problem based on the family of criteria - deemed relevant somehow by actors - and inter-criteria information

Source: Roy, 1990a.

Inspecting this table and keeping in mind the MAUT-MCA approach presented here thus far, ambiguity nevertheless remains. The MAUT-MCA approach is generally regarded as typical of a decision-making approach, but the difference concerning the 'well-shaped' preferences did not appear in the discussion above as 'mess reduction' and 'inconsistencies in preferences' ranked high. Relating to the alternatives and the mathematical formulation Roy emphasises that the MCDA approach is more open than its counter part. Decision aid is geared towards helping stakeholders either to shape, and/or to argue, and/or to transform their preferences, or to make decisions in accordance with their goals. Decision-making approaches allegedly take alternatives and preferences as fixed and aim at 'optimal' solutions.

The point here seems to be somewhat more subtle than mere technical precision however. Von Winterfeldt and Edwards (1986, p. 421) explain the main benefits of using the MAUT-MCA approach. They argue that the major benefit of performing a MAUT analysis is *not* in precise measurement⁵⁸ but more in:

- 1) eliminating dominated options, i.e., by reducing the risk of choosing disastrous options, and
- 2) being more certain about the fact that the options are really close calls.

The main advantages of MAUT according to Von Winterfeldt and Edwards lie in structuring the problem, finding new options, and sometimes collecting better information:⁵⁹ a striking resemblance with the aims of the decision aid approach that Roy advocates.

These observations may act as a warning not to exaggerate the differences between value function approaches and outranking approaches (Stewart and Losa, 2003). The differences

⁵⁸ Also because of the inevitability of using averages at many points in the analysis.

⁵⁹ See also Bana E Costa et al. (1997, p. 30): "...[MCA] facilitates learning about the problem and the alternative courses of action, by enabling people to think about their values and preferences from several points of view." Remember that critics of CBA saw this as a major weakness of CBA; e.g., Niewijk (1994, p. 67): "...people must *construct* a value rather than simply retrieve a previously formed value...."

between outranking and MAUT-like approaches do not seem to follow the division decision aid versus decision-making as a matter of principle;⁶⁰ MAUT-like approaches can be used in a decision aiding style.⁶¹

If they are used in a decision aid style, performing a MAUT analysis is “a constructive process of dis-aggregation and re-aggregation in a new reality of the problem, building a new awareness of [...] preferences” (Stewart and Losa, 2003, p. 658). Outranking approaches, according to Stewart and Losa, are marked by a fundamental prudence to present no stronger conclusions than those that are credible in the light of the preference information.⁶² MAUT analysis first generates a seemingly precise answer and then inform decision-makers of the limitations of the suggested precision. The conclusion must then be that there is no fundamental difference between the MCDM and MCDA methods as such. The difference lies more in how the techniques are used and how limitations are presented.⁶³

4.5.4 Assessment of the judgement orientation

4.5.4a MCDM and MCDA difference is not a matter of technique – Decision aid is the correct style of using MCA in the sustainability context

The most important conclusion for our purposes is that when using MCA – *any* technique of MCA – the sustainability context requires that one adopt a decision aiding approach. Different techniques can be used under that heading: both MAUT and outranking. The important but subtle change is that when moving from a decision-making to decision aiding approach of MCA, the understanding of the main objective of MCA changes. If one acknowledges the difficulties of ‘over-articulation’ in MCA in a sustainability context, one is led to a different understanding of MCA. MCA is at most an aid in decision-making; providing full judgement may not be possible. MCA cannot provide decisions; it can only facilitate the decision-making processes. Practically speaking, adoption of a decision aid approach in MCA should lead to a greater receptivity to include a degree of non-compensation between criteria in the model; and it may result in a degree of incomparability between alternatives (Stewart, 2003, p. 658).

By being more aware of the limitations, the analyst will focus more on decision aid and be very attendant to structuring and framing the problem situation as the priority phase in decision-making (Wiedeman, 1978). The essential function of MCA then is to help to shape,

⁶⁰ See Hammond et al., 1999 for a clear ‘borderline’ exposition of MCA as both decision-making and decision aiding.

⁶¹ Compare Roy’s observation “solutions obtained by solving well-formulated MCDM problems constitute a fundamental background for MCDA” (1990a, p. 34).

⁶² Obviously people can articulate their values to an extent, but it seems reasonable to accept that there are limitations. Not every analyst is aware of this fact, however. Fischhoff discussed the hazards of what might be called asking for under- and over-articulation of values. At the ‘over-articulation’-side of the spectrum, Fischhoff speaks of the paradigm or philosophy of articulated values. “Investigators working within this paradigm have enormous respect for people’s ability to articulate and express values in the most diverse topics.” (Fischhoff, 1991, p. 839). If we look at the MAUT-MCA approach and assess its usefulness in the sustainability context, the danger of over-articulation is quite relevant. The prudence of the outranking approach is in line with Fischhoff’s view of trying to prevent over-articulation of preferences (Fischhoff, 1991).

⁶³ Again a quote from Roy may be illuminating (1990a, p. 17): “MCDM bases its claims to legitimacy on a framework in which these limitations [on objectivity] *are left aside*” (italics inserted). Compare also Woltjer (2004, p. 53) who comes to the same conclusion on the importance of ‘style’ in the field of planning problems: ‘It seems that many aspects of the rational planning approach should be safeguarded when adopting a communicative planning style based on consensus building.’

and/or to argue and/or to transform preferences. In a decision aid approach MCA is “designed to facilitate decision-making by constructing a set of keys which might enable the actors to go forward” (Bana E Costa and Pirlot, 1997, p. 565). Simplicity and the related issue of the possibility of communicating the results of the analysis are prerequisites for an effective use of MCA (Ibid, p. 564-565).

4.6 Conclusions: MCA’s performance in the sustainability context

In the final part of section 4.2 conclusions about MCA’s performance in the sustainability context were grouped and divided into those that had become clear and others which require further discussion. The whole chapter can now be assessed and conclusions drawn about MCA’s performance in the context of accounting for sustainability (see Table 4. 3 below).

Our findings concerning the issues further explored in sections 4.3 to 4.5 generally indicate that the sustainability context is too demanding to warrant the simple use of MCA in a decision-making style.

Stakeholder involvement is no panacea in the sustainability context, as the number of potential stakeholders is large and representation is usually not realistic under such a condition. Stakeholder involvement can at best aim at reaching a broadness of perspectives that are useful towards discovering consensus areas in judgement, weights and measurement (section 4.3).

The sustainability context moreover is conducive to large value trees, which in turn undermines the effectiveness of MCA. Focus on the most important criteria seems to be the most feasible modus operandi; by so doing one can also avoid redundancy (double counting) of criteria (section 4.4).

We see that the sustainability context requires a decision aid approach in which several MCA techniques may be used (section 4.5), but that, in the end, only partial judgement is a realistic expectation.

Table 4.3 Conclusions on MCA's performance in the sustainability context (per stage)

Conclusions: MCA
Stage one: Identify objects and functions of the evaluation
(1) MCA does not stress the need for a 'nul' alternative
(2) MCA reflects on the functional context and tailors the evaluation to the project
Stage two: Identify stakeholders
(3) Involvement of stakeholders makes the political and communication skills of the analyst a critical factor for successful evaluation
(4) Practical impossibilities limit representative stakeholder involvement; broadness of perspectives can be achieved especially for measurement consensus
Stage three: Elicit and organise criteria
(5) Triple E impacts can easily and equally be incorporated into MCA
(6) Values are structured in the process of MCA
(7) MCA depends on measurement (options) for clarifying and structuring judgement criteria
(8) Large value trees make MCA less effective – Focus on the most important criteria may help
(9) Interdependency and double counting is hard to avoid in a complex setting
Stage four: Assess the relative importance of criteria
(10) Diametrically opposed interests hamper finding a convincing common perspective
Stage five: Estimate scores of alternatives on low level criteria dimensions
(11) Achieving comparable measurement implies information loss and affects judgement
(12) Interdependency between knowing the relevant range of impacts and giving weights is not ideal for accountability
(13) The preferred personal interactivity of stakeholders with MCA analyst is a complicating factor
(14) The 'open' treatment of impacts spread over time within MCA contrasts sharply with CBA's discounting practice
Stage six: Aggregate scores with relative weights
(15) Partial judgement will often be the maximum result of an MCA
(16) Especially higher-order weights may be difficult to agree upon; weights within lower-order criteria are often easier
(17) Structuring and framing of the problem situation – decision aid – is the main realistic purpose of MCA in a sustainability context
Stage seven: Perform sensitivity analyses
(18) Additional analysis may give some extra judgement possibility in case of trade-off difficulty
(19) MCA is commonly viewed as one of many steps in a decision-making process

CHAPTER 5: MCCBA: INTEGRATING CBA AND MCA IN THE SUSTAINABILITY CONTEXT

5.1 Introduction

This chapter will synthesise the findings of the previous chapters. While chapters 3 and 4 explored CBA and MCA individually, the current chapter will consider CBA and MCA simultaneously. The analysis has thus far revealed that CBA and MCA as evaluation tools in a sustainability context have pros and cons; we have highlighted the limitations of CBA and MCA and discussed their proper use. The concluding tables of chapter 3 and 4 illustrate an overview of the findings; they list 24 conclusions on CBA and 19 on MCA. The discussion in this chapter is based on those conclusions. In the text below, the conclusions will be referred to by the number they have been given in the tables to which either CBA or MCA is added before it.

We present an integration of CBA and MCA in this chapter that seems to be a fruitful addition to the evaluator's toolbox – a useful tool in the sustainability context.¹ This new tool will be called 'MCCBA', which stands for 'multi criteria cost benefit analysis'. The name suggests that the approach is a new combination of *existing* elements of CBA and MCA rather than something altogether new.

MCCBA is discussed in a threefold way. First, we identify the common ground between the two methods that allows for the integration of CBA and MCA. Second, three major building blocks are presented that further specify the style of integration. Third, the stages of the new approach will be explicated.

5.2 Common ground for integrating MCA and CBA in MCCBA

5.2.1 Full judgement is often impossible in a sustainability context

The discussion of the judgement concept in the complex sustainability context has shown that an analyst's high expectations are likely to be frustrated.

CBA aspires to have the 'utility' of all individuals in an impact population as the judgement criterion. However, utility as a broad category is hard to compare across individuals in general, and the sustainability context assures that a very broad range of widely differing welfare impacts are relevant (CBA10, 16). Although it is ideal in limited cases (CBA 12, 13, 14, 19) CBA's procedure of measuring utility or welfare changes through monetary valuation is fraught with difficulties when applied in the broader sustainability context (CBA 9, 11, 7, 17, 20). If the analyst wants to remain practical, he or she must lower the level of ambition, and a less general judgement perspective (based on everyday economic realities) has to be

¹ Integration of CBA and MCA is mentioned by several authors as a possibility but is not worked out. Janssen and Padilla (1999) may be seen as a rare – practical – example in the literature combining CBA and MCA. Liu et al., (1998) may also provide some inspiration, while Lichfield's (1996) Community Impact Evaluation – which stretches CBA with an elaborate distribution analysis of impacts – may also be seen as a combination approach. Compare Howarth et al. (2001, p. 176): "...CBA might be combined with some scoring and weighting procedures (MCA) for those impacts that are not readily monetized."

realised; the use of CBA should largely to be limited to 'efficiency' and market-related or financial impacts (CBA 4, 12, 13, 21, 22).²

In a way, MCA is more practical, since it starts with the specific concerns of decision-makers and stakeholders, as well as the aspects they think are relevant to the project evaluation at hand (MCA 2, 3, 6). However, building a commonly accepted judgement criterion in the sustainability context with many different stakeholders worldwide who have diametrically opposed interests is not a realistic goal (MCA 4, 10), the more so since the MCA literature clearly shows that judgement is closely related to measurement options (MCA 7). Therefore, 'full judgement' or finding a judgement concept that integrates every concern of every stakeholder, will often not be possible (MCA 12, 13, 15).

Perhaps the most important lesson to be learned from the discussions of CBA and MCA is that determining which is *the best* project alternative in a sustainability context may be unachievable; there may often be no objective or otherwise satisfactory way to judge what is best for everyone affected.³ Metaphorically speaking, using the CBA or MCA 'ladders' may very well help the analyst acquire a better perspective on the project impacts, but these ladders will not allow him or her to reach for the sky, with its clouds of 'social welfare at large' or 'all concerns of all stakeholders'.⁴

5.2.2 *Decision aid is the realistic aim*

Now where do we go from here? Aiming for the realistic target of decision aid seems to be the best way forward (MCA 15, 16, 17; compare O'Connor, 2000a). The analysis has shown that it is impossible to pass full judgement on projects in a sustainability context, but some important partial judgement is well within reach.

CBA is often seen as a valuable tool for assessing efficiency, but how to handle the distributional aspects in CBA is a fierce topic of debate. The approach of giving distributional weights to impacts affecting different stakeholder groups really seems to turn CBA into a form of MCA: It diverts CBA away from its objective stance and its focus on actual everyday economic behaviour. The distributional weights have to be given by an individual decision-maker or a small group of stakeholders. Giving these weights may be very difficult and CBA practice in recent years has settled for second-best (CBA 4, 22). The giving of weights is no longer advocated: mere description is often preferable. Mere description, or rather, careful description can be seen as keeping 'an MCA branch' or 'a specific area of concern' separate; alternatively, it implies a movement from full to partial judgement. The underlying reason for this movement seems to be that it is now unreasonable to accept preferences as given; they must be discovered and formed during the evaluation process.⁵

MCA's difficulties with higher-order value judgement in complex decision-making settings mirror CBA's problems with distributional impacts (MCA 11, 12, 15, 16). In the sustainability context partial judgement will commonly be the maximum result of an MCA. For lower order

² A disaggregated format may have major advantages in these circumstances (Nash et al., 1975) but as will be argued below, this would best be one of the steps of an integrated MCCBA approach.

³ Compare Morgan and Dowlatabadi (1996, p. 337): "We argue that many conventional tools for policy analysis are not adequate to deal with the climate problem."

⁴ See Doane (2000, p. 3): "...an all-encompassing and essentially meaningless set of stakeholders."

⁵ Zerbe (1998) sees given preferences as a prerequisite for CBA, and Richardson (2000) who argues that CBA is 'stupid' precisely for that reason.

weights, these value judgement problems need not arise because a wide consensus on how to provide – partial – judgement is often possible (MCA 16). This situation implies that the aggregation of individual impacts can be pursued as long as the aggregates remain conceptually clear to broad groups of stakeholders and the aggregation weights are largely undisputed (compare Moss et al., 1996).⁶

MCA and CBA alike require both the structuring of impacts and the measurement of effects in ‘physical’ units; these elements are needed before the impacts are unified to a common measurement or judgement scale (stage 2 and 3 of CBA and stage 3 and 5 of MCA).

The structuring of impacts requires listing all impacts and organising them hierarchically without redundancy and double counting. Physical measurement in natural units is the first step before trade-off valuation takes place.⁷

Although CBA and MCA theory naturally focus heavily on valuation issues, evaluation theory and practice have shown that increased problem understanding is commonly the major result of any evaluation: in other words, it is not so much full judgement that is offered by evaluation but decision aid.⁸ The two elements of structuring impacts (or judgement criteria) and physical quantification are major steps in this process of increased problem understanding.

5.2.3 The common ground

The common ground for integrating CBA and MCA can now be formulated. The main point underlying this integration is that the complexities of the sustainability context make both MCA and CBA at best an aid in decision-making (MCA 17, 19; CBA 10, 22). Trying to determine what is best is often beyond the capabilities of the techniques; this underlines the danger of over-ambition in trying to find too simple solutions for passing judgement in a complex evaluation context. It is increased problem understanding that should be the aim of a judgement-oriented evaluation approach.

Understanding this point will lead to a different style of conducting MCA and CBA, in which both CBA and MCA tools, or elements thereof can be used for structuring judgement information and framing the problem situation.

Structuring and framing judgement-oriented physical information is the major aim whenever passing ‘full’ judgement is too difficult. The structuring and framing can be pursued aiming for more and fuller judgement, as long as broad groups of stakeholders understand and agree upon the partial judgement elements.⁹

⁶ A parallel situation is seen to have arisen in corporate performance measurement, where financial measurement is combined with non-financial measures (as for instance in the Balanced Score Card) without formula based aggregation to one final indicator of performance. Instead, room is left for discretionary subjective evaluation (Ittner and Larcker, 1998, p. 228).

⁷ If both methods highlight an impact as being relevant, but physical quantification proves difficult, then the application of both methods is hampered, although MCA can more easily switch to qualitative assessments.

⁸ Blair et al., 1997 argue that in a decision-making context with uncertainty, incomplete information and risk, a shift occurs from solving the problem or producing a ‘correct’ model of the decision to gaining insight into a difficult decision.

⁹ See Rosenhead and Mingers (2001, p. 2) on problem structuring methods: “the aim ...is both more modest and more ambitious than that of the previous generations of optimising methods. More modest, because they do not set out to capture a single truth about the situation from which the one best answer can be derived. More ambitious, because their aim is rather to provide useful assistance to those processes of dialogue and debate which prepare the way for decisions that significantly affect future prospects.” Compare also Jepma and

To some of the more theoretically geared CBA and MCA analysts this limitation may not be easy to accept; should not the aims and ideals be high? Those analysts may be helped by considering the issue of the use of evaluation findings. Evaluation literature has shown that evaluation use is in general rather low (Rossi et al., 1999). There has been prolonged debate on why this situation occurs and how it can be improved (Chelimsky, 1997, p. 15), but the major point is that evaluation practice is far from a luxury position where any judgement information will easily and rapidly be assimilated by decision-makers and stakeholders. If it were in such a position, it would be wise to aim high. The current position of evaluation use is quite different, however (MacCrimmon, 1999). Evaluators have to struggle to get small pieces of their information through and to prove being understandable and useful to users. In an illustrative study Nyborg (1997) showed that the majority of Norwegian politicians, if they used CBA at all, they used it as a mere indicative screening device: an alarm signal, a rough selection or a starting point for questions (Nyborg, 1997, p. 387).¹⁰ In MCA literature Bouyssou has argued convincingly that despite the final operational definitions of criteria, it is usually estimating an ‘order of magnitude’ of the impact that is the only realistic result (Bouyssou, 1990, p. 68), while Wenstøp and Seip (2001) find that the quality of the evaluation says nothing about its significance to decision-makers (compare also Weiss, 1988 and Voogd, 2004). Starting from this low-use position, realism in the evaluation aims is clearly a better course than overly ambitious ideals.

The central research question of this thesis was twofold:

How can CBA be combined with MCA to make a better tool for accountable project evaluation in a sustainability context?

The sub-questions to be answered in Chapter 5 were: Can a project evaluation approach be formulated that minimises the number of cons and maximises the number of pros of both CBA and MCA? What would be the essence of such a Multi-Criteria Cost-Benefit Analysis (MCCBA)? By now, the first question can be answered with a positive yes. CBA and MCA can be combined on the common ground of clear partial judgement accessible to broad groups of stakeholders. We will next specify in detail what the essence of such an MCCBA approach looks like.

5.3 Major building blocks for integrating MCA and CBA in MCCBA

5.3.1 Introduction

On the common ground of looking for structuring and framing of judgement information CBA and MCA can be integrated. Before turning to a stage wise description of this approach three

Munasinghe (1998): ‘When public or collective decisions must be made, it is essential to have a transparent set of decision-making criteria and tools around which a consensus can be *developed*’ (italics inserted).

¹⁰ See also Little and Mirlees (1994, p. 225): “We have found that the extent to which [the project evaluation rules and procedures] are used and have real influence is not great.” Furthermore, Boardman, et al., 2001 may be illuminating. They describe an ex-post evaluation of two different CBA analyses of the same project and conclude (p. 486) that: “This exercise is somewhat humbling.”

major building blocks of this integration can be identified:

- Stakeholder involvement
- Standardisation
- Judgement and measurement.

5.3.2 Stakeholder involvement useful for broadness of views and for checking on consensus

CBA is silent on stakeholder involvement: no involvement is necessary as the analysis is designed to be objective and it mostly uses existing revealed preferences. Although it is not widely acknowledged, the final basis for CBA is the consensus among stakeholders that its procedures are correct and useful as a judgement tool (CBA 4). In the sustainability context this consensus can easily be absent or may only be partly available.

Conversely, MCA is deeply grounded in decision-makers' involvement, which has over time extended to broader groups of stakeholders. However, in the sustainability context identification and limitation of the relevant stakeholders to a 'representative' set of stakeholders does not guarantee success and is often impossible. The list of (potential) stakeholders in the sustainability context may be far too long, their interests may conflict too strongly, and new stakeholders with new values may pop up continuously (MCA 4, 10, 13).

Stakeholder involvement in MCCBA is seen as a tool that aims at incorporating a broadness of perspectives. In management science, the stakeholder perspective was introduced because it put names and faces on otherwise abstract and distant groups in society (Freeman, 1984). Involving a limited group of stakeholders with a broadness of perspectives can then in the context at hand best be seen as a mechanism for checking the extent to which stakeholder consensus exists.¹¹ Realistically speaking, the focus for finding consensus in the sustainability context should be more on (higher and lower order) measurement issues than on (higher order) judgement issues.¹²

5.3.3 Standardisation of global long-term Triple E criteria structure

The benefits of a standardised format seem to be well established in CBA, but far less so in MCA. The most telling example of the standardisation of CBA is the government prescription of a discount rate.¹³ Furthermore, CBA practice has many elements that are more or less standard practice: estimating income effects for instance or not including both income and employment impacts. In as much as CBA moves away from its everyday economic basis standardisation becomes more difficult. This situation can be seen most clearly in the use of CVM valuation methods for environmental impacts. CBA in this respect closely resembles MCA and the

¹¹ See O'Connor (2004a, p. 190): "A successful valuation study requires that the analysts identify, interpret and confront the concerns of actors in multiple ways, so as to permit appraisal of whether there is a dominant legitimacy order used as a reference by several, if not all, actors (including the researchers themselves) or whether the situation is marked by concatenation of different legitimacy orders." See also Munda (1996, p. 166) who mentions 'extended peer communities' as a means for assessing quality in evaluation, and O'Hara (1996) for a more in depth treatment of an 'extended peer' discourse with "generating consensus among participants, at least on part of their agenda" as an important element.

¹² See also Ward et al. (2003) on 'opening up' citizen juries, and difficulties of aiming for consensus.

¹³ Despite, as we have seen, theoretical arguments against standardisation.

development of CVM methods seems to move CBA clearly in the direction of MCA. The results of a CVM-based valuation are best understood in the specific evaluation context in which they were derived,¹⁴ and attempts at standardisation and transferring values from one context to another have failed dramatically.

MCA has far more difficulties concerning standardisation, as not only the criteria will differ for every evaluation, but the scaling of the measurement will be also unique for every evaluation. The scaling of variables is closely connected to the giving of weights, and it involves many subtleties and interdependencies, which are not always well understood.

The sustainability context has several standard elements that can be summarised as concern about global, long-term Triple E impacts. MCCBA introduces these impacts as standardised elements of its evaluation concerns. This approach is more down-to-earth and specific than CBA's notion of broadly defined welfare changes of the complete impact population. At the same time MCCBA provides a more standardised criteria structure than MCA, which usually depends on newly developed structures for each and every evaluation.

5.3.4 Judgement and measurement are closely related

The third major building block of the MCCBA approach is the close relation between judgement and measurement. This close union is often stressed in MCA literature, and it plays an important role when defining criteria and giving them weights. However, not all stakeholders involved in MCA will understand the impact of measurement issues on judgement. The discussion of CBA can be seen as showing the problems one encounters when the two – judgement and measurement – do not match easily. CBA on the judgement side tries to capture all concerns of the complete impact population, while on the measurement side it is in practice often restricted to monetary measurement.

MCCBA evaluation in the context of accounting for sustainability requires both judgement and measurement to be understandable to broad groups of stakeholders.

5.4 Stages of the MCCBA approach

5.4.1 Introduction to the eight stages of MCCBA

We recall that both CBA and MCA had seven stages. The integrated approach of MCCBA has eight stages, which combines elements from the equivalent stages of either CBA, MCA or both, usually in a somewhat adapted form (see Table 5.1). Table 5.1 also shows the 'equivalents' of the stages of CBA and MCA by putting them on the same row. It shows that stage one and two of CBA and stage one and three of MCA easily match, although they do differ in important respects. For stage two and four of MCA there is no equivalent in CBA, while stages four, five and six of CBA are the equivalent of stage six of MCA. The table shows further that stage seven of the MCCBA approach seems to have no equivalent in either CBA or MCA, while stage 8 is common to all evaluation approaches.

Putting equivalents on the same row facilitates formulating the new stages for the MCCBA approach. These stages are shown in the first column of table 5.1. The eight stages comprise a

¹⁴ Although we have seen that serious difficulties remain even here. Formulated differently: the lack of transferability across studies/contexts has shown that values are often misunderstood in the specific individual studies/contexts.

standardised approach. Naturally in any specific evaluation setting other elements may be added, and other tools may be used. The elements sketched below nevertheless seem to be standard for any evaluation of the types of projects considered in the context of accounting for sustainability. They can be seen as the minimum elements that the analyst may prepare for.

Table 5.1 The eight stages of the MCCBA approach – compared to the stages of CBA and MCA

MCCBA	CBA	MCA
Stage one: Identify function, project alternatives and scale of the evaluation	Stage one: Define the project (or policy) and impact population	Stage one: Identify objects and function(s) of the evaluation
Stage two: Involve a broad group of stakeholders		Stage two: Identify stakeholders
Stage three: Organise judgement criteria on Triple E impacts	Stage two: Identify project impacts	Stage three: Elicit and organise value dimensions/ attributes
		Stage four: Assess the relative importance of value dimensions/attributes
Stage four: Quantify impacts physically	Stage three: Quantify relevant impacts physically	Stage five: Estimate scores of alternatives on low level dimensions
Stage five: Aggregate monetary scores consensus based	Stage four: Value relevant effects monetarily	Stage six: Aggregate scores with importance values
Stage six: Aggregate non-monetary scores consensus based	Stage five: Discount costs and benefits	
	Stage six: Apply the Net Present Value test	
Stage seven: Interpret trade-offs		
Stage eight: Perform sensitivity analysis and reconsider project alternatives	Stage seven: Perform sensitivity analysis	Stage seven: Perform sensitivity analysis

The content of these eight stages will be discussed and explained in turn. In chapter 2 we noticed that overestimation of the applicability of a evaluation method is quite a common mistake. Therefore, in the discussion below focus will be on the major elements of the MCCBA approach that give *direction* to the evaluation. To facilitate grasping the essentials of every stage easily, key words and sentences are underlined.

5.4.2 Stage one: Identify function, project alternatives and scale of the evaluation

In this stage important preliminaries are specified. This stage comprises three elements:

- identify function of the evaluation
- define preliminary project alternatives and
- determine the spatial or organisational scale of the analysis.

Identify function of the analysis

MCA at the start of the analysis explicitly addresses the functional limitations of the analysis (MCA 2). In the MCCBA approach, this element is seen as extremely useful. As the common ground (see section 5.2) has made clear, the major functional limitation in the context of accounting for sustainability is the absence of an objective way to determine ‘the best’ alternative. The analyst has to understand that the most one can do is to clear as much ground as possible. To undertake this the analyst has to work with judgement concepts which are necessarily partial, but are acceptable and clear to many stakeholders¹⁵. The function of the evaluation will, by necessity, be limited to increasing problem understanding and to clarify what can and cannot be integrated and understood.

MCCBA implies a style of evaluation in which both CBA and MCA tools can be used for structuring judgement information and framing the problem situation. This structuring and framing process aims at increased and expanded judgement, as long as broad groups of stakeholders understand and agree upon the partial judgement elements.

Define preliminary project alternatives

Both CBA and MCA require the definition of project alternatives, and naturally in MCCBA this is also an important step.

In the MCCBA approach two elements may be more explicitly added. First of all, CBA puts more stress on the formulation of a nul-alternative, which is closely connected with the long time horizon adopted in CBA. This practice, although not uncommon in MCA, is not always explicitly adopted in MCA. The specification of a nul-alternative against which the project alternatives are assessed, seems to be very useful to standardise and is therefore an element of the MCCBA approach. The second element is the preliminary nature of the alternatives. Given the functional limitations, MCCBA should by itself be seen as an important element in finding new alternatives (MCA 19).

Determine the spatial or organisational scale of the analysis

In the sustainability context many interdependencies exist that are not complex or difficult to understand as such, but are poorly visible because they occur on different spatial levels. Choosing a relevant spatial or ‘impact-population’ level is quite natural to CBA. Assessment at different spatial scales, although not commonly practiced, provides no problem to CBA (CBA 2). MCA is more open in choosing how to evaluate (MCA 2, 6). It is however perfectly possible to have different scale levels in MCA, and this is practiced too (see for example Clayton and Radcliffe, 1997; Janssen and Padilla, 1999; Becker and Soloveitchik, 1998). The MCCBA approach adopts strictness in the evaluation at different spatial levels, as it will help clarify impacts in the sustainability context and prevent unnecessary and unwarranted

¹⁵ Remember Keeney’s ‘jointly held objective conceptions’ from chapter 4 and Mishan’s ‘ethical consensus’ from chapter 3. Already in 1975, Nash, et al. argued that using CBA in a disaggregated format (like Lichfield’s Planning Balance Sheet) was better suited to meet the growing demands of public participation.

perceptions of complexity among stakeholders.¹⁶ Globalisation of the economic system has made worldwide impacts a common everyday phenomenon, thus global level assessment is a prerequisite for sustainability evaluation.

The global level is standard in the MCCBA approach and by obviously being the highest spatial level, it acts as a standardised insurance against overlooking impacts. If after serious consideration no Triple E impacts are deemed realistic above a certain high international/spatial level, for example Europe, then this lower level can of course be the highest evaluation level.

However, in most evaluations there will be other spatial levels closer to the decision-making, including the regional, national level or organisation. To remain practical the scale of the analysis should stick close to what the decision-makers feel is the natural scale for the analysis. In many instances, the global level will seem somewhat far-fetched. The MCCBA approach remains practical by adopting another level in the evaluation, which stays close to the 'natural' level of decision-making¹⁷.

Other scale levels can be added depending on the circumstances of the evaluation but things rapidly become complicated and hard to oversee as more levels are added.

5.4.3 Stage two: Involve a broad group of stakeholders

In this stage of MCCBA a broad group of stakeholders is involved. Above it was already argued that representative involvement of stakeholders in MCCBA is not a realistic option. Direct involvement of stakeholders may be very useful however (compare Patton, 1997). Actively involving stakeholders is very suitable in the 'open' decision-making context for which the MCCBA approach is developed and it may increase accountability as it implies making values and judgement criteria explicit (MCA 6). In the MCCBA approach the involvement of stakeholders should primarily focus on achieving a broadness of perspectives, which may act as a mechanism for exploring consensus among stakeholders. This seeking of consensus has to start from lower-order measurement and judgement issues and work upwards. The highest-order judgement issues will usually be beyond consensus.¹⁸

5.4.4 Stage three: Organise judgement criteria on Triple E impacts

In this stage the judgement criteria are organised. This stage combines the MCA procedure of organising criteria, with CBA's insights in – market related – causality and avoidance of double counting.

MCA is well suited to incorporate separate Triple E elements in a common judgement structure. MCA normally uses multiple criteria which are often hierarchically structured. Triple E judgement can easily be merged into an MCA structure, in which the separate Triple E elements are high in the hierarchy of criteria and are treated in a comparable way (MCA 5).

¹⁶ Refer to the recommendations of Turner et al. (2000).

¹⁷ Compare Wilbanks (2002, p. 102): "Sustainability science needs to be sensitive to *multiple* scales rather than focused on a single scale."

¹⁸ Wenstøp and Seip (2001) discuss an MCA on oil spills in Norway: "There was no agreement on how to directly measure damage to life in the sea. Therefore, tons of oil discharged in the sea was used as a proxy variable."

CBA in the sustainability context has more difficulties than MCA in achieving comparability of Triple E concepts and measurement. These difficulties became clear in the discussion of the incorporation of environmental impacts in CBA (CBA 16, 17). Environmental non-use impacts are not naturally expressed in monetary terms. These impacts in the CBA evaluation have in practice become somewhat disadvantaged compared to everyday economic impacts. MCCBA in the sustainability context has two standardised ideas that give direction to organising criteria:

- Define separate Triple E criteria without redundancy and double counting
- Analyse short and medium-term impacts separately from long-term impacts

Define separate criteria for Triple E impacts without redundancy and double counting

Economic development, environmental degradation and extreme poverty are common concerns within the sustainability context. The extent to which CBA can cover the relevant impacts naturally depends on the specific evaluation. However, generally speaking, it has become clear that the measuring rod of money cannot always capture all these impacts and that for a substantial part of these impacts MCA has greater potential. Because the concerns of environmental degradation and extreme poverty are quite standard, the analyst should attempt to standardise treatment of these impacts in an MCA structure. The common MCA approach of continuously developing an evaluation from the beginning, without aiming for standardisation, simply is not very efficient (Brouwer et al., 1999, p. 48). As judgement and measurement are closely related, one should try to use or develop standardised measurement of the criteria; this standardisation preferably holds on different spatial scales.

Analyse the short and medium-term impacts separately from the long-term impacts

In the sustainability context long-term impacts are seen by many as very relevant, especially for environmental impacts. Two specific problems occur: uncertainty and how to treat the time preference. Can a form of discounting be applied for all Triple E impacts like the one that is common in CBA?

The discounting practice in CBA is a very effective way of aggregating impacts spread over time and has no equivalent in MCA (MCA 15). This practice is furthermore widely understood by broad groups of stakeholders in as much as it relates to everyday economics. However, it was argued above that even for economic development impacts the logic of treating long-term impacts differs from those of the short and medium-term. The reason for this difference is that not much of the everyday economic discounting, which large groups of stakeholders – including policy makers and CEOs – are familiar with, stretches over the long-term: uncertainty looms large in the current dynamic system. The concept and rationale of discounting for assessing profitability or efficiency measurements over long time periods is quite different from those over the short and medium-term (Revesz, 1999; Atherton and French, 1998). This situation may make it useful to distinguish short-term and medium-term impacts, say no longer than 15 years, from the long-term impacts, which go beyond 15 years.¹⁹ It was shown that for environmental impacts and extreme poverty impacts, the discounting has no generally (or widely) accepted basis (CBA 19, 20). Although there may be some time-preference in these areas of concern – i.e., beneficial impacts are preferred to occur as soon as possible and negative impacts as late as possible - this fact as such does not lead to a specific discount rate that can be determined and applied in a standardised way. The best way to

¹⁹ Compare Atherton and French (1998) who explicitly use and derive values for different time frames.

proceed is to make time considerations explicit, but keep the overall structure simple by distinguishing between long-term impacts and short and medium-term impacts.

To summarise in MCCBA, impacts are separated into:

- 1) short/medium-term impacts and
- 2) long-term impacts.

Ad 1) For the short and medium-term impacts on economic development, the CBA practice of discounting can be applied, but the other Triple E impacts are shown in their own right. Ad 2) All long-term Triple-E impacts are shown in their own right too. Conceptually these impacts, because of their long-term relevance are on a more equal footing without discounting, and probably a more MCA-like treatment may yield better results.

Further treatment in MCCBA has to depend on the specifics of the particular evaluation; more experience with MCCBA at this point may lead to a more refined standardisation.

5.4.5 Stage four: Quantify impacts physically

In this stage the measurement of impacts in their natural dimensions takes place. As was observed earlier this ‘fact finding’ phase is probably the most important phase in any evaluation, although methodologically not much needs to be said about it.²⁰ Evaluation is based on measurement. Aspects that cannot be measured have no place in a formal judgement-oriented MCA that wants to be accountable to stakeholders.²¹ This limitation also gives the analyst the opportunity to streamline the concerns and different values of stakeholders: they may be asked for the measurement they think is suitable for their values. As measurable criteria are a subset of all relevant concerns, the analyst has a possibility here to lower the number of criteria, and clear out minor differences in meaning that stakeholders attach to different criteria (MCA 7).

Stage four ends with the performance matrix, as it is called in the MCA literature. Measurement of the impacts is in their ‘natural’ dimensions. In a very basic way, this can be regarded as the most crucial stage of the evaluation. The raw data of the evaluation are gathered here. If for some reason one of the criteria is measured qualitatively here, instead of quantitatively, that will always hinder the judgement possibilities. Still, MCA has several possibilities to accommodate this type of measurement. The performance matrix then shows the impacts for the three Triple E areas²² in which the results for the different scale levels and time blocks may be presented separately. The assessment at different scale levels presents a major challenge to physical measurement in the environmental domain (e.g., Faucheux et al., 1993b, p. 73; Nijkamp and Van den Bergh, 1997).

In the sustainability context this performance matrix will usually be too large and thus too complicated to make clear interpretations. Stages five and six therefore become crucial as they reduce the information to a smaller number of dimensions.

5.4.6 Stage five: Aggregate monetary scores consensus based

In this stage of MCCBA a CBA is performed for the impacts that fit well within it. The

²⁰ Compare the ‘two arms of evaluation’ of Scriven: data-gathering and values (Rossi, 1999, p. 21).

²¹ See for instance Cohon (1997).

²² Separate in as far as there is non-double counting necessity of course.

limitation to impacts that ‘fit well’ is the reason why it is ‘consensus based’.²³

What should be included in the CBA part? To many stakeholders it is unclear as to how far CBA is able to grasp the impact on social welfare at large. By now, the MCCBA approach can be much more specific. The ground on which the MCCBA approach rests is the notion that CBA’s attempt to grasp welfare at large is hard to defend in the sustainability context, but that CBA remains a powerful evaluation tool in as much as it relates to everyday economic notions such as efficiency and monetary measurement that stay close to market related valuation.²⁴ Only those impacts should be fitted into CBA that broad groups of stakeholders see as well captured within the monetary measurement possibilities of CBA.²⁵

The projects in the sustainability context for which the MCCBA approach is designed have important economic development impacts that warrant the use of CBA as an evaluation tool. In MCCBA the type of CBA proposed is pragmatic. The basis of CBA should be the everyday economics common to many stakeholders. In fact, this implies that the analyst explicitly define something like efficiency as a judgement criterion. Efficiency is a concern so deeply rooted in the current market economy that one may safely state that there is consensus among stakeholders that it is an important criterion. To clarify the difference with mainstream CBA the MCCBA approach uses the term ‘limited CBA’, which underlines the limited scope of this type of CBA.²⁶ Despite being limited, no doubt should exist about the robust quality of this type of CBA. A pure MCA approach, as discussed above, often lacks the crucial economic realism that is well captured in CBA²⁷.

In using CBA one should be aware that discounting over the very long-term, even of efficiency impacts, very often does not have much realism²⁸ or familiarity to many stakeholders.²⁹ This is the reason why in the MCCBA approach these impacts are preferably split into short/medium-term impacts and longer-term impacts. For longer-term economic impacts relevant criteria for every evaluation may be judged, as uncertainty is substantially different, and different concerns may be relevant.

MCCBA’s consensus based valuation with CBA can to some extent use the straightforward but labour intensive procedure of asking for willingness to pay or willingness to accept money values. This procedure is limited to changes that can, in Pigou’s words, “be brought into relation with the measuring rod of money” (Mishan, 1975, p. 166). The approach of asking for willingness to pay is on the one hand a powerful and straightforward way of asking how

²³ Ashby (1980, p. 1180) asks: “.. is it better to restrict economic analysis to values that are unquestionably quantifiable, leaving the decision-maker to integrate the imponderables into the decision by subjective means? Or is it better to stretch economic analysis to cover values that have to be stripped of their human meaning in order to quantify them?”

²⁴ Ashby, 1980 states: “[CBA] sets out to answer the question: what is *efficient* for society?...the public asks:...What is *good* for society?”

²⁵ See Howarth et al. (2001, p. 182): “The efficiency status cannot be determined ... without some form of CBA. Thus CBA can and should be an input to MCA. Proceeding to MCA without CBA can ... be dangerous.”

²⁶ See Liu et al., 1998 for an example of a very limited CBA with results on a few other objectives. See also Cortenraad et al. (1986) for an example of using limited CBA with other criteria for infrastructure planning.

²⁷ Howarth et al. (2001, p. 171) agree that: “What should be avoided is the view that if we cannot monetize everything, nothing should be monetized. This view amounts to rejecting valuable information about people’s preferences. Moreover ... rejecting monetization simply raises all kinds of other problems.”

²⁸ In the sense that real foreseeable alternative opportunities can be counted on.

²⁹ Wenstøp and Seip (2001, p. 59) when discussing a choice between MCA and CBA type of measurement in a specific evaluation setting: “It was concluded, however, that it would be even more doubtful to rely on such [CBA type] judgements for end-impact criteria *so far removed from normal experience of the public*” (italics inserted).

important something is to people. On the other hand, awareness that one is moving away from the more objective and real behaviour-based foundation of observed market data is necessary, as well as knowing that comparability will easily suffer (CBA 16). The further one is removed from common everyday economic valuations, the less reliable WTP or WTA valuations will be, and the more that MCA-like measurement techniques should be preferred.

5.4.7 Stage six: Aggregate non-monetary scores consensus based

In stage six of the MCCBA approach the second possibility of reducing the size of the performance matrix is applied: performing a consensus based MCA on the remaining impacts. The results of the CBA can be seen as a separate MCA branch to be combined with the other branches of the remaining impacts.

The MCCBA challenge here is twofold:

- to reduce the number of criteria to a minimum – in a CBA style
- to use consensus based judgement criteria and measurement.

Reduce the number of criteria to a minimum – in a CBA style

The clarity of the judgement concept will be easily blurred when stakeholder involvement leads to value trees that are far too large (MCA 8). Furthermore, and in relation to that, the complex interdependencies that exist between different criteria hamper the achievement of a judgement structure that is well built and understandable. If the interdependencies between criteria are unclear, very uncertain or very differently perceived by different stakeholders, the application of MCA may likewise become arduous (MCA 9). It is at this point that the MCCBA approach takes its inspiration from CBA's practice when avoiding double counting and incorporating causality. CBA *practice* typically uses indicators that are at the end (or at one point) of a causal chain and avoids using intermediate indicators – as they would easily lead to some double counting. It is only fair to observe however that the inspiration can come partly from the MCA side too, as a thorough analysis of causality to avoid redundancy of criteria is a well-established part of MCA *theory*.

As the final method of reducing criteria, the MCCBA approach focuses on the most important criteria or criterion and leaves aside criteria with either a low weight or with unclear or small impacts. Naturally completeness suffers from this procedure but the gain is the increased likelihood of understanding for many stakeholders – including the understanding of limitations. The selection of the most important criteria should certainly rest upon the increased problem understanding that the evaluation has brought.

Use consensus based judgement criteria and measurement

The second challenge of MCCBA is to perform an MCA that makes use of judgement concepts and measurements about which there is consensus among broad groups of stakeholders, or about which such a consensus can be developed.³⁰ Compared to the performance matrix what happens here is that several sub-criteria can be scaled to make up an overall criterion.

There is a clear role for expert knowledge here. It may not be possible to reduce the Triple E impacts to three single criteria. Still, the possibilities for useful reduction by use of regular MCA-like approaches is far greater in the separate branches of this MCA structure than in the

³⁰ See Woltjer (2001) for an analysis of theory and practice of consensus building.

overall structure.³¹ If one thinks of the impact on biodiversity or the impact on the health condition of extremely poor people, it may be very helpful to involve experts to assess the various aspects and sub-impacts to an aggregated format or to a few indicators.³² This procedure mirrors asking economic experts to assess the impact on efficiency. Naturally choices and assumptions have to be made in these processes, but broad groups of stakeholders may leave these choices uncontested.³³ It is usually the higher order valuations that are most disputed; these contested valuations are avoided as much as possible in the MCCBA approach.³⁴

The end result of the MCCBA is ideally an aggregation of the performance matrix, based on a broad consensus among stakeholders.

5.4.8 Stage seven: Interpret trade-offs

What type of analysis is possible in these final stages of MCCBA? In the standard CBA and MCA approaches the only remaining stage is sensitivity analysis. The major difference between the end result of stage five and six of MCCBA and the more typical CBA and MCA approaches is that not enough judgement is provided. This thesis has argued that although this may be a regrettable state of affairs, in many instances it will be the optimum and realistically feasible result, as accountability is best served with this approach, and the provision of more judgement will likely become uncertain and thus subjective and contestable. However, although in the end one may not be able to pass full judgement on the various project alternatives, there are more possibilities for judgement, while also staying within the confines of an objective and accountable approach. One should take small steps, as non-involved stakeholders need to understand what has happened, and moreover the limitations of the analysis should remain clear to all.

The first goal of this stage is to interpret the consensus-based aggregation of the performance matrix in several ways aimed at analysing trade-offs.

Generally, an important analysis will be analyses of trade-offs by means of ratio-analysis. In such an analysis, ratios are calculated on the basis of the aggregated performance matrix which looks at, for instance, net-CBA outcomes per outcome on another criterion (Stewart, 2003). This type of analysis may be closely related to cost-effectiveness analysis.³⁵

Ratio-analysis may gain in strength if comparisons across projects can be made.³⁶

³¹ See Pomeroy and Barba-Romero (2000, p. 317): “There is no obligation in multi criteria decision to end up with an aggregation.”

³² Many examples could be mentioned here, e.g., Faucheux, 1998; Ten Brink, 2000; Wackernagel, 2002.

³³ Compare Vincke (1992, p. 58): “What is attempted in these [outranking] methods is to enrich the dominance relation by some elements which *suffer no discussion*, by strongly established preferences” (italics inserted).

³⁴ See Lind (1997) about the use of a multi-objective framework without weights.

³⁵ Cost-effectiveness analysis focuses from the start on one non-monetized effect and assesses cost-effectiveness ratios (costs per unit of effectiveness). Cost-effectiveness analysis focuses on the effect on which the intervention is aimed. It may for instance focus on the reduction in the number of students who drop out due to a new educational project or on the gain in travel time due to the development of an infrastructure link. The choice of this one effect is of course crucial for the social relevance of the analysis. As will be clear, the MCCBA approach is wide open at the outset and includes all effects; and, depending on which and how many criteria remain non-monetized in the end – and which prove to be crucial – results may resemble the outcomes of a cost-effective analysis. For more information, see Levin and McEwan (2001) and Petitti (2000).

³⁶ See Keeney (1996) but also consider the CVM discussion of chapter 3 in which transferability across projects was very important.

Standardisation of measurement in project evaluation enhances the possibilities for comparisons across projects.³⁷ For a clear example of this standardised measurement, the Quality Adjusted Life Years (QALY) from health economics is noteworthy.³⁸ Often this type of analysis will not lead to the designation of a 'best alternative'. However, it will provide extra judgement information and will often be the crucial element of the structuring and framing of decision-making central to the MCCBA approach.³⁹

Another related type of analysis in this stage may be to perform what is called a stakeholder perspectives analysis or what Söderbaum (1998, p. 65) has called 'position analysis'; that is, to explicitly adopt the perspectives of specific stakeholders, to assess the importance of different criteria to these stakeholders (what if priority was given to this criterion over these?), and to present what would seem to be a preferred option from these perspectives⁴⁰. Several elements from the MCA toolbox can be used for this analysis⁴¹. Sometimes one will realise that that no matter which perspective is adopted – from those that seem logical – the result remains the same (Alexander, 2001).

5.4.9 Stage eight: Perform sensitivity analysis and reconsider project alternatives

Uncertainty is often central in the sustainability context and therefore uncertainty in the MCCBA evaluation is preferably not treated in a sensitivity analysis that functions as a footnote to the main results (CBA 24). As argued above, uncertainty is preferably integrated in the evaluation by adopting a more open and less ambitious style of conducting evaluations. Still even in the MCCBA approach in the end uncertainty should be brought to the fore.⁴² Stage eight has an important function for this task.

The overall aim of this stage of MCCBA is to share the increased problem understanding developed during the evaluation. Sensitivity analysis is a useful tool in support of this aim. It should be emphasised that the sensitivity analysis may have an important function for accountability. It can act as a check for stakeholders to oversee, as an indicator of which impacts and assumptions matter most to the outcomes, and show how other assumptions – that may be closer to stakeholders' own convictions or estimates – work out on the overall outcomes. This knowledge can be a crucial element in judging the value – and limitations – of

³⁷ Compare for instance Cortenraad et al. (1986, p. 9): "...it is important that each criterion is measured in a standardised manner, so that the effects of various projects can be compared with one another."

³⁸ See Klose (2000) who speaks of cost-utility analysis when QALYs and Cost-effectiveness analysis are combined. For indicators for environmental policy see for instance Adriaanse (1993), Pethig (1994) and Faucheux and O'Connor (1998). For an interesting mixed ratio of environmental costs and economic value, see Vogtländer (2002).

³⁹ An example of Wenstøp and Seip (2001) shows the subtleties of meaningfully calculating such ratios in a misleading comparison of efficiency (cost per energy unit) of different power plants with conflict (and NOT: conflict per energy unit) leading to equal ranking of plants that differed 20 times in energy production – as they had equal efficiency and conflict scores. See also Janssen and Padilla (1999) for interpretation of efficiency (CBA) and environmental (MCA) scores, and Jepma and Munasinghe (1998, p141-146) for a trade-off analysis concerning green house gas emissions on the one hand and the total of mitigation and damage costs on the other..

⁴⁰ 'It is therefore much better to show the consequences of a variety of standpoints, rather than a single interpretation of what would be an artificial, non-existent, 'best weight set' (Voogd, 2004, p214).

⁴¹ The community impact assessment (or planning balance sheet) of Lichfield (1996), with its disaggregated presentation of CBA outcomes to different stakeholders, can also be seen as a perspectives analysis.

⁴² See Fischhoff, 1977 (!), p. 196: "Perhaps the most important aspect of informing the public is for the analysts themselves to point out the limits of their craft."

the evaluation.

Practical experience with sensitivity analyses shows that they can only play the role of sharing problem understanding to an extent.⁴³ Therefore, the MCCBA approach uses another powerful tool for sharing the increased problem understanding: it explicitly asks for a reconsideration of the project alternatives.⁴⁴ The following (types of) questions should be posed:

- Does the increased problem understanding show that new alternatives should be considered?
- Does the increased problem understanding show that existing alternatives should be combined?

Linked to these are questions about the consistency and feasibility of the goals to be reached with the different project alternatives: the increased problem understanding may shed light on this too. The spirit of MCA may be seen as closer to this stage of MCCBA than that of CBA. MCA is commonly viewed as one of many steps in decision-making, and is more commonly used for interactive rounds of decision-making, in which several MCAs may be performed (MCA 19), and this stage of MCCBA explicitly reconsiders stage one. Of course, CBA is in actual fact also part of decision-making rounds, and a sequence of CBAs may in principle be performed. This fact however has little bearing on the focus of the analysis as such and there is no stage in CBA addressing the consequences for the set-up of the CBA.⁴⁵

Ultimately, this final stage of MCCBA leads to a greater focus on increased problem understanding and the discovery of new alternatives as a central part of the evaluation.

⁴³ Compare section 3.2.8: “The trouble with sensitivity analysis seems to be that from the analyst’s perspective more often than not a sensitivity analysis is bound to be somewhat superfluous. It is very hard for an analyst to make the message of the sensitivity analysis something like: the main results can hardly be trusted.”

⁴⁴ A sensitivity analysis might be interpreted as shedding more light on the functional limitations of the evaluation. These limitations, the reader may recall, were part of stage one of the evaluation of which the definition of project alternatives was also part.

⁴⁵ An exception might be found in CBA-based Community Impact Assessment (Lichfield, 1996).

PART THREE – USING MCCBA

THREE CASE STUDIES

CHAPTER 6: INTRODUCTION TO PART THREE

Three chapters about applicability

Part three of this thesis has three chapters in which different applications of the MCCBA approach are discussed. Chapter 7 addresses the evaluation of a major nature conservation project in the Netherlands to transform agricultural land into nature reserves and nature friendly farming areas. Chapter 8 discusses the evaluation of the building of a flood barrier close to the Waddensea in the German Ems River with the double target to prevent flooding and allow for cruise ships of the Meyer shipyard in Papenburg to sail the Ems River to the deeper waters of the North Sea. Chapter 9 treats the evaluation of Sustainable Corporate Performance in general and illustratively sketches an evaluation of the EurepGAP protocol, an initiative of major European food-retailers¹ that aims at safer and more sustainable fresh products.

The main aim of the forthcoming three chapters is to provide insight into the applicability of the MCCBA approach. Although all three chapters contribute to this main aim, they differ in character. They differ first in the aspects of the MCCBA approach most highlighted, and they differ in concreteness and function.

Essential aspects: Improving upon either CBA or MCA

The following three chapters highlight important aspects of the MCCBA approach in practice and in so doing, they also address its strengths and weaknesses. However, the selection of the cases was made because – taken altogether – they emphasise the most crucial strengths of the MCCBA approach. The chapters reveal aspects not normally tackled by either CBA or MCA. Part three could then be understood as an explanation of the major strengths of the MCCBA approach from a practical perspective. What are its major strengths?

The first is its capability to combine the value of non-monetary measurement with monetary measurement; this strength stands out most when compared to the use of CBA only. The evaluation in Chapter 7 shows the benefit of using non-monetary measurement of the impact on nature separately from other monetarized impacts. Unlike CBA, the MCCBA approach can easily accommodate non-monetary measurement. Choosing how to measure depends on the best way of presenting judgement-oriented impact information to broad groups of stakeholders.

The second strength is the accessibility to more spatial scale levels. Chapter 8 shows that structuring the impacts in this way facilitates a clear analysis and presentation of impacts. This element may be seen as an improvement upon either CBA or MCA by using it as a standard approach; familiar to both techniques, it is simply an element that is especially useful in the sustainability context and thus useful to standardise.

The third strength is the value of causal (economic) analysis. Chapter 9 shows how a clear analysis of the (economic, market related) causality – and (again) the evaluation at a high spatial level – may help to reduce the number of criteria in an MCA. Chapter 9 can therefore be seen as an improvement upon MCA by combining it with CBA.

¹One of them the Dutch Ahold concern.

The character of the following chapters

Part two of this thesis argued that MCCBA may be a welcome addition to the evaluator's toolkit; its line of reasoning was mainly theoretical. However, part of the theoretical line of reasoning used here was inspired by earlier practical evaluations of the current author. Some of this work is presented in part three and results in a peculiar situation regarding the function of (especially) chapters 7 and 8. On the one hand, these chapters should preferably simply show how the MCCBA approach works in practice, but on the other hand, these applications were MCCBA applications 'avant la lettre' and for that reason alone they show a mix of elements – some that do and others that do not – follow the MCCBA approach.

To tackle this problem chapters 7 and 8 are structured such that, after a brief introduction, major sections (7.2 and 8.2) present elaborate summaries of the original evaluations. Sections (7.3 and 8.3) assess the applicability in light of the ideal MCCBA approach. Their focus is to discuss the applicability of the ideal approach. There will undoubtedly be several aspects that have proved to work well – as that is the main reason for presenting these applications. However, deviations from the ideal are also inevitable, given the peculiar situation mentioned above. Of special relevance will be to identify *structural* deviations that indicate non-accidental limitations to the applicability of MCCBA.

The function of chapter 9 is somewhat different in that it explores the wider applicability of MCCBA. The chapter first addresses SCP theory and practice and the drawbacks of existing evaluation tools. MCCBA is not completely worked out in chapter 9, but is shown illustratively for a concrete evaluation case. Table 6-1 summarises the differences of the three chapters.

Table 6-1 Distinctions among the three chapters of part three

	Chapter 7	Chapter 8	Chapter 9
Essential aspects of MCCBA illustrated	Value of non-monetary measurement	Value of more scale levels	Value of causal (economic) analysis and high scale level
	Improvement upon CBA sec	Improvement upon CBA and MCA	Improvement upon MCA sec
Concreteness	Concrete	Concrete	Abstract (only a sketch of a more concrete case)
Function in part three	Shows applicability of MCCBA approach	Show applicability of MCCBA approach	Explore further MCCBA application

CHAPTER 7: EVALUATION OF THE DUTCH ECOLOGISCHE HOOFDSTRUCTUUR (EHS): THE VALUE OF NON-MONETARY MEASUREMENT

7.1 Introduction

The evaluation in this chapter was conducted in 1995, five years after the Dutch government presented the Natuurbeleidsplan (Nature Policy Plan; LNV, 1990) to the Dutch Parliament. The primary aim of the plan was “sustained preservation, restoration and development of nature and landscape values.” In order to counteract the decline and crumbling of nature areas, the plan heralded a new strategy; its most important factor was the ambition to implement an Ecologische Hoofdstructuur (EHS; literally: Ecological Main Structure) within the time period 1990-2020. The evaluation focussed on this EHS.

The EHS ($\pm 700,000$ ha) aimed to consist of all existing nature areas (450,000 ha) and add to the so-called Relatienota-reservaatgebieden (nature conservation areas on agricultural land with already relatively high nature values; 99,400 ha), nature development areas (nature conservation areas on agricultural land needing intensive ‘development’ of nature values; 44,200 ha.), and regulated agriculture (100,000 ha). This meant that 244,000 hectares of new nature were to be developed on what was until then mainly conventional farmland.¹ The total area of all Dutch agricultural land and nature areas together was, at that time, approximately 2.8 million hectares, of which more than 80% was agricultural land.

The aim of the evaluation was to accurately assess the costs and benefits of implementing the EHS. To that end the act of realising the EHS was compared with the situation without it. The study was carried out on a national level; impacts for individual provinces or other regions were largely left out of the analysis, as the decision-making level was mainly national.

The evaluation combined CBA and MCA: MCCBA *avant-la-lettre*. Special attention was paid to quantifying the nature values that EHS was to realise, because insufficient attention was given to these values in earlier research. The increase in nature values was maintained as an independent criterion. Costs and benefits that could be expressed in monetary terms were integrated in a Cost-Benefit Analysis. The financial costs and benefits to government were analysed as an independent criterion, separate from the CBA results.²

The evaluation will be discussed by first presenting a summary of the original results³ (section 7.2) and then by assessing from these results the applicability of the ideal MCCBA approach in a sustainability context (section 7.3).

¹ In the Natuurbeleidsplan this figure is 250,000 ha. The difference lies mainly in yet-to-be allocated national reserves.

² In a standard CBA these costs would be considered a mere transfer; that is, a double count that should be left out of the analysis.

³ The results of this evaluation were published in Sijtsma and Strijker (1995a; 1995b); Strijker and Sijtsma (1996a; 1996b; 1996c) and Strijker, Sijtsma and Wiersma (2000) and Strijker, Sijtsma and Bettels (2000).

7.2 Original results

7.2.1 *Brief summary of results*

In this evaluation of the Ecologische Hoofdstructuur (EHS), great effort was made to assess how nature values in nature areas differ from the nature values on agricultural land. To this purpose nature values were quantified by means of the occurrence of a large group of ecologically valuable plant and animal species: the so-called ‘target species’. The conclusion was that, within EHS as a whole, a hectare of nature area contains five times more ‘nature value’ than a hectare of agricultural land. The results showed that regulated agriculture within EHS contains about twice as much nature value than conventional agriculture, while the relation document areas and the nature development conservation areas scored about seven times higher than conventional agricultural land. Based on these findings it could be estimated how much the national nature value would increase due to the implementation of EHS. Given several limitations it could tentatively be concluded that, due to EHS, nature value in the Netherlands would increase by 15-20%.

Part of the evaluation investigated the degree to which the yield of agricultural land changed into nature differs from other farmlands, as the former might have a structurally lower agricultural quality. The conclusion reached was that productive capacity of EHS land was about 10% lower than average Dutch agricultural land.

In order to realise EHS, Dutch society would incur costs. Costs are for instance loss of economic activity in agriculture and agricultural industry (benefits foregone), but costs also involve management/restoration expenses for the nature areas. The total net amount of costs and benefits – that could be expressed monetarily – for the whole project, measured and discounted over a 50-year period, were calculated at €1.5 billion (net present value in 1990; discount rate of 4%). This figure was substantially lower than the net costs arising from other research,⁴ although the method of calculation hardly deviated; differences mainly stemmed from higher precision in assessing the costs.⁵

The realisation of EHS would also tax the government budget. In the EHS evaluation that figure is calculated at €2.3 billion – again a net present value (in 1990) of discounted costs over a 50-year period. This result was more in accordance with the conclusions of other studies.⁶

7.2.2 *Results and method detailed*

7.2.2.1 Nature values

In this evaluation vigorous effort was made to analyse nature values in the best possible way. Until this study no insight has existed into the different types of nature (and their location) that were to be realised under EHS. A comprehensive survey was therefore held among regional policy making bodies and (future) nature area managers in order to develop a systematic classification of the different types of nature used to describe which of them were to be realised. This classification into so-called target nature types was based on an existing manual

⁴ See Slangen (1994) and Oskam (1994).

⁵ See Sijtsma and Strijker (1996b) for a detailed comparison of differences.

⁶ Commissie Financiering Natuurbeleid, 1992; Werkgroep Financiële dekking SGR, 1993; Berenschot, 1994.

with nature reference types (Bal et al., 1995) to which various types of regulated agriculture were added. In total 131 target nature types were identified as relevant within the evaluation: for the ‘with situation’ (realising EHS) 118 nature types were potentially relevant, while in the ‘without situation’ (not realising EHS) 13 nature types – on agricultural land – were used.

The questionnaire results provided the specification of nature types relevant in both ‘with’ and ‘without’ situations, and were interesting in themselves: it was the first time that the probable nature results of EHS had been specified. In the evaluation these results were the basis of a further operationalisation of nature values.

The operationalisation proceeded by using a database specifying which species occur in which nature types.⁷ Not all species were represented: the evaluation used a set of 564 so-called target species distributed over different species groups (see table 7-1).

Table 7-1 Distribution of target species across species groups

Species group	Number of target species
Plants	405
Birds	64
Mammals	16
Amphibians	7
Reptiles	5
Dragonflies	20
Butterflies	47
Total	564

Source: Bal et al., 1995.

Measuring (changes in) nature value was based on the data about the occurrence of target species in different ecosystems/landscapes or nature target types. The criterion for linking species and nature types was whether the nature type was important to the survival of the species. The occurrence of different target species, and the frequency of their occurrence, were viewed in this study as indicators of the quality of the condition of nature.

For every target nature type, three nature scores were established to form three different indicators for nature value. When determining these nature scores we assumed that the more that target species occur in a certain type of nature, the higher its nature value would be.⁸ A second assumption was that nature value rises in proportion with number of hectares realised. The three scores were established as follows:

- (1) Target species per se: number of target species that occur
- (2) Target species with optimality: as (1), but with extra value/weight to target species that occur under optimal conditions
- (3) Target species with optimality under degree of threat: as (2), with extra

⁷ This database was built from two sources: the underlying database of Bal et al. (1995) kindly provided by IKC Natuurbeheer, and additional expert judgement (also by IKC experts) about species occurring in the agricultural target types added to the existing database especially for this evaluation.

⁸ Given the fact that the scores are calculated *within* the nature target type system, anomalies of zoo-like densities of species are excluded.

value/weight to higher rarity, more decline in occurrence and more international significance of the species.

Each hectare of a target nature type was assigned a nature score for the above three indicators. Aggregation of the nature scores of the relevant target nature types produced the average nature score per hectare for EHS as a whole, while the same procedure was applied to the situation where EHS was not implemented. A comparison of the two outcomes shows the improvement of nature value as a result of the realisation of EHS. The three different nature scores are now explained in further detail.

Nature score 1: Target species per se

When calculating the nature score based on target species, every hectare is awarded points equal to the number of target species occurring on that hectare. Table 7-2 shows the results.

Table 7-2 Average nature score per hectare for EHS and non-EHS sites, based on target species per se

	Nature score based on number of target species per se	
	Average score per hectare	Relative index (Non-EHS/agriculture = 1.0)
EHS	58	3.0
Non-EHS	19	1.0
EHS:		
All EHS minus regulated agriculture	79	4.1
- Relation document conservation areas	74	3.8
- Nature development	90	4.7
Regulated agriculture	28	1.4
- Strong regulation	32	1.6
- Light regulation	24	1.2

Source: Bal et al., 1995; own calculations

The database specifies which species would ideally occur in the different nature types. The criterion designed is to simply count database-species in the different nature types: the more species the better.

Adding everything and averaging this indicator on a hectare basis the analysis shows that in the ‘without’ situation – the regular agricultural situation – some 19 species would ideally occur. In the ‘with’ situation of the realised EHS, approximately 58 species would occur on every hectare; this indicator is easy to understand and there is wide consensus about its relevance.

When calculated in this way, a difference of a factor three can be seen between the nature score of conventional agricultural land and that of land when EHS is realised there. Thus on average three times as many target species occur within EHS. Regulated agriculture shows only limited relative benefits of 1.2 and 1.6. EHS ‘without’ regulated agriculture scores four times higher than conventional agriculture; the section nature development scores nearly five times higher.

Nature score 2: Target species with optimality

In order to account for the extent to which a species occurs under either optimal or sub-optimal conditions within a certain target nature type, the optimally occurring species were

multiplied by a factor two (the sub-optimally occurring species were left at a factor one). Table 7-3 shows the results. The absolute score is less straightforward than the first nature score, as it is a synthetic index with an ‘artificial’ weight. The relative scores of the separate EHS elements set as an index against the ‘without’ situation of conventional agricultural land are easier to understand. The difference in these relative nature scores of EHS is somewhat larger than previously (3.9 instead of 3.0). Obviously this difference resulted from the fact that the number of target species occurring in optimal conditions on conventional agricultural land turns out to be relatively small. The initially small difference between the scores of regulated and conventional agriculture has increased somewhat (from 1.4 to 1.8).

Table 7-3 Nature score with optimality for EHS and non-EHS realised sites (based on target species with optimality)

	Nature score based on amount of target species with optimality	
	Average score per hectare	Relative index (Non-EHS/agriculture = 1.0)
EHS	79	3.9
Non-EHS (conventional agriculture)	20	1.0
EHS:		
All EHS minus regulated agriculture	108	5.4
- <i>Relation document conservation areas</i>	105	5.2
- <i>Nature development</i>	116	5.8
Regulated agriculture	36	1.8
- <i>Strong regulation</i>	43	2.2
- <i>Light regulation</i>	29	1.5

Source: Bal et al., 1995; own calculations.

Nature score 3: Target species with optimality and threat

Following a nature value assessment model of Centrum voor Landbouw en Milieu (Centre for Agriculture and the Environment; CLM, 1995), a scale which indicates the nature value, or ‘degree of threat’, of a target species was used. This model measures the nature value of species on a 0 to 100 scale, on which a high score indicates that the species is very valuable in terms of rarity in the Netherlands, international significance and decline over previous years. The CLM nature value is compiled from a rarity value (R), a trend value (T) and an international significance value (I) which are multiplied.

- The rarity value is defined as the maximum number of 1677 *uurhokken* (grid squares of 5 by 5 kilometres on the topographical map of the Netherlands), divided by the number of *uurhokken* in which the species is actually found:

$$R = 1677 / \text{number of } \textit{uurhokken} \text{ where species occurs (minimum of 1; thus } 1 \leq R \leq 1677).$$

- The trend value is defined as the change in population size, expressed as a percentage of the population size in the past.⁹
- $T = (\text{past population size} - \text{present population size}) / \text{past population size} * 100$ (thus $-\infty < T \leq 100$, although negative values were not foreseen).
- The international significance is a value attached to the importance of a species' occurrence in the Netherlands, which assesses the importance of the Netherlands for its global occurrence:

$I = 1 + (0 \text{ to } 0.5)$ depending on the amount of criteria for international significance the species meets (thus $1 \leq I \leq 1.5$).

In order to reach a rating of species on the basis of the three factors expressing the value of species, CLM argued that two arithmetical principles, addition and multiplication, can be used. CLM argued for multiplication because there are strong connections between the components; when a species is rare a downward trend will be worse, and this downward trend will worsen again when the presence of that species has great international significance. Mere addition of the components would express these connections insufficiently. Calculation of the CLM nature value is consequently done by multiplying all components.

CLM assigns greater weight to the components' rarity and trend than to international significance. A species can receive a maximum of 0.5 points for international significance, making the maximum value 1.5. The points value is hence calculated by multiplying the components: $I * R * T$. The minimum conceivable value is 1 point, the maximum conceivable value about 250,000 points. Because these widely varying values are not very practical, CLM has scaled them down to values ranging from 0 to 100. However, an arithmetical downscaling ($I * R * T * 0.0004$) would lead to very few species with a high value, and numerous species with an extremely low value. The CLM therefore applied a logarithmic reduction. Expressed in a formula in which the score of a species (j) is called the Ecological Score (ES):

$$ES_j = 18.5 * 10^{\log(I_j * R_j * T_j)}$$

The ES_j value ranges theoretically from 0 to 100.¹⁰ The value 0 will seldom be reached in practice.

The formula for the nature score (NS) of a specific nature target type i then reads:

$$NS_i = \sum_{j=1}^{Sr} P_{ij} * ES_j$$

i the different types of nature areas (131 in all)

j the different number of species occurring in that type of nature area (subset from the total of 564 in all)

P degree of occurrence of species j in nature type i (dummy variable: 1 for sub-optimal or 2 for optimal)

⁹ The fact that negative values were not foreseen in the CLM method might be a source of error: compare van Diggelen et al. (2005) where, in an ecological-economic analysis of the Dutch province of Drenthe, which included all available species, many species with increasing occurrence were found.

¹⁰ The tenth power logarithm of $I * R * T$ indicates to which power 10 must be raised to reach $I * R * T$. If $I * R * T$ is 1000, then the $^{10}\log$ equals 3; if $I * R * T$ is 100, then the $^{10}\log$ equals 2, etc.; if $I * R * T$ is 1, then the $^{10}\log$ equals 0.

Table 7-4 shows the results of using this calculation method. Again in this table the score of conventional farmland was indexed at 1. EHS as a whole then scores a nature value of 4.9. This score is considerably higher than the 3.0 (target species *per se*) found earlier, or 3.9 (target species with optimality). This difference indicates that agricultural land does relatively little for highly threatened species. The value of regulated agriculture is only marginally higher than in the previously calculated alternatives: regulated agriculture does not offer much extra to highly endangered species. The difference between EHS conservation areas and regulated agriculture is large (7.0 compared to 1.9).

Table 7-4 Nature score with optimality and threat for EHS and non-EHS sites (based on target species with optimality under degree of threat)

	Nature score based on number of target species with optimality and threat	
	Average score per hectare	Relative index (Non-EHS/agriculture = 1.0)
EHS	4423	4.9
Non-EHS (conventional agriculture)	902	1
EHS:		
All EHS minus regulated agriculture	6294	7.0
- <i>Relation document conservation areas</i>	6088	6.7
- <i>Nature development</i>	6759	7.5
Regulated agriculture	1741	1.9
- <i>Strong regulation</i>	2133	2.4
- <i>Light regulation</i>	1342	1.5

Source: Bal et al., 1995; CLM; own calculations.

Limitations of nature scores

The nature scores determined above are merely a quantified indication of the actual changes in nature.

For instance, the list of target species in the database is mainly concerned with whether a target species *could* occur, not whether it *actually does*. This problem turned out to be unsolvable in the present study, although it does not seem unsolvable in principle – given more data. As this bias is relevant both in the ‘with’ and ‘without’ situation, its impact is probably limited. However, within the framework of a result-orientated nature policy, it might be better to base the indicators on an accurate density per hectare of target species in a specific target nature type. In this evaluation a list of target species without highly endangered species was used to test the result of focussing on actually occurring species; the conclusion was that such a list had a very limited effect on the results.

The size of the target species list could also influence results. The target species list consisted of about 20% of all species occurring in the Netherlands, for which a more or less complete description existed at that time. The target species were furthermore selected as representational of other species not on the list. The chance that another, more complete list of species would render a substantially different outcome is therefore small.

The distribution of species might not be taken sufficiently into account when determining the nature score per target nature type, particularly if target species *per se* are used as the nature value criterion. The reason for this fact is that – as seen above – the EHS is of significance, in

real terms, for an *average* of 58 target species per hectare. However, these 58 species are not in every different nature target type: in the Netherlands as a whole nearly ten times as many, 558 target species, benefit from EHS. This situation differs starkly from conventional agricultural land. As seen, approximately 19 target species benefit from an average hectare of conventional agricultural land, but these are largely the same species: over the whole of the Netherlands fewer than twice as many – only 33 – target species benefit. Using the average number of target species – as with the target species per se criterion – is thus shown to provide an incomplete picture. However, this aspect is accounted for, at least partly, when using the other criteria, especially when using the criterion based on threatened target species with optimality – since a high rarity of species cannot be combined with occurrence in common nature types.

A final noteworthy limitation is the assumption that the nature value increases proportionally with the number of hectares; this assumption is quite reasonable for small increases in surface area. Larger surface areas pose problems, however, because in theory the nature value of extra surface areas drops off gradually. The impact of this problem is reduced because the nature target type system also has a scale element. However, the problem could not be completely tackled in the current evaluation.

National impact indication

Based on the results of table 7-4, an approximation could be made as to how much Dutch nature values gain by the implementation of EHS. Due to limited data availability on the nature value of the complete Dutch area, the results are only rough estimates. This change in Dutch nature value is calculated by multiplying the scores from table 7-4 with the corresponding hectares and subsequent comparison of the results for the situations ‘with’ and ‘without’ EHS. The increase in nature value in the Netherlands then becomes 15-20%. Because of the problems mentioned in the previous section, this estimate is of course only an indicative result. This increase, calculated at 18%, originated for nearly 11 percentage points from Relation document conservation areas, a good 5 percentage points from nature development, and for nearly 2 percentage points from regulated agriculture.

7.2.2.2 Agricultural production

Introduction

In this part of the evaluation, CBA style of reasoning was followed, in which the underlying notions of market related causality and prevention of double counting rank high. Concerning the agricultural production impacts, it was argued that the main impact on which to focus was the impact on production value or value added. The major judgement criterion was production value and the change therein because of EHS. The measurement of the impacts mainly required knowledge of the market situation and market perspectives for different crops,¹¹ knowledge of the economic drivers of farmers¹² and knowledge of the Dutch agricultural land-use and geographical situation.¹³ The analysis had a strong everyday economic basis and could be followed by any farmer or other non-expert stakeholder. This economic basis showed that although EHS might be planned on land now used for milk cows, the result at national level

¹¹ For instance, about the EU milk quota policy.

¹² Leading to a replacement scheme on the basis of gross yields per hectare.

¹³ Leading to an assessment of the general productive capacity of EHS designated on agricultural lands on the basis of, for instance, water availability, and parcelling structure.

would be that a certain amount of wheat production would be lost. Furthermore, it led to the insight that the labour and capital involved would not be permanently lost for productive use (outside agriculture). Slangen (1994) had overlooked both these points in an earlier study.

The realisation of EHS may involve heavy emotional impacts when farmers have to sell land on which they had worked hard to improve productivity, with the result of turning it into 'wild' and unproductive area. The land may have, in some families been used for generations. Although this aspect seems very relevant to many stakeholders involved, it was not included in the analysis separately. The estimate was based on aggregate market values: farmers were not asked for their willingness to pay or their willingness to accept payments for certain changes. The aggregate market value approach followed here was an important element in the policy debate to many stakeholders.

Details

When establishing the consequences to agricultural production of converting agricultural land into nature, three important concerns were foremost: first was to determine which crops or livestock were cultivated or held on the lands to be converted; second, to examine the productive capacity of these lands; and third, to ask which agricultural production will be relocated, and which production would disappear.

Present agricultural land-use – Data on what is presently being cultivated on EHS lands were unavailable. In order to form a 'land-map' of these lands despite lack of data, the physio-geographical types¹⁴ in/of every municipality of relevant EHS regions was investigated. On the basis of percentage distribution of categories of physio-geographic types, it was possible to assess what was, in 1995, probably produced on EHS lands.

Productivity – The productivity of EHS lands can differ from the Netherlands average as a whole, because EHS lands are not distributed evenly across the country. In order to estimate this difference, the level of agricultural productivity in the various agricultural regions (as charted by the Centraal Bureau voor de Statistiek [CBS; Central Statistical Office]) was determined first using basic indicators. Analysis based on these data showed that the productivity of EHS lands was at least 2% below average. However, EHS could also incorporate relatively poor land within the agricultural regions. To ascertain the possible impact of this factor, survey respondents were asked for an indication of the productivity of the relevant land. Quantification of these indications led to an additional difference in productivity, making EHS land in total about 10% less productive than average.

Disappearance and allocation of agricultural production – Which type of agriculture production will be relocated and which will disappear? It was assumed that crops /production (crops henceforth) that generate a high added value per hectare will push out other existing crops on non-converted land. Crops subject to quotas such as milk (cows) and sugar beets thus remain, due to their relatively high margin, and products such as grain are partially removed. The eventual loss of production as a result of EHS therefore lies mainly in wheat, green maize, other cereals and roughage.

Loss of added value

In this study the initial drop in net added value of Dutch primary agriculture, as a result of converting farmland into nature, was calculated at €98 million per year. The drop also affected the added value in agricultural industry, which had been estimated at €120 million yearly.

¹⁴ E.g., higher sand areas, sea clay areas, river areas.

These estimates account for losses due to relocation, which constitute for instance temporary stagnation of growth in production as a result of years of uncertainty due to EHS planning/policy process. The overall loss can be put into perspective by comparing it with the yearly national added value for agriculture and agricultural industry, which amounted in 1995 to approximately €20 billion.¹⁵

When determining the social costs these losses ought to be balanced with the alternative production taking place elsewhere in the economy by the production factors labour and capital involved – the factors initially becoming obsolete by the EHS project. Naturally this process of finding alternative uses will take time. A formal temporal scheme for finding alternative productions was implemented: a scheme developed by Oskam (1994). The scheme differentiates between primary production impacts and agribusiness impacts. For primary production the scheme runs as follows: The production value allocated to the production factor land (25%) would be permanently lost. As for the labour and capital part, half would be in alternative use after one year – yielding the same productive value as before – and the other half would find an alternative productive value over the course of 10 years (5% every year).¹⁶ There is no permanent loss for the agribusiness part of production value; 75% of the former production value is realised in other productive uses after one year. After five years this has increased to 100% (again 5% every year).

Environmental value

The fact that implementation of EHS diminishes the area of agricultural land implies a decrease of options for manure placement, which could cause losses to added value in intensive livestock breeding. But damage to the environment would also decrease because the costs that water companies face for distributing clean drinking water would diminish. Since the effects are: 1) both positive and negative 2) inconclusive as to the size of the impacts, and showed that 3) the valuations and locations for the long-term were indeterminate, the change in other environmental values as a result of EHS was noted pro memoriam.

7.2.2.3 Impact on recreation

The EHS changes the nature and landscape of parts of the Netherlands. The Nature Policy Plan regarded improvement of the experience value of the landscape – the enjoyment of the scenery by inhabitants and visitors – as crucial to nature conservation and development. However, an exact description was not delineated and a method to measure landscape experience value was lacking.

At the time of the EHS evaluation a general model for experience value of the landscape was being developed in order to predict the experience value people attach to different landscapes on the basis of its characteristics (building on Kaplan and Kaplan, 1992, 1995, Kaplan et al., 1989) and their functional relationship to it. A few preliminary results could be used in the evaluation for only one drastic landscape change (for final results see Van den Berg, 1999). The results showed that landscape experience value of an initially unattractive area could improve by 15-35%, provided the area is re-landscaped in the most highly valued way. The landscape experience indicator was based on a questionnaire, asking people how they

¹⁵The estimation of this study relied heavily on the *current* situation – in 1995 – in the agricultural sector. Compared to parts of the WRR (Netherlands Scientific Council for Government Policy) analysis (WRR, 1992), this estimation could be considered conservative.

¹⁶ (Sub-) sector specifics were not taken into account.

appreciated different landscapes – mainly on the basis of reference photographs.¹⁷

Recreational impacts are closely connected to landscape experience value. The impact on recreation would be caused by the increased attractiveness of the Dutch landscape. To determine the impact the evaluation focussed on recreational expenditure,¹⁸ which was not regarded as particularly sensitive to an increase in the ‘supply’ of more attractive landscapes as such. Relative preferences might change, however. No net national effect was assumed for day-trip recreation; realising EHS could lead to regional re-allocation of flows but an overall increase in expenditure was deemed unrealistic. Tourist expenditure in the Netherlands was also thought not to increase because of the small increase in landscape attractiveness of the Netherlands. Summer holiday expenditure (i.e., destination choice) and winter sport holidays of the Dutch were also not thought to be sensitive to EHS realisation. The class of recreation expenditure that could be affected were short holidays to Belgium, Germany and France taken by the Dutch. On a yearly basis this expenditure amounted to approximately €330 million. Several experts were consulted and different literature was studied but no convincing way of estimating this effect could be found. Furthermore, large parts of EHS will see relatively little change in the landscape as a result of its implementation. In fact, large improvements from a landscape perspective could only be expected in nature development areas. But these areas form only one-fifth of EHS, and merely 1.5% of the total area of agriculture and nature in the Netherlands. Improvement in the landscape experience value for the Netherlands as a whole therefore seemed to be limited.

EHS also has negative impacts on recreation however. EHS invoked a more restrictive policy towards recreational activities in certain EHS areas. More than before, the EHS area was seen as a pure nature area and not as a mix of both recreation and nature.

Because of the existence of both positive and negative impacts and their limited size, the evaluation put the estimate of the impact at zero – acknowledging the possibilities of regional relocation of recreation activities.

7.2.2.4 Costs of purchase, management and restoration

Government purchased a large part of the agricultural land to be converted into nature from the farmers. From a cost-benefit perspective this involves a transfer which does not seem to affect community welfare. The purchase would nevertheless require public funds. For this study – conducted in 1995 – an amount of €18,150 (Hfl. 40,000) per hectare was assumed. The price was somewhat higher than the land price of €15,880 (Hfl. 35,000) used at that time in longer term government planning (LNV/VROM, 1993).

The management of the purchased land is fundamentally different from that of functional farmland. The purchased lands are overseen by site management organisations: Staatsbosbeheer (National Forestry), Natuurmonumenten (the largest private nature conservation organisation) and the Provinciale Landschappen (the 12 provincial nature conservation organisations). The managerial costs per hectare of these organisations were

¹⁷ The question might arise as to whether one might use a WTP approach here, and ask people what they would be willing to pay for certain landscape changes. Although certainly some support for this approach is valid, it seems that a well-designed WTP questionnaire requires a thorough understanding of how and why landscape is appreciated by different people. Thus, the general model would have been necessary anyway, as it could facilitate decisions about who to ask for what. Compare the trend from CVM to Conjoint Analysis (see chapter 3).

¹⁸ No need to say: a rather down-to-earth proxy for broad welfare changes related to recreation, which, although based on aggregate data illustrated the willingness to pay stance of CBA.

calculated at €360 (Hfl. 800) annually (overheads included). Third parties financed part of this sum.¹⁹ The annual expenditure for the government was €190 (Hfl. 415) per hectare.

In the case of regulated agriculture, the land remains under control of the farmers, who receive government compensation for restrictions to their management of these lands. The effect of enforced restrictions determines whether the regulations are called strong or light. Light regulation is compensated at an average rate of €230 (Hfl. 500) per hectare annually and strong regulation receives €500 (Hfl. 1,100). When assuming that strong and light regulation is divided equally, the average compensation for regulated agriculture is €360 (Hfl. 800) per hectare annually; the costs are entirely covered by government.

Purchased lands require restoration for the adequate development of nature values. Standard figures estimated by the government vary between the different types of land. On average, this once-only amount adds up to nearly €1,820 (Hfl. 4,000) per hectare. Once again, the costs are borne entirely by the Dutch government.

7.2.3 Overall evaluation of the results

7.2.3.1 Introduction

The elements discussed thus far are not easily combined in an overall evaluation. A number of the impacts (agricultural production, recreation) can be converted into effects on community welfare that are quantifiable in monetary terms by valuing and aggregating them into a Cost-Benefit Analysis (CBA). The effect on public finance, however, is different. Although it can be valued in monetary terms, it cannot be dealt with in the CBA, because payments by government do not have automatic consequences for social welfare: they are largely a mere transfer among different citizens. The changes in nature values cannot be included in the CBA either, because nature values are not valued in monetary terms.

In this research a Multi-Criteria Cost-Benefit Analysis – *avant la lettre* – is used as an instrument so that the various elements can indeed be assessed jointly. Three criteria play a role in the Ecologische Hoofdstructuur (EHS) project: the balance of the CBA, the increase in nature value, and the acceptability of the demand that EHS places on public funds. The latter should be viewed as a pre-condition rather than as an independent criterion for which a trade-off relation can be established; it will therefore be treated first.

7.2.3.2 Government finance impacts as a threshold criterion

The total Net Present Value of the discounted costs over 50 years was €2.3 (Hfl. 5.1 billion). The yearly costs average at €90 (Hfl. 200 million). The maximum in year 2020 is €170 million (Hfl. 375 million).

The criterion of the government finance impacts as mentioned above has a different character than the other two. As a criterion with a vaguely defined ‘threshold’ or ‘ceiling value’, it should not be exceeded. Government budgetary policy requires that the total amount required is important, but so is the spread of costs over time. An analysis of the foreseen annual flows for the next 50 years was performed. The non-discounted flows were expressed as a share of the expenditure in the overall budget (0.1%), and the share that different elements of EHS had within the total outlays was examined.

¹⁹ This part is excluded in the calculation for the demand put on government finance.

The maximum annual (non-discounted) costs amount to €170 million (12.1% of the present budget of the Ministry of LNV (Ministry of Agriculture, Nature and Fisheries). Over 50 years, the cost is on average €100 million annually (approx. 0.1% of the total national government budget). Table 7-5 recapitulates these figures.

Table 7-5 Government finance impacts by EHS

Total claim on public funds by EHS (discounted over 50 years)	€2.3 billion
Annual development of costs (not discounted)	average over 50 years: €100 million maximum in 2020: €170 million
Overall tentative judgement	essentially feasible

In all, looking at the 12.1% maximum and the 0.1% of the total state budget, the expenditures seem feasible; although the amounts needed are substantial and they might continue to arouse debate, the absolute amount could be fitted in. In that sense the – not exactly specified – ceiling does not seem to have been exceeded. Obviously this procedure is rather informal and can easily be challenged.

7.2.3.3 The two major criteria in perspective

If one accepts that the size of government expenditure does not exceed an acceptable amount, then this criterion can be dropped and focus on the remaining criteria: the nature value and the net balance of the CBA is possible. The two criteria can be compared without formal weighing. However, to have a clear perspective of the evaluation results at a national level, it may help to view these criteria in the context of relevant national totals.

Regarding the nature value, the calculated indicative gain in national biodiversity due to realisation of EHS is relevant here. As seen above, realisation of the EHS yields an 18% estimated gain in national nature value. However, the aspect of sustainability of these nature values is not weighed here and the estimated degree of occurrence of the target species could also be improved upon.

The results for the relevant elements in the Cost-Benefit Analysis (including restoration, management and agricultural production) are calculated over 50 years, based on a discount rate of 4%. This leads to a negative balance of €1.5 billion (Hfl. 3.4 billion) for the entire project.²⁰

The total net cost may be related to the total volume of domestic production of more than €255 billion in 1992. The calculated total discounted ‘once-only’ costs of €1.5 billion are 0.6% of this figure; the average annual costs over 50 years amount to 0.03% of GDP.

A permanent gain of 18% in national biodiversity could thus be achieved with a once-only cost of 0.6% of the domestic product, or with costs spread over a 50-year period of 0.03% of the GDP (see table 7-6).

²⁰ Net Present Value of discounted flows over 50 years.

Table 7-6 Results of Multi-Criteria Analysis at national level

Nature value	18% permanent gain national biodiversity
CBA of effects that can be incorporated therein	(once-only) 0.6% GDP (annually over 50 years) 0.03% GDP

Since costs largely consist of loss of agricultural production, they can, apart from this national picture for the entire economy, also be viewed in light of the total volume of primary agricultural production, or better still, the production of the entire agricultural industry.²¹ The total loss of added value to agriculture (agricultural industry included) due to the realisation of EHS is 1.5%. The production factors that ‘suffer’ from EHS could largely be put to alternative use – thereby gradually diminishing the loss to society – but this alternative use is not necessarily within the agricultural sector.

7.2.3.4 Further interpretation per component of EHS

Table 7-7 shows the scores for the separate criteria of the various components of the EHS.

Table 7-7 The overall impacts of EHS, on the three MCCBA criteria and the relative contribution of different EHS components

	Hectares	Criterion 1: Nature value	Criterion 2: Costs (limited CBA)	Criterion 3: Claim on public finances
EHS TOTAL (Total effect)	243,535	18%¹	Euro 1.7 bln (discounted)	Euro 2.6 bln (discounted)
EHS TOTAL (%)	100%	100%	100%	100%
EHS conservation area	59%	90%	71%	81%
Relation document conservation area	41%	60%	45%	53%
Nature development	18%	30%	26%	28%
Regulated agriculture	41%	10%	29%	19%
Strong regulations	21%	7%	20%	13%
Light regulations	21%	3%	9%	6%

1 – 18% gain in national nature value, based on the nature value indicator that weighs rarity, trend in occurrence and international significance of species

EHS conservation areas comprise 59% of the hectares generate 90% of the nature values against 71% of the costs to society and 81% of the costs to government finances. Regulated agriculture comprises 41% of the acreage and leads to 10% extra nature values, against 29% of the costs to society and 19% of the total costs to government. These figures can also be

²¹ According to LEI figures, primary agricultural production comprised 2.8% of the national income (€7.1 billion) and production of the entire agricultural industry - primary agriculture included - 6.9% (€17.5 billion) (LEI-DLO, 1995b). It was calculated that the initial loss of added value to the agricultural industries (primary agriculture included) resulting from the formation of conservation areas amounted to €217 million (1.2%). The loss to regulated agriculture of €36 million could be added to this figure.

used to calculate ratio indicators of the relative efficiency of the various EHS components.

Efficiency with regard to costs to society

The efficiency of the components in terms of social costs is then of interest first of all. The contribution of EHS to national nature value amounts to 18%, as seen above. When subsequent costs to society of a component of EHS is divided by the gain in national nature values realised by that same component, a picture emerges of the cost efficiency of the separate components. Table 7-8 shows these results.

Table 7-8 Cost efficiency of EHS components

	Costs to society per percentage point national gain in nature value	
	Euro million per percentage point	index (EHS = 100)
EHS	86	100
of which:		
EHS conservation area	68	79
Relation document conservation area	65	75
Nature development	75	88
Regulated agriculture	251	291
Strong regulations	234	272
Light regulations	298	346

On average, a gain of one percent in national nature values is realised by the entire EHS for an estimated €86 million in social costs. Of the separate components, the *Relation document* conservation areas have the lowest social cost per percentage gain in national nature values, followed closely by the *Nature development* areas. The difference between these components is caused by restoration costs, which are much higher in the case of nature development sites compared to those of conservation areas.²² The worst score in terms of social costs per percentage gain in nature value is reached by lightly regulated agriculture. Regulated agriculture with strong restrictions scores better.

The last column in table 7-8 shows the costs per nature value reinterpreted as an index of the average for the entire EHS. The separation between conservation areas and regulated agriculture is clear. The costs to society of regulated agriculture are three to four times higher than those of EHS conservation areas.

Efficiency with regard to government finance

In a similar way, the efficiency of EHS components with respect to public funds used can be shown (see table 7-9).

²² Noteworthy is that contribution to nature values by nature development areas might be underestimated, because initial nature values of development sites are lower than those of conservation areas. On the other hand, realisation of nature values in nature development areas will take longer.

Table 7-9 Efficiency of EHS components with regard to government finance and nature values

	Costs to government Per percentage point national gain in nature value	
	Euro million per %	index (EHS = 100)
EHS	130	100
of which:		
EHS conservation area	117	90
Relation document conservation area	115	89
Nature development	120	92
Regulated agriculture		
Strong regulations	251	193
Light regulations	234	180
	298	229

Despite the high costs of purchasing the sites and restoration, in proportion to nature values realised, the formation of conservation areas is more cost efficient than regulated agriculture. The order of the various components is the same as regulated agriculture with regard to the social costs: the *Relation document* conservation areas score the highest, closely followed by *Nature development* sites. Regulated agriculture with light restrictions scores the lowest, while agriculture with strong restrictions scores better than its light counterpart.

Sensitivity analysis

A sensitivity analysis of the CBA elements in the evaluation was performed. For several elements of the CBA, the estimates of the size of the impact was varied with a +50% and a -50% change, respectively. The analysis showed that the most sensitive cost items were management costs of conservation areas and nature friendly farming areas.

Another part of the sensitivity analysis was performed by comparing outcomes of this evaluation with outcomes and assumptions of two earlier evaluations (Oskam, 1994; Slangen, 1994). To this purpose some of the differences with the other evaluations were classified as minor – differences in discount rate, precise number of hectares and length of the period of analysis. Differences were certainly very important to the outcomes, but they represented style differences or differences in knowledge about standardisation. A total of 12 other assumptions were classified as important because they differ between the three evaluations. Major deviations were caused by adoption or non-adoption of a replacement scheme for the production factors in agriculture, and the scheme showing the relocation of agricultural production on the basis of yield per hectare (see 7.2.2.2).

7.3 The Ecological Main Structure evaluation in light of the MCCBA approach

7.3.1 Introduction

The evaluation of EHS will be discussed here within the MCCBA approach defined in chapter 5. This evaluation is a typical case of project evaluation in the sustainability context (see chapter 2) and important strengths of MCCBA have been used. Compared to CBA alone, the first strength of MCCBA proved to be the nature impact measurement; both its elaborate separate treatment and its aggregation clearly indicate how MCA elements may operate to provide judgement insight. Compared to MCA alone – the monetary aggregation and non-

double counting of the several economic impacts reveal a second strength, as it elegantly reduces the number of criteria.

The applicability of the MCCBA approach will be assessed in detail below by delineating the eight MCCBA stages. The constant focus will be on establishing how practicable the approach is. To this end the analysis will outline how well the case has followed the ideal approach, and will especially note whether deviations from the ideal have occurred for more structural reasons.

7.3.2 Stage one: Identify function, project alternatives and scale of the evaluation

7.3.2.1 The function of the research

The function of the EHS evaluation was first to assess both costs and benefits, and second, to use the results as a reference point in the ongoing debate in Dutch policy on the merits of EHS. With this function the evaluation, came close to the ideals of the MCCBA approach outlined in part two, as the MCCBA approach explicitly aims to be accessible to many stakeholders.

The function of the actual research differed from its initial function as proposed by Vereniging Natuurmonumenten; the major private nature conservation organisation of the Netherlands. Realisation of EHS implied a substantial growth in nature conservation areas; it would halt a negative trend that had been going on for decades. However, fierce debate on the (high) costs of EHS had been underway for years prior to the evaluation. Natuurmonumenten decided to fund new research on the costs of EHS, as they felt confident that a new study would show that the costs would be lower than other studies. Different economists, c.q., institutes were approached for a research proposal with this function.²³ The function in the winning proposal of the evaluation was changed; although researchers tended to agree that earlier studies had overestimated costs involved, they also agreed that a different approach would be more fruitful: if the EHS aim was to improve nature in the Netherlands, an evaluation of that policy should attempt to evaluate the nature results as well. The function of the evaluation was therefore changed to a double function: to assess both costs and benefits and to act as a reference point in the debate.

Acceptance of this change meant that not only was the initial function as foreseen by the initiator still followed, but also the two new functional elements at least had a chance of being beneficial.

7.3.2.2 The project alternatives

The selection of the project alternatives does not require much explanation: the nul-alternative was defined as not realising EHS, the policy alternative as reaching it.

However, during the research the evaluation team decided to try and estimate costs and benefits for different parts of EHS corresponding to the different policy instruments used. As shown above, EHS was divided into areas where nature friendly adjustments to regular farming are practiced and into nature conservation areas. Both categories were subdivided into two groups. Although initially there was only one policy alternative, in practice different options for improving nature quality in the Netherlands were evaluated by distinguishing costs

²³ The funds available were approximately €40,000.

and benefits of the separate policy instruments.

7.3.2.3 The scale of the evaluation

In EHS evaluation the scale of the analysis was the national scale only. Was this a structural deviation from the ideal MCCBA approach or would the inclusion of other spatial scales have been better?

In the EHS evaluation inclusion of the regional scale was explicitly considered, as was inclusion of the European scale. The evaluation focussed on the national level because EHS was a national plan and all the major decisions about it were made at the national level.

The evaluators rejected the regional level because, according to them, it did not add anything extra, and although this choice was made early in step 1 of the research, it was not absolutely given. If, during the research, impacts relevant to specify at the regional level had been found, the evaluators would have reconsidered.²⁴ Non-treatment of the regional level does not therefore imply a significant deviation from MCCBA

To not take into account the European level is a more complicated issue. When looking at impacts and impact population, there were basically two reasons for not including this level. First, EU agricultural policy (the existence under the Common Agricultural Policy of milk quota for instance, and the Less-favoured areas policy) were relevant to the project. Second, the nature values realised in the Dutch context might be examined from a European or international perspective.

To not include the EU level for the first reason seems justified. At the EU level there neither would be a relevant change in budget flows nor any substantial change in impacts related to broader CAP instruments. Perhaps it is rather superfluous to include a high international level if, in the end, it does not indicate an important impact.

Not including the EU level for the second reason is more questionable, as the international dimension in nature values is relevant and interesting. It is a well-known fact that national nature conservation generally aims at protection of only species or biotypes that are rare nationally, but are abundantly available across country borders. Including the EU level is something for which the MCCBA approach – if it had been fully available at that time – could have helped. Although measuring international impacts would have been problematical, on a national scale some important data were available on the international importance of the Netherlands for different species.

To summarise, it seems that the ideal of including a high international or global level in the MCCBA approach is not irrelevant for this case; it could have been given practical content and possibly important outcomes too.

7.3.3 Stage two: *Involve a broad group of stakeholders*

As noticed above, the function of the research was to act as a reference point in the discussion about EHS among many stakeholders. Active involvement of stakeholders was also practiced. In the EHS evaluation the initiator of the research arranged an external panel of seven

²⁴ For instance, all the newly gathered provincial data on the areas of different nature target types to be realised were published in the study – mainly as they were very informative to some stakeholders. However, these regional level data were of no ‘judgement-oriented interest’ to the evaluation on a national scale.

members who supervised the research and followed its progress.²⁵ The external panel participated closely, gave important contributions and facilitated data-availability and in-depth knowledge of the subject matter. Members were selected from the most relevant stakeholder groups and showed a broad array of views. Natuurmonumenten chaired the panel.²⁶ At first glance, stakeholder involvement in this evaluation seems to meet the ideal of broad stakeholder involvement of the MCCBA approach. But did it really work in practice?

The EHS evaluation case showed two aspects one 'positive' and one 'negative' concerning the applicability of stakeholder involvement.

The negative aspect involved a representative of the agricultural sector and was especially relevant to the evaluation. This involvement seemed to assure the widest spectrum of views, as farmers and nature conservation organisations were often in direct opposition around the EHS; the agricultural sector was the main 'loser' of the realisation of it. This case clearly showed that in this type of setting the word 'representative' has no formal meaning. When results were published nothing prevented the agricultural sector or individual farmers from disproving or challenging the results. The general point is that involvement of a stakeholder who is bound to lose in an economic sense is a complicated matter, and may have a limited impact on commitment to evaluation results.

The positive aspect that became clear is that stakeholder involvement in the evaluation can really make a difference: it can change the content. During the evaluation the researchers strived to quantify the impact on nature. When the phase was entered of giving quantitative nature scores to different nature target types, the initiator of the research objected strongly. The initiator was sceptical about this type of quantification of nature values, as nature seemed to lose an essential part of its charm when reduced to simple numbers. The stakeholder panel, however, was far more positive but not uncritical. During heated discussions of the panel, the initiator was 'overruled' by the other stakeholders; this occurred in a setting without formal rules as a back up for decision-making. Clearly without the support of the other stakeholders, the research would not have quantified the nature impact to the extent that it did. Although no formal rules were agreed, this type of stakeholder involvement may create its own dynamics and thus have unforeseen impacts on the content of the evaluation.

7.3.4 Stage three: Organise judgement criteria on Triple E impacts

To assess the applicability of the Triple E structure in the MCCBA approach, three aspects require discussion.

First, noteworthy is that only two of the three Triple E impacts are covered in the final criteria structure; impacts on extreme poverty were not covered. They were not purposely neglected; they were simply never discussed because no one thought them to be relevant. Is this a flaw in the applicability of MCCBA? Or would attainment of the ideal have led to a different way of conducting the evaluation? The only possible impact concerning extreme poverty is that loss

²⁵ Also involved was an internal panel of scientists of the University of Groningen: J. Oosterhaven, D. Wiersma and T. Zuidema from the Economic Faculty, and T. Schoot Uiterkamp of the Center for Energy and Environmental Studies (IVEM).

²⁶ G. Beers, *Landbouw Economisch Instituut*, LEI (Institute for Agriculture and Economics), The Hague; G.A.M. van der Grind, *Landbouwschap* (Agricultural Board), The Hague; R. Rabbinge, Agricultural University of Wageningen; Ch.M. van Schaik, *Vereniging Natuurmonumenten* (Nature Monument Society), Graveland (Chairman); P. van der Veer, *Stichting Natuur en Milieu* (Nature and Environment Foundation), Utrecht; J.T. Woldring, Province of Groningen; F.J. van Zadelhoff, *IKC - Natuurbeheer* (Information and Knowledge Centre for Nature Conservation), Wageningen.

of agricultural production may provide further opportunities for farmers outside the Netherlands, some of whom may be poor. Inclusion of the global or higher international level – as discussed above – would then have covered this point, at least partly. However it seems highly unlikely that any non-negligible impacts would have been discovered. Abiding more strictly to the ideal approach and greater focus on extreme poverty impacts would have been perceived as irrelevant and highly artificial in this evaluation situation. Thus we can conclude that the Triple E impact structure is relevant if serious impacts can be perceived or expected. If close scrutiny yields only negligible impacts, there is no need to pay any further attention to them.

The second point clearly shown here is that measurement options will determine the final criterion structure. When looking at the environmental degradation part, and importantly, contrary to common practice in MCA, there was no clear idea about measurement when the research began. The notion of what could be achieved – at minimum – was that more specific information about the types nature being realised could be gathered. The broad aim was to replace the very broad policy aims of 250,000 hectares of nature with more insightful and judgement-oriented information. In the end the success in developing a measurement tool to cover a large number of impacts in one aggregated index determined the criteria structure.

For the landscape impact – the change in experience value – the situation was basically the same. The researchers tried to measure the impact in a meaningful judgement-oriented way without a clear plan on how to do so. Although nature and landscape impacts are often mentioned together, they have quite a different relation to welfare. Landscape impacts are perceived to be much more associated with individual usage, such as with pleasant scenery for housing and recreation. Meanwhile, biodiversity or nature concerns of most people are more collective welfare notions. The option to use monetary measurement of landscape impacts is therefore a serious choice to make. There was little reason in this particular case for finally including the physically measured landscape impacts, because the size of the impacts was regarded as minor. If this had not been the case, then measurement possibilities would have determined whether these impacts were included under CBA impacts or some other criterion. Both of these points clearly indicate that the criteria structure in stage three is preliminary and that size of impacts and their measurement options determine the final structure. This procedure may be seen as somewhat subjective and accidental, but the EHS case illustrates that this procedure may in fact also have important benefits, in that it can act as a means of ‘unvaluing evaluation’ (Alexander, 2001). The need to assign weights – and its concomitant problems – is reduced by considering only criteria that really matter in an empirical sense.

Finally, the criterion of the financial burden to the Dutch government should be discussed briefly; in a standard CBA this type of impact would often be a ‘mere’ transfer and therefore not included. In the EHS case financial burden was structured as a separate criterion apart from CBA elements, although the impact was measured monetarily. Given the importance in policy debate around EHS on the size of the financial burden to government, this has obvious advantages. The practical strength of MCCBA, highlighted in this treatment, is its ability to either include impacts in a CBA or use them as a separate criterion.

7.3.5 Stage four: Quantify impacts physically

Physical quantification of the impacts – as is often the case – required the most time of the researchers, a situation not unique to the MCCBA approach. Therefore, the EHS case does not illustrate much about the applicability of MCCBA as such. Two points that do concern the

applicability of MCCBA are: 1) the role of stakeholders for physical quantification, and 2) the practice of disregarding unimportant criteria.

It was crucial to first know more about the types of nature to be realised in EHS. Stakeholders were very important here and a detailed list of nature types was used. Nature conservation organisation managers and provincial civil servants – local and regional experts – were asked to estimate how much area of which nature type they thought would be realised. A full response to the questionnaire was critical to the success of the research.

Despite considerable uncertainty these stakeholders had to be persuaded that they were in the best position to make these estimates. The initiator of the research helped somewhat here, but several respondents had to be persuaded to participate in this unfamiliar affair – and the last province could only be convinced because the other provinces had already answered.

In the theoretical chapters stakeholders were mostly discussed regarding their role of providing value judgements. However, this particular EHS evaluation case illustrates that stakeholder involvement may also be crucial in the physical quantification stage (Patton, 1997).²⁷

The second point is about disregarding criteria. For impact on the environment, determining the physical impacts proved to be the decisive stage. A negative impact of EHS on the environment might be that the options for ‘placing’ manure from intensive farming systems diminish. Because the possibility for placing manure is essential to the survival of this type of farm, some authors had claimed to translate it into an extra €135 cost per hectare (Oskam, 1994). This loss might be even higher if loss of agribusiness were to be taken into account. Conversely, agricultural production as such causes an environmental burden, and realising EHS might decrease some of it. Diminishing for instance the contamination of water reserves might lead to prevention of extra costs for drinking water production. No clear positive or negative environmental impact could be determined and therefore the criterion was dropped altogether. This procedure of disregarding unimportant criteria is of course not uncommon in the CBA and MCA literature, although in MCA literature the criterion of being important is often applied in advance when choosing the criteria that really make a difference (compare chapter 4). The approach to skip criteria that do not matter in the end is less often proposed.²⁸

7.3.6 Stage five: Aggregate monetary scores consensus based

In stage five the explicit valuation part begins. From section 7.2 it may be clear that the approach of aggregating various impacts in a CBA worked here to an important extent. The CBA part of the analysis incorporated various impacts – from the loss of agricultural production possibilities to the cost of nature development. Some of the complicated aspects of these impacts were reduced by the everyday economic stance of the CBA analysis, such as the gradual replacement of agricultural production factors to other economic uses or the disregard of emotional attachment to agricultural lands to be converted. Without CBA many of these impacts would have been treated as separate criteria in an MCA. With the help of CBA all these aspects – spread over time – were aggregated to one monetary net result: a net present value.

Two questions on the applicability of the approach have to be raised however. The first is how

²⁷ The use of the IKC database on the occurrence of species in different nature target types was also facilitated by a stakeholder (of the IKC) in the external panel.

²⁸ One exception being Alexander (2001)

well are impacts treated that CBA would call ‘transfers’? Second, how practical is the notion ‘consensus based’ aggregation?

To many stakeholders the costs of EHS are synonymous with purchase value of agricultural lands, and to another (probably overlapping) number of stakeholders they are synonymous with budget costs to government. Both types of impacts are normally considered to be transfer payments within CBA (see chapter 3).

Starting with the purchase value of the land, this price is clearly a potential candidate for the estimation of the economic impact on agricultural production. However, it was shown that production on a specific piece of land might move to another one. Taking the price of land as a basis for estimating this impact then becomes more difficult than starting from the value of the agricultural production of the land. As argued above, in the evaluation the choice was made for this last option and may be seen as convenient in a specific setting and not one of principle. However, one could not have chosen both options as this would have led to severe double counting of impacts. Noteworthy too is that although this choice was made on convincing analytical grounds, a choice for estimating impacts through the price of land would have more intuitive appeal to many – non-analytically geared – stakeholders.

The treatment of budget costs has already been discussed; these costs to government are primarily transfers and give a double count with agricultural production impacts. The budget costs are nevertheless an important concern for many stakeholders in the policy debate on EHS and were therefore taken into the evaluation. Having no place in a CBA, they were used as a separate criterion in MCCBA. However, in the valuation stage of the evaluation it is unclear how to judge outcomes on this criterion. In the EHS case it was treated mainly as a precondition for the other two criteria. This treatment may not be *generally feasible* and the evaluation may end up with several separate criteria. The less successful the analysis is in reducing the number of final criteria, the greater the need for some formal MCA approach to give weights and scale the outcomes.

The second issue to be addressed is the extent to which the approach is ‘consensus based’. The discussion thus far makes obvious that the MCCBA approach will not be consensus-based in that stakeholders agree to it the first time they hear about it.

For instance, being able to follow the CBA-based line of reasoning and not for instance, looking at the costs to government, requires some persuasion, or effort to delve a bit deeper. Nevertheless, the position to be persuaded to is not difficult to understand as such: in this case the problem was to examine the long-term impact on production value and compare this value in the ‘with’ and ‘without’ situation. The procedure was easy to grasp by many stakeholders and resulted in a broad consensus that the indicator used was very important. However, the analysis of how different aspects have an impact on this indicator requires specific expertise and logical reasoning that is largely beyond the knowledge and intuition of most stakeholders. Furthermore, as shown above, the CBA-based analysis leads to a reduction of aspects taken in. The completeness of the analysis can therefore always be challenged.

The conclusion seems to be that no consensus can emerge about the completeness of the analysis, but that consensus about relevance of criteria²⁹ and (relative ease of) understanding measurement of impacts is indeed possible.

²⁹ Relevance of criteria can also be understood as importance of criteria, but importance not in the (MCDM) sense of a specific weight given in relation to other criteria, but importance in the (MCDA) sense of (at least) having a minimum importance level.

7.3.7 Stage six: Aggregate non-monetary scores (consensus based)

The results presented in section 7.2 clearly show the practical strengths of using separate criteria for important impacts which are hard to fit within a CBA.

The main issue to clarify here is to know whether the aggregation can be considered consensus based. In the evaluation three indicators were presented for the impact on nature value. An important trade-off situation becomes visible if one examines this set of three. The third indicator, which covers many different and relevant aspects of nature value – occurrence, trend and international importance – is the most complete. The calculations require giving an implicit weight to these three factors and the resulting index is rather complex and synthetic. However, the first indicator – the number of target species – is the easiest to understand for a broad group of stakeholders.

On a small scale, the EHS case reveals the problem hindering many MCA applications: trying to be complete leads to a criteria structure and a way of measurement that is complicated (see chapter 4). Consensus about measurement may then become more difficult.³⁰

In the presentation of the results of the EHS study, the most refined nature indicator was emphasised, although the first and second indicators were also used. Applying the consensus based perspective of MCCBA more strictly would probably have led to greater attention on the first indicator.

Another aspect relevant here is the scale level of the analysis. If MCCBA had been applied more strictly – as discussed above – the analysis would have introduced a high international level. However, with the most refined third nature value indicator the international aspect came in too, and the relation with the national level of the analysis was – in hindsight – not 100% clear. A separate international level might have been preferable, which would have focussed only on the nature scores from an international perspective.

Although MCCBA was fruitfully applied, a more strict application of it would have led to results that could have met with broader consensus among stakeholders.

7.3.8 Stage seven: Interpret trade-offs

How applicable is stage seven of the MCCBA approach if one looks at the EHS evaluation?

To assess this, one may look at the MCCBA final results. Contrary to the standard approach in CBA and MCA, stages five and six combined provide only limited aggregate results, but no final one-answer as such.

In the results presented in section 7.2, a further interpretation using different ratios took place. The ‘cold’ results were put in the perspective of their national totals to have a better idea of the order of magnitude of the impacts (table 7.6). Furthermore, several ‘efficiency’ interpretations looked at benefits per unit of costs (tables 7.8 and 7.9). To many stakeholders these interpretative ratios were the main results of the evaluation.

Even from this one study, it may become clear that this type of interpretation in stage seven of the MCCBA approach has important merits; it may increase the value of the judgement information tremendously to relate the outcomes to relevant benchmarks.

The options for ratio-analysis and interpretation of ratios strongly increase in the case of standardisation of impact measurement. If greater numbers of projects with nature impacts are

³⁰ Related to this issue is the fact that the evaluation focussed merely on the most important criterion, thus making the judgement-information easier to understand but covering less ground.

evaluated using the same indicators, then the order of magnitude of the impacts of these projects becomes far easier to judge. This is certainly one of the major strengths of the CBA part of the analysis, where monetary outcomes are easily compared to other projects.³¹

7.3.9 Stage eight: Perform sensitivity analysis and reconsider project alternatives

This stage in the MCCBA approach aims at sharing increased problem understanding, using sensitivity analysis and providing a reconsideration of project alternatives.

The sensitivity analysis in the EHS study had two elements: 1) a sensitivity analysis of CBA elements and 2) a comparison of results and assumptions of two earlier evaluations. The first mainly highlighted the importance of estimates of management costs of new EHS areas. The second part of the sensitivity analysis showed the importance of differences in insight of the impact on the agricultural sector: which crops disappear and what happens with agricultural labour and capital initially becoming unemployed. The differences refer to a debate about economic – market-related – causality.

The nature score measurement achieved limited attention in the sensitivity analysis. The primary reason for lack of attention in this particular part of the evaluation is the fact that the evaluation used three different scores, which in themselves showed much sensitivity of the different underlying assumptions. Furthermore, because the methodology was new, many choices and assumptions had been tested earlier for their sensitivity in the evaluation report; they were discussed throughout the evaluation.

Apparently, the sensitivity analysis did focus on the most important trade-off relations and sharing increased problem understanding – at least in as much as it concerns the CBA part of the analysis. However, although in the debates following the publication of the evaluation these outcomes were well discussed, the treatment in the main evaluation report may have had too much an appendix character: this is not the MCCBA ideal.

There was no explicit reconsideration of project alternatives. An available and explicated MCCBA approach would certainly have contributed to such an explicit reconsideration. Only one alternative was evaluated in this evaluation; in hindsight the result of a reconsideration might have meant that it seemed useful to evaluate 1) a project alternative that focussed less on combining regular agricultural activity with nature aims, and 2) a project alternative that would focus more on realising high international nature values.

7.3.9 Epilogue

Several years have passed since the evaluation of the EHS and options for ex-post evaluation of results have arisen. A complete evaluation and interpretation will not occur here on what has happened since the evaluation finished. However, a number of points useful to the assessment of the applicability of MCCBA will be touched on.

³¹ Perhaps in the MCCBA approach this ease of comparison is mistakenly applied when CBA aggregation is far from consensus based.

CBA results

The new CBA results of the evaluation deviated strongly from earlier studies, but remain largely undisputed. Some debate has ensued among economists (Strijker and Sijsma, 1996a,b,c) but most issues have been settled. In the years since the evaluation no new debate has arisen about the costs to society of EHS.

Quantifying nature values and cost-effectiveness

The quantification of nature values was fairly new, as was the measurement of cost-effectiveness in relation to nature goals; these elements are closely associated with the MCCBA approach, and the evaluation has contributed to and partly provoked debate around these elements.

The years following the evaluation have shown a steadily increasing interest in both science and policy. In nature policy more attention was given to the nature results³² – also within and between nature conservation organisations. The nature measurements have been (further) developed (Van Wenum, 2002).

The cost-effectiveness of different instruments remains a major point of discussion. The relative performance of nature-friendly farming versus nature conservation was largely confirmed in later analyses (Van Wenum et al., 1999). The role of specific types of nature (meadow birds) provided by agriculture and more recently the landscape ‘provision’ of the agricultural sector, received increased attention in the discussion. The nature scores developed in the EHS evaluation played a role in the first development because it could quantify scores on the meadow birds mentioned; but with regard to landscape provision quantification of impacts is still difficult, and the evaluation seems not to have played a role.³³

Looking at the lack of debate around the CBA results as well as the input of the evaluation on the nature measurement and cost-effectiveness issue, one is tempted to find support for the argument that MCCBA can possibly ‘clear as much ground as possible’.

Government budget

In the years after the evaluation, agricultural land prices have risen – even higher than assumed in the evaluation. An explanation of the reasons for this development is not within the scope of this thesis, but suffice to say that the main change in the evaluation outcome would not be changes in estimated loss of agricultural production, but rather would focus more on the required amount of government budget.³⁴ Generally, it is noteworthy that most of the extra budget needed was provided and thus seems to acknowledge the correctness of the assessment that the financing was feasible. The relevant point here is not this fact as such; it is instead that the rather vague criterion of ‘feasibility’ can work (it would be difficult to specify the maximum amount of feasible financing).

³² Regularly overestimating the steering ability of this type of ‘production’.

³³ The hope by some to come to payments to farmers for landscape services – which, incidentally, might have warranted a CVM approach in the evaluation – was not realised.

³⁴ Another way to say the same thing is that these high prices are more changes in the size of transfers and less a ‘real’ economic change from an agricultural perspective.

CHAPTER 8: EVALUATION OF THE GERMAN EMSSPERRWERK: THE VALUE OF MORE SCALE LEVELS

8.1 Introduction

The evaluation in this chapter was carried out in 1998, a couple of months before the German government decided to build the Emssperrwerk.¹ The Emssperrwerk had two official purposes, first to act as a (movable) dam in case of extremely high water, thus preventing water from the sea threatening the hinterland in the EmsDollard region. Second, to facilitate large cruise ships built at the Meyer shipyard in Papenburg (approximately 30 kilometers upstream the river Ems) in order to reach the deep waters of the North Sea.² There has been heated debate about the Emssperrwerk build. Many thought that the only real reason for building the Emssperrwerk was for the purposes of the Meyer shipyard. This evaluation follows the official stance of the two major aims.

The Emssperrwerk was planned in the river Ems between the villages Gandersum and Nendorp. The total length of the Emssperrwerk would, according to the plans, be 477 meters. There would be nine large concrete pillars, with movable steel doors 50-meters wide in between them. There would be two passages for ships, one for inland vessels and one for sea-going ships, 50 and 60 meters wide respectively. Whereas most of the steel doors were planned to hang above the water when in rest, the one of the sea-ship segment was planned to lie on the river bed. The sea-ship passage also had greater depth: minus 9,00 meters NN. Part of the plan was the option to build a lock, which would give the possibility of *permanent* ship passage, as ships might pass even when the Emssperrwerk was closed.

The aim of the evaluation was to assess costs and benefits of building the Emssperrwerk and possible alternative actions for it. One alternative for the Emssperrwerk, a high water dam, was to reinforce the dikes of the relevant area. Alternatives for the Emssperrwerk such as a cruise ship facility, were to move the shipyard closer to deep water – combined with the renaturalisation of the Ems river. Two options were reasonable with regard to the relocation: the German Emden or the Dutch Eemshaven. A final alternative – but of a different order to many stakeholders – was the alternative of renaturalisation combined with closing the shipyard. Future orders would then be taken over by shipyards elsewhere in Europe.

The evaluation again took the form of an MCCBA *avant la lettre*. The impact on nature values was an independent criterion. The social costs and benefits possible to express in monetary terms were classified in a limited CBA. Here too, as in chapter 7, special attention was paid to quantifying the nature values that the Emssperrwerk would affect. More dominant in this evaluation was the explicit use of different spatial levels: Niedersachsen, Germany and the global level. The first two levels were chosen because decision-making and funding existed at these levels: it was primarily the decision of Niedersachsen, but not without the consent of the German central government – that had to pay a substantial amount of the costs.

¹ Building the Emssperrwerk was planned for June 1997. The major sources used in this chapter for the plans of the Emssperrwerk were the official project plans of the Bezirksregierung Weser-Ems.

² The huge Meyer shipyard in Papenburg has specialised in building cruise ships. These boats have become ever larger and deeper, outgrowing the depth of the river Ems. With the Emssperrwerk the water level in the river Ems can temporarily be raised when once or twice a year a cruise ship has to sail to sea.

In this chapter the evaluation will be discussed first, by presenting a summary of the original results (section 8.2) and second, by assessing from these results the applicability of the ideal MCCBA approach in a sustainability context (section 8.3).

8.2 Original results³

8.2.1 Brief summary of results

This evaluation attained results on three spatial levels: The province/state of Niedersachsen, the national of Germany, and at the global level.

German level results will be discussed first, since it seems to be the best compromise between defining a correct impact population and defining an evaluation level that stays close to the decision making level. At the level of *Germany*, calculations showed that building the Emssperrwerk would give a net negative balance of the limited CBA. The benefits of extra growth of Meyer shipyard did not outweigh the extra costs of building and maintaining the Emssperrwerk. Even according to the cheapest estimate, there was still a negative net cost of €98 million (net present value in 1998 of discounted flows – at a 4% real discount rate – over a 50-year period). The change in nature value caused by the Emssperrwerk was estimated as negligible.

The alternative of renaturalisation gave a positive net balance of benefits and costs, and also showed a gain in nature value. At the German national level a relocation of the Meyer shipyard to Emden was the best option. But even the relocation to the Dutch Eemshaven showed a positive balance to Germany. The alternative of renaturalisation, combined with the closing of the Meyer shipyard and its production taking place elsewhere in Europe, indicated large losses to Germany.

At the level of *Niedersachsen* the picture was very different. Here the cheapest estimates of the Emssperrwerk had a positive balance; the reason being that most of the economic benefits accrued to Niedersachsen, while it had to pay only half the costs of building the Emssperrwerk.

At the *global* level the outcomes are yet again quite different; it was assumed that the demand for cruise ships will always be realised, either by being built by Meyer or by shipyards elsewhere in Europe. The extra growth in the shipbuilding industry because of the good market prospects for cruise tourism will therefore in every case be realised somewhere in the world (read: Europe). All alternatives give net costs on a global level – compared to the nul-alternative. From a global perspective the Emssperrwerk alternative has low attractiveness: substantial costs have to be made for benefits that can be realised without much effort elsewhere. At the global level all renaturalisation alternatives gave costs too, although these costs were always lower than the Emssperrwerk alternatives. The costs were lowest when Meyer's production stops and is taken over by shipyards elsewhere in Europe.

³ The original study: Bettels and Sijtsma, 1998.

8.2.2 Results and method in more detail

8.2.2.1 Alternatives

In this section the alternatives are fully described. The safety of the coastal areas and the navigability of the river Ems for large cruise ships were the official reasons for building the Emssperrwerk. In this evaluation the alternatives were specified following three different policy aims:

- Extra coastal safety⁴
- Extra coastal safety and navigability of the river for cruise ships
- Extra ecological quality in the river Ems by renaturalisation

The third policy aim did not arise from the project plans for the Emssperrwerk as such but came from existing policy for this Wadden area – which tries to redevelop or maintain natural estuaries in the Wadden area. According to the nature oriented lobby organisations, the Emssperrwerk in particular might intervene with this third policy objective.

Under these three headings the following alternatives were specified, evaluated and compared to the nul-alternative, that is, the ‘business as usual’ situation:

Two alternatives for extra safety only:

- 1a - reinforcing the dikes
- 1b - constructing a flood barrier

Four alternatives⁵ for extra safety and navigability by:

- 2a/b - building an Emssperrwerk without a lock (two cost estimates)
- 2c/d - building an Emssperrwerk with a lock (two cost estimates)

Three alternatives for extra safety and extra ecological quality by:

- 3a - renaturalisation and moving the shipyard to Emden
- 3b - renaturalisation and moving the shipyard to Eemshaven
- 3c - renaturalisation and closing the Meyer shipyard (and letting other shipyards outside Germany, elsewhere in Europe build cruise ships)

8.2.2.2 Nature impacts

In order to calculate the nature impact, use was made of the same database on nature types and species used in chapter 7 (based on Bal et al., 1995). One would ideally prefer a database especially designed for Germany or Niedersachsen. However, this was no option within the scope of this evaluation. As a second best option it was analysed whether the Dutch database could be used. This analysis showed that the ecological situation in Niedersachsen, and more generally the situation in the low-land coastal area of Germany proved suitable for using the Dutch data; the same species and nature types occur, and roughly the same species and ecosystems were considered as highly valuable.

⁴ The business as usual situation – the nul alternative – could also be judged as safe, as the dikes were judged safe by the relevant authorities (die Bezirksregierung Weser-Ems, in das ‘Generalplan für Küstenschutz’ von 1997) prior to the Emssperrwerk plan.

⁵ In retrospect, the treatment of different cost estimates as minimum and maximum estimates would have been better than treatment as different alternatives.

The area where direct impacts would be felt by the different alternatives was approximately 14000 hectares, which could be subdivided into two main areas: the downstream area with the tidal water area, and the upstream area of the river and its natural surroundings. Data on the types of nature in the area were translated into the nature target types of the Dutch database, which enabled the calculation of the available nature value in the nul-alternative. Table 8-1 shows these results.

Table 8-1 Calculation of nature value in the nul-alternative according to nature score method

Area	Number of hectares	Average nature value (per hectare) indicator 1***	Average nature value (per hectare) indicator 2***	Average nature value (per hectare) indicator 3***
Tidal water area*	9035	84	86	4808
River area**	4781	59	78	4805
Total	13816	75	83	4807

* subdivided into six smaller nature types

** subdivided into five smaller nature types

***See table 8-2 for a short description; see chapter 7 for a more elaborate description.

For all the various alternatives the impact on the nature values was analysed by accounting for the possible change in river area and tidal water area, and more specifically, by analysing changes in the subdivision of nature type areas within these two aggregate categories on the basis of substantial foreseen impacts on species living within them. For the renaturalisation alternatives this assessment was made on the basis of reference nature types near the Elbe.

The analysis showed that a direct nature impact of the Emssperrwerk and dike alternatives was absent. Although some temporary impacts may occur during the realisation phase, no permanent impacts were foreseen.⁶ Therefore, the nature scores in these alternatives show no change (compare tables 8.1 and 8.2). Substantial impacts on nature values were relevant only in the renaturalisation alternatives. As can be seen in table 8-2, the change in nature value in these alternatives was estimated to increase between 20 to 25%, depending on the indicator used. From the first indicator we see that the number of (target) species occurring on a hectare would, on average, increase from 75 to 94.

⁶ For the Emssperrwerk alternative this result was established only after a substantial research effort to contact different experts about the maximum time species living in the water can survive in different salt conditions. The closure of the Emssperrwerk proved short enough to guarantee survival.

Table 8-2 The nature scores of the different alternatives

		Nature value (average per hectare)		
		1*	2*	3*
Nul-alternative	abs	75	83	4807
Alternatives - Extra safety	abs	75	83	4807
Alternatives - Emssperrwerk	abs	75	83	4807
Alternatives - Renaturalisation	abs	94	100	5935
	relative change (in %)	25%	20%	23%

* 1 = number of target species

2 = number of target species with optimal occurrence weighed double

3 = like 2, but species scored 0-100, depending on rarity trend and international importance

The calculation above was made for the project area. However, because the evaluation was performed at the level of Niedersachsen and Germany, it required calculation and interpretation of changes in nature values for these two scale levels (see appendix 2 to this chapter). Without detailed data for Germany, the calculation at this level was only a very rough approximation. At the global level calculation of totals on the basis of essentially Dutch data was not possible.

8.2.2.3 Economic impacts

Introduction

Costs and benefits that could be monetarized were specified for a 50-year period and discounted with a 4% (real) discount rate. Net present values were calculated for the year 1998. First the separate cost items and their estimation will be discussed below. The benefit items that can be monetarized are treated afterwards.

Costs of dike improvement

The costs of the higher dikes were estimated as follows. If the Emssperrwerk would not be built then the environmental impact study showed that the dikes would have to be raised/strengthened over a distance of 110 kilometres. The estimate of costs was provided at €150 million. Because in this case it would merely be a measure to increase coastal safety, the cost would be divided between Niedersachsen and the German government on a 30:70 basis (that is, €45 mln : €105 mln). No extra costs are foreseen for maintenance, as in the nul-alternative the dikes also have to be maintained too (see table 8-3).

Table 8-3 Costs of dike improvement in Euro mln

	Investment costs	Maintenance costs	Total
Nominal total	150	none (extra)	150
Discounted total	128	none (extra)	128

Costs of a flood barrier

Another option for increasing coastal safety is to build a flood barrier. For an independent

estimation of the costs of such a measure,⁷ an analysis was made of the costs of the Dutch – single purpose – flood barrier in the Nieuwe Waterweg near Rotterdam, which was completed in 1997. The costs of this option were estimated at €456 mln.⁸ As for the maintenance costs certain considerations were made; in the Dutch type of flood barrier, the large doors are normally on the shore and maintenance can be performed on land, whereas the Emssperrwerk doors are in and above the river. The assumption was therefore that for a Dutch type of flood barrier maintenance costs were lower (as a percentage of the total building costs) than those of the Emssperrwerk. Maintenance costs were therefore put at 3% per year (€14 mln per year). Again, Niedersachsen and Germany divide costs at a 30:70 ratio; Niedersachsen €137 mln investments plus 4.2 mln annually for maintenance and Germany €319 mln investments plus €9.8 mln annually for maintenance.⁹

Table 8-4 Costs of a flood barrier in Euro mln

Flood barrier	Investment costs	Maintenance costs	Total
Nominal	456	656	1112
Discounted	456	278	634

Costs of the Emssperrwerk

The costs of the Emssperrwerk comprise the costs of building the Sperrwerk and the costs of maintenance.

There were two official cost-estimates. A study for the Niedersachsen Ministry of Economic Affairs in 1992 estimated the cost of the Emssperrwerk (without a lock) at €242 mln¹⁰ and the yearly maintenance at €20 mln (that is 8.0%). Corrected for inflation this would be 272 mln DM for construction and €22 mln for maintenance at the time of the evaluation.

In 1997 the Bezirksregierung Weser-Ems estimated that the costs for the Sperrwerk would be substantially lower, at €176 mln for construction, and yearly maintenance at €6 mln (3,6%). This latter estimate seemed rather low. The costs would be split in half: 50% borne by Niedersachsen and 50% by the German central government.¹¹

With or without a lock

In the evaluation it was initially assumed that the Sperrwerk would be built with a lock to facilitate regular ship movements during closure of the Sperrwerk for the Meyer shipyard. The extra costs of a lock had been estimated earlier by Planco (a German consultancy organisation; Planco, 1997); they specified three lock alternatives. In the current evaluation the middle estimate in terms of costs was used (inclusive of investment and maintenance costs), which resulted in a sum of €60 million extra for a lock.

During the evaluation doubts about the feasibility of realising the lock increased; Therefore,

⁷ That is, without depending on the calculations for the (double function) Emssperrwerk.

⁸ These costs are relatively high compared to the Emssperrwerk; this could be caused by the more complex design of the Dutch barrier or by a relatively low estimate of the costs of the Emssperrwerk.

⁹ Maintenance costs are the added cost of all years until the end of the 50-year evaluation period (discounted at 4% per year), but maintenance starts only after finishing realisation (i.e., after three years).

¹⁰ In the study – undertaken before the Euro replaced the Deutsche Mark – an unofficial but largely correct exchange rate of 2 Deutsche Mark for 1 Euro was adopted.

¹¹ Compare Ostfrisische Zeitung, 04.07.1997.

the costs without a lock were also specified. Furthermore, to do justice to the fact of two official cost-estimates for the Sperrwerk, both were included in the evaluation. To recognise this aspect more easily, these estimates are treated as alternatives or options, so there are four options for the Emssperrwerk:

- Option 2a: Weser-Ems-'97 cost-estimate with a lock
- Option 2b: Weser-Ems-'97 cost-estimate without a lock
- Option 2c: Niedersachsen-EZ-'92 cost-estimate with a lock
- Option 2d: Niedersachsen-EZ-'92 cost-estimate without a lock

For all options the investment costs were spread over the three year build period, maintenance costs were spread over a 50-year period. Both cost item flows were discounted to a net present value. Table 8-5 shows the costs of all four options.

Table 8-5 Overview of costs of the Emssperrwerk to Germany in Euro mln (net present value)

Emssperrwerk	Investment costs	Maintenance costs	Total
Option 2a (W.-Ems + lock)			
nominal total	237	400	637
discounted total	227	166	393
Option 2b (W.-Ems no lock)			
nominal total	176	299	475
discounted total	169	123	293
Option 2c (EZ NS + lock)			
nominal total	332	1.248	1.580
discounted total	319	517	836
Option 2d (EZ NS no lock)			
nominal total	272	1.023	1.295
discounted total	261	423	685

Costs of renaturalisation

The renaturalisation alternatives have several cost items relating to: 'building' new nature, the inconvenience to shipyard employees of a longer travelling distance, infrastructure costs, and the shifts in production value.

Renaturalisation costs – Renaturalisation of the river would give nature development costs, while a more nature-friendly management of the river shores and river forelands would give extra costs to farmers. Dutch data was used for the estimation of these costs. Overall, this resulted in €17 million in costs (discounted total: 2 for development, 15 for extra management).

Moving the shipyard – These costs concern the rebuilding of the main construction hall and offices and moving the machinery from Papenburg to Emden or the Eemshaven; the costs were estimated at €100 million.

'Inconvenience costs' – For the Emden and Eemshaven renaturalisation alternatives a once-only compensation to employees for the inconvenience of the movement and the changed travel scheme was included. For the Emden alternative this compensation was set at €7500 per employee. For Eemshaven it was set at €15000.

Adjustment of infrastructure – Before a shipyard of this size can move some adjustments in infrastructure would be needed. For Emden the estimated costs are higher (€30 mln) than at Eemshaven (€5 mln)

Shifts in production and value added - Depending on the renaturalisation alternatives (Emden, Eemshaven or Elsewhere Europe), shifts in the location of value added would occur. It was first assumed that 70% of the value added could be ascribed to labour, and 30% to capital. In the Emden variant it was assumed that all remuneration for labour and capital, both direct and indirect, would remain within Niedersachsen. In the Eemshaven alternative 15% of the direct employment and 30% of the indirect employment would become Dutch; all the capital rewards would remain within Niedersachsen. In the Elsewhere Europe alternative all rewards for labour and capital would be received outside Germany and Niedersachsen.

It was furthermore assumed that, because the shipyard would be moved, a fixed percentage of 5% of the direct value added would be lost and the region losing the shipyard would suffer an equally large loss due to loss of the complex providing vital economies of scale to some activities. For both items the argument that these costs would fade over a long time period seemed to be a realistic notion – as for instance, unemployed former employees leave the potential workforce eventually. Both costs items were therefore assumed to decrease to zero over a period of 25 years.

Benefits of the different alternatives

The different alternatives have different benefits, which are discussed below.

These benefits partly overlap:

The benefits of the higher dikes are:

- Increased safety

The benefits of the Emssperrwerk alternatives are:

- Increased safety
- Lower costs for dredging the river Ems
- Secured growth possibilities for the Meyer shipyard at its current location

The benefits of the renaturalisation alternatives are:

- Increased safety
- Lower costs for dredging the river Ems
- Secured growth possibilities for the Meyer shipyard at another location (or another shipyard in Elsewhere in Europe)

The benefit of the increased safety level

No information could be found on the increased safety level that could specify the extent to which extra safety would be realised. Based on studies initiated by the project team 'Emssperrwerk', it became clear that in the event of a storm tide, the water level at the river would be substantially lower, while the water level at the (German) sea-side of the Ems-Dollard estuary might increase by some 7-10 centimeters. In the Dutch part of the estuary another study showed increased sea levels at approximately 2.5 centimeters.

The safety level gain in the renaturalisation alternatives comes from the increased storage capacity due to building larger flood areas, but the exact size of the increase was not calculated.

On the basis of this information it seemed predictable that the safety level in the region would increase with all alternatives, although it is not known by how much; trying to estimate this impact within this evaluation more precisely was not possible. Extra safety was therefore included in the MCCBA as a pro-memory post.

The benefit of lower costs for dredging

The second benefit consists of lowered dredging costs. In the years before the evaluation about one million m³ of river bed was dredged (Herr et al., 1997 p. B4/5). The costs associated with this dredging were estimated at €2.5 /m³/year, which totals €2,5 million yearly. No dumping

costs were taken into account here, thus the real costs would have been higher. It was reasonable to assume that because of the Emssperrwerk, dredging activity could decrease. Lowering the activity to the level prior to 1986 (the year Meyer started building cruise ships) did not seem realistic, however. Exactly how much dredging activity would remain necessary was unclear: the assumption was that the activities could be halved compared to the nul-alternative due to Emssperrwerk. This approach implied a yearly cost saving of €1.25 mln. In the renaturalisation alternatives no dredging costs are made.

The benefit of secured growth possibilities for the Meyer shipyard

The Emssperrwerk gives the Meyer shipyard the opportunity to follow the trend of building larger and larger cruise ships. Instead of 7.30 meters (the depth of the Ems at the time of the evaluation) the Meyer can then use the depth of 8.50 meters. In the planning of the Emssperrwerk 8,5 meters is the maximum depth, probably because of the large tunnel crossing the Ems (BR Weser-Ems, p. 7, 1997). In the nul-alternative a stand-still was assumed concerning the depth of the ships, which still allowed Meyer to follow a ‘business-as-usual’ development.¹²

It was uncertain how long the extra depth provided by the Emssperrwerk would sufficiently help the Meyer shipyard. In this evaluation it was assumed that eventually the growth of the Meyer shipyard would be hampered again by the limited depth of the river Ems; it is difficult to estimate exactly how long Meyer can profit. In the evaluation it was assumed that after the realisation of the Emssperrwerk in 2002, the Papenburg shipyard can follow the trend in depth growth of cruise ships for 10 extra years. In order to assess the sensitivity of this assumption, results were also calculated using a 20-year period of extra growth.

The Meyer shipyard does not give information to third parties about its turnover. Because no direct information was available, the evaluation worked with a rough approximation, building on the number of workers at the shipyard (2000 employees). For the estimate of the impact of the Sperrwerk, firstly the value added of the Meyer shipyard was important; this value added was estimated using the gross value added per worker in industry in Niedersachsen, which was €45950 in 1996. It was further assumed that if Meyer no longer experienced limitations concerning the depth of its ships, its value added per worker will show extra growth. This growth would then also hold for the related value added of the suppliers: the indirect employment related to the Meyer shipyard.¹³

How much extra growth would Meyer experience? The market for cruise tourism is generally seen as a growth market. However, this does not mean it is free from changes in the economic climate: luxury products are often quite sensitive to business cycles. While the growth of the Meyer shipyard may rise and fall, the estimate assumes a constant growth path – as long as it is not hindered by the depth limitations.

Because of the favourable market perspectives, the activity within the Meyer shipyard would generate extra growth in value added which would not otherwise occur. For a starting point when estimating the extra growth of Meyer, the average real production growth in the German

¹² One might challenge the realism of this combination of stand-still in depth and continuing to reach the current turnover level; it might require extra marketing effort or even development of extra quality in other aspects of the cruise boat design. These requirements were not calculated.

¹³ A multiplier for the shipyard activities was calculated (of 2,3 implying 1,3 of indirect employment for every 1,0 of direct employment) using data from the Dutch regional input-output model concerning the Groningen shipyard industry (Eding et al., 1995). The Meyer shipyard has a direct employment of 2000, the indirect employment was thus estimated at 2600.

economy in the last decade was taken. This real growth rate in the years 1970-1996 was 2.3% per year; it was argued that the employment and capital active in the Meyer shipyard have other possibilities – they could be used in other economic activities (showing average growth). The estimate in this research was that Meyer would generate an extra growth of 0.5% above the long-term growth of the German economy: thus a 2.8% growth in value added per year of all employment connected with Meyer. The *extra* growth is the economic benefit to society of the possibility for unhampered growth of the Meyer shipyard. Furthermore, it was argued that this extra growth will continue for 10 years; after that, the depth of the Ems will once again be inadequate. Meyer then does not generate extra growth in comparison to the alternative uses. The extra growth of 0.5 % is seen as a maximum estimate, the 10 years as a minimum. Table 8-6 also shows the results of a 20-year higher growth period. The indirect employment of the Meyer shipyard was included in the estimates.

Table 8-6 gives an overview of the benefits at the German level and the Niedersachsen alike¹⁴ of the Emssperrwerk and renaturalisation alternatives. The benefits were the same for all four Emssperrwerkoptions (2abcd), as the options differed only with respect to the lock and the costs involved.

Table 8-6 Benefits of the Emssperrwerk alternatives (2abcd) and the renaturalisation alternatives (3abc) in mln Euro (German level; discounted totals)

	Increased safety level	Lower dredging costs	Extra growth Meyer 10 years	Extra growth Meyer 20 years	Extra growth (50 years)
Emssperrwerk (2abcd)	n.a.	24	170	278	
Renaturalisation					
3a	n.a.	*			450
3b		*			353
3c		*			0

* No dredging costs are made with these alternatives, but some extra costs for managing the flooding areas and summer dikes etc. will be needed. The gain in dredging costs is assumed to equal these extra costs.

8.2.3 Overall evaluation of the results

8.2.3.1 Results at the German level

Table 8-7 shows the final results at the level of Germany. According to the calculations, the Sperrwerk gives a net negative balance of the CBA. The benefits of 10 years of extra growth of Meyer do not outweigh construction costs and maintenance of the Sperrwerk. Even the cheapest estimate for costs gives a negative balance of €98 mln. The change in the nature value due to the construction of the Sperrwerk as such is probably negligible. Renaturalisation of the river and moving the shipyard to Emden or Eemshaven will give a positive balance, and also have a positive impact on the nature values. At the German level the Emden option is the

¹⁴ It was assumed that all suppliers come from Niedersachsen, which is obviously not the case. As no detailed decomposition was available warranting estimation, and as it was thought that at least more than half the supplies would come from Niedersachsen, this 'second best' choice was deemed defensible.

most attractive; it gives higher nature values and the balance of the cost-benefit analysis is most advantageous. The reason for this result is that the expenses of moving the shipyard, compensation of employees, etc., are lower than expenses for construction and maintenance of the Sperrwerk. The benefits are larger than with the construction of the Sperrwerk because Meyer can generate high growth for a much longer period of time (until the end of the 50-year period of analysis). These extra benefits fall to (inhabitants of) Germany. In the renaturalisation option Eemshaven, benefits fall partly to the Netherlands; however, relocation to the Eemshaven still gives a positive balance to Germany. In the renaturalisation option Elsewhere Europe (Meyer ceases to exist) the costs to Germany are large. The complete production of the large shipyard with the accompanying suppliers disappears from Germany; however, the expenses of this ecological variant are still lower than the expenses for the Sperrwerk according to cost estimates of Niedersachsen in 1993, which underlines the importance of a correct cost estimate for construction and maintenance of the Sperrwerk. As seen above, the official estimates of the Bezirksregierung Weser-Ems were rather low.¹⁵

Table 8-7 Summary of results at the German level ('+' means net benefits, '-' means net costs)

Germany		
	Benefits minus costs (Euro mln, net present value)	Change nature value (% change of total German nature value)
<i>Extra Safety</i>		
Dike	-128	0%
Flood barrier	-634	0%
<i>Emssperrwerk</i>		
Option 1 (W.-Ems plus lock)	-197	0%
Option 2 (W.-Ems without lock)	-98	0%
Option 3 (EZ NS plus lock)	-640	0%
Option 4 (EZ NS without lock)	-489	0%
<i>Renaturalisation</i>		
Eco 1 Emden	+288	+ 0.009 %
Eco 2 Eemshaven	+69	+ 0.009 %
Eco 3 Elsewhere Europe	-395	+ 0.009 %

A sensitivity analysis was performed due to remaining major empirical uncertainties about the results at the German level. Several 'what if?' questions were asked. The analysis showed that the most crucial question was: 'What if the Meyer shipyard can have 20 years of extra growth instead of 10?' A prolonged growth period gave results for the Emssperrwerk greater than €100 mln higher for all Emssperrwerk options, leading to a positive balance of the cheapest option (option 2) at the level of Germany.

¹⁵ An alternative estimate of the maintenance costs (e.g., 7,2% instead of 3,6%) would yield an extra amount of costs to the sum of €166 mln for the cheapest Emssperrwerk alternative.

8.2.3.2 Results at the level of Niedersachsen

When the results of the evaluation are examined only at the German level it is difficult to understand why constructing the Emssperrwerk would be considered at all. However, at the level of Niedersachsen the situation is very different.¹⁶ After all, a large part of the expenses is not covered by Niedersachsen; Germany pays 70% of the costs for the Extra safety alternative (flood barrier or dikes) and 50% of the Emssperrwerk; it was assumed that there is no German contribution to the renaturalisation expenses for nature construction and relocation.

According to the cheapest cost estimates, the Sperrwerk without a lock gives a positive balance at the level of Niedersachsen (see table 8-8). The variant with a lock (cheap estimate) gives a negligible negative balance. Even at the level of Niedersachsen, however, the renaturalisation alternatives Emden (and to a lesser extent Eemshaven) might be preferred above construction of the Emssperrwerk. The difference between the Eemshaven variant and the Sperrwerk without a lock is, however, very small.

Table 8-8 Summary of results at the level of Niedersachsen ('+' means net benefits, '-' means net costs)

Niedersachsen		
	Balance of benefits and costs (Euro mln, net present value)	Change in nature value (%)
<i>Extra Safety</i>		
Dike	-38	0%
Flood barrier	-190	0%
<i>Sperrwerk</i>		
Option 1 (W.-Ems + lock)	-1	0%
Option 2 (W.-Ems without lock)	+48	0%
Option 3 (EZ NS + lock)	-223	0%
Option 4 (EZ NS without lock)	-147	0%
<i>Renaturalisation</i>		
Eco 1 Emden	+288	+ 0.08 %
Eco 2 Eemshaven	+69	+ 0.08 %
Eco 3 Elsewhere Europe	-395	+ 0.08 %

8.2.3.3 Results at the global level

At the global level the results are again very different. As noted above, it was assumed that the cruise ships will be built in all cases: either by Meyer (wherever) or by competitors elsewhere in Europe. The extra growth in shipbuilding – due to good market prospects for cruise tourism – is thus realised in all cases.

The costs and benefits of the relocation or shift of the shipbuilding activity at the global level are almost completely balanced, because it was assumed that the gain of one region is largely

¹⁶ Compare Oosterhaven (1982) for another CBA study in which spatial scale levels were crucial to the outcomes. In this study several non-monetized criteria were taken in, with plus or minus scores.

the loss of the others. The attraction of the different variants is then determined by the other impacts: costs of the Emssperrwerk, expenses of renaturalisation of the river, costs of moving the shipyard, compensation of employees, and other infrastructure adaptations in the case of the renaturalisation variants. All alternatives give net costs at the global level compared to the nul-alternative¹⁷ (see table 8-9). From a global perspective the Emssperrwerk is not very attractive as large expenses are required in order to receive profits that can ‘effortlessly’ be realised elsewhere.

All renaturalisation variants give expenses; in all cases these costs are lower than those of the Sperrwerk. Obviously the most striking outcome is that the disappearance of Meyer from Germany has the lowest negative balance of all variants. The fact that there are some costs is because part of the production factors – even after a number of years – cannot find an alternative use. There will be extra costs with competing shipyards when taking over orders that would otherwise have been built by Meyer. It seems reasonable that these expenses will always be lower than the expenses that must be made in the Eemshaven where there is no shipyard of this size active.

The renaturalisation variants do give gain in nature value; it can, however, not be calculated as a percentage of the total global nature value, as no such figure is available. On a global level both benefits and costs seem to be small. Whether on a global level the (small) nature benefits outweigh the small amount of costs is difficult to say, as on the global level the nul-alternative of no growth for Meyer is also a realistic alternative.

Table 8-9 Summary of results at the global level

Global		
	Balance of costs and benefits (Euro mln, net present value)	Change in nature value (% of total at scale level)
<i>Extra Safety</i>		
Dike	-128	0%
Flood barrier	-634	0%
<i>Sperrwerk</i>		
Option 1 (W.-Ems + lock)	-393	0%
Option 2 (W.-Ems without lock)	-293	0%
Option 3 (EZ NS + lock)	-836	0%
Option 4 (EZ NS without lock)	-685	0%
<i>Renaturalisation</i>		
Eco 1 Emden	-207	n.a. (but positive)
Eco 2 Eemshaven	-192	n.a. (but positive)
Eco 3 Elsewhere Europe	-45	n.a. (but positive)

8.2.3.4 Results at all levels

In the original study an overview table, table 8-10, was presented which shows how the two

¹⁷ The nul-alternative: stand-still of Meyer, extra growth elsewhere, no effects on nature values.

most important alternatives – that at the least met both policy targets of extra safety and extra growth for Meyer – scored on different spatial levels. The scores were given in qualitative judgements indicating whether an alternative is ‘desirable’ or ‘to be considered’ or ‘undesirable’. The following assessment rules have been followed. If an alternative or option yields nothing ecologically, and the economic cost-benefit analysis gives expenses only, then it is undesirable. If economic benefits or costs and ecological benefits or costs have a mixed sign, then the option can be considered. Benefits for both the economy and nature are desirable.

Table 8-10 Comparison of two alternatives (Emssperrwerk and renaturalisation) at three spatial scale levels

Judgement of the alternatives per spatial level			
	Niedersachsen	Germany	Global level
<i>Emssperrwerk</i>			
Option 1 (W.-Ems + lock)	±	-	-
Option 2 (W.-Ems no lock)	+	-	-
Option 3 (EZ NS + lock)	-	-	-
Option 4 (EZ NS no lock)	-	-	-
<i>Renaturalisation</i>			
Eco 1 Emden	+	+	±
Eco 2 Eemshaven	+	+	±
Eco 3 Elsewhere Europe	±	±	±
<i>Economic impact</i>	<i>Ecological impact</i>	<i>Classification</i>	<i>Symbol</i>
<i>gain</i>	<i>gain</i>	<i>desirable</i>	<i>+</i>
<i>loss/no change</i>	<i>gain</i>	<i>to be considered</i>	<i>±</i>
<i>gain</i>	<i>loss/no change</i>	<i>to be considered</i>	<i>±</i>
<i>loss/no change</i>	<i>loss/no change</i>	<i>undesirable</i>	<i>-</i>

The assessment ‘results’ of this table are of course merely indicative, as there is no convincing argument for the implicit 50% weight of the two final criteria. An overview of the final results is furthermore hampered by a lack of clarity about the (measurable) safety impacts of the different variants. The main benefit of the table is that it gives a rough summary of the positive and negative overall impacts at different spatial levels. Section 8.3 will return to the merits of table 8-10.

8.3 The Emssperrwerk in the light of the MCCBA approach

8.3.1 Introduction

In the discussion below, the applicability of MCCBA will be assessed by proceeding through its eight stages. The main focus as before will be on the practical applicability of the approach and it will especially be noted if deviations from the ideal seem to occur for more structural reasons.

8.3.2 Stage one: Identify function, project alternatives and scale of the evaluation

8.3.2.1 Function of the research

The evaluation was conducted on the request of the Waddenvereniging, an association active as a lobby organisation for preserving the Waddensea – the largest nature reserve in the Netherlands. Little funding was available.¹⁸ The Waddenvereniging was concerned about the building of the Emssperrwerk and thought that an evaluation study might yield insights into the costs and benefits, and that it might even help by raising arguments to halt the building of the Emssperrwerk.¹⁹ The research team's function in the evaluation was to test the applicability of the (then) more rudimentary MCCBA methodology.

8.3.2.2 Project alternatives

The choice of the alternatives was crucial to the entire study. Naturally any evaluator would evaluate building the Emssperrwerk against the nul-alternative of not building it. Due to the knowledge of the Waddenvereniging about international Waddensea policy and their expectations about possible future developments of the Waddensea, the alternative of renaturalisation of the river came to the fore; renaturalisation of the river required relocation of the shipyard.

8.3.2.3 Scale of the evaluation

The scale of the evaluation posed an interesting decision. For political reasons Niedersachsen might have been chosen, although the much smaller Weser-Ems region could also have been an option. Mainly because of CBA's theoretical stance of following the impact population did Germany really prove to be the better candidate. Because of the dominance of Niedersachsen in the political discussion about the Emssperrwerk, the choice was made to analyse the level of Niedersachsen as well. The global level was accepted as part of the analysis because the

¹⁸ The research team mainly consisted of two people. Apart from the current author, Ing. Katrin Bettels (sadly deceased in 2003) was involved. She was both an agricultural engineer and (then) a student in environmental sciences who did her Master thesis on the evaluation.

¹⁹ The Waddenvereniging was familiar with parts of the evaluation approach because a methodology for evaluation lending from the MCCBA approach had been designed on their request in another research project (Sijtsma, et al., 1998).

underlying rationale²⁰ for the methodology was not unfamiliar to the Waddenvereniging, and furthermore, because the relevant economic scale of the market – especially the competitor side of the market – granted a high international level of analysis.

8.3.3 Stage two: Involve a broad group of stakeholders

This evaluation typically shows the difficulty of involving a broad group of stakeholders. Other than the Waddenvereniging, there were no other stakeholders involved in the evaluation except university staff, which is a serious deviation from the MCCBA ideal. Many stakeholder groups existed and many were active; German environmental organisations, the local German population in villages near the Emssperrwerk, the Papenburg population, the unions, other shipbuilders, the Meyer shipyard itself, to name a few. Availability of the MCCBA approach would no doubt have helped the involvement of more stakeholders, but in this case no other parties were considered.

Nonetheless, involvement of more stakeholders might not have been easy; first, the country border represented a major divide among stakeholders. Although the German and Dutch stakeholders could physically be very close, the distance in terms of decision-making relevance and ease of being heard in the policy debate was far larger for any Dutch stakeholder than for any German stakeholder. Second, to assure a maximum of broadness of views, it would have been attractive to include a representative of the Meyer shipyard in the committee that followed the progress of the research. However, the Meyer shipyard chose not to provide basic figures about its turnover, let alone involve itself in a neutral evaluation; it had a clear position: it lobbied for the Emssperrwerk; it had no need for extra information that could potentially be detrimental to their cause.

By following the MCCBA approach, it is clear that more effort should have been put into involving stakeholders with widely different views. However, by so doing, it is doubtful whether it would have then been possible to involve the most preferred parties.

8.3.4 Stage three: Organise judgement criteria on Triple E impacts

Because the research methodology had been somewhat formalised by the time of this evaluation, organising judgement criteria followed the MCCBA structure more closely. Clearly, more than one scale level needed to be specified, at least the global or high international level; economic impacts would have been given plenty of attention as would environmental degradation. Landscape impacts might have been considered, but evaluating extreme poverty impacts was (again, as in chapter 6) not foreseen.

This procedure seems defensible to the extent that Meyer's competitors are within Europe. Competing cruise ship shipyards (2004) are located in Italy, France and Finland, but there is some competition from low and middle-income countries such as China and Korea. This aspect could have been given more attention. The main difficulty here is that including this impact would have required a more thorough analysis of the competition and structure of the shipbuilding market in and around the cruise ship market segment. Lacking specialised expertise within the research team, this particular point of analysis would have required a substantial amount of time. The effort required in this case does not seem to have been beyond the practical possibilities but, perhaps more importantly, the fact of neglecting the impact does

²⁰ That is, that sustainability cannot really be assessed without a global scale.

reveal a weakness in the applicability of MCCBA, which can be formulated as follows. Since the analysis of the global level is often not the primary focus of the evaluation, assessing impacts on the global level then requires existing knowledge and expertise within the evaluation team. If this expertise is lacking, there is the danger that an overly large claim for scarce research time will be required in order to assess impacts deemed relatively unimportant at the decision-making level.

8.3.5 Stage four: Quantify impacts physically

Again, physical quantification of the impacts took up most of the time of the research team. A summary of the main results and methods was described above; at the moment two points deserve special attention: the impacts at different spatial levels, and the treatment of the landscape impact.

1) Probably the most striking feature of the evaluation was specifying costs and benefits at different spatial scale levels. It was in the stage of physical quantification that the final choice for the more local spatial level was made, and it seems worthwhile to briefly discuss this choice.

Several spatial levels were potentially relevant: the city of Papenburg, the Bezirksregierung of Weser-Ems, the 'land' Niedersachsen, and finally Germany. All these levels might have been included, but it would have been difficult to calculate and communicate results. Following an MCA approach, Niedersachsen would have been the most preferred level, as most of the political decision-making seemed to be at that level. Following a CBA approach, Germany might have been the better candidate, as the impact population clearly included Germany as a whole; the German taxpayer would pay a substantial amount of the costs involved and the central government's consent proved crucial to the project's funding. That Papenburg (as such) was too low a spatial level was obvious, as most of the impacts on the river Ems would not then be included. The Weser-Ems level also seemed too low, as the costs were only to a small extent borne by Weser-Ems.

Thus it was decided to work with two spatial levels: Germany and Niedersachsen. Very basically this dual-choice shows the tension of the analyst between trying to follow the political reality on the one hand and the demands of a correct analytical approach on the other. Although analysing at different spatial levels is obviously more difficult than working at one, there should be no doubt that specifying the impacts at these levels greatly assisted in thinking about the project. The simple gathering and structuring of information correctly at the given levels was probably the major benefit of the evaluation. Generally speaking, it seems that the use of more spatial levels in an evaluation could be very practical and fruitful.

2) As for the landscape impact, it was decided to discontinue analysing this impact. Although environmentally-focused people might see the Emssperrwerk as a major degradation of the landscape, Dutch experience has shown that this type of major infrastructure project may also generate a substantial stream of visitors. Therefore, a priori it was not clear whether a positive or negative impact would result. Furthermore, there was a lack of a clear methodology to assess this impact.

The line followed was that if – after some research – there is neither a clear positive or negative effect, nor a clear methodology for its measurement, one can skip the criterion. Beyond that, if a criterion is not seen as having a high degree of importance, then it may be better to ignore it.

8.3.6 *Stage five: Aggregate monetary scores consensus based*

The consensus based valuation and aggregation in the Emssperrwerk evaluation focused on the benefit side of the impact on the production value that Meyer realised in different alternatives (the value added of Meyer and its suppliers) and on the investment and maintenance costs on the cost side. The reader may recall the conclusion in chapter 7 about consensus based: “no consensus has to emerge about the completeness of the analysis; the consensus should concern importance and relative ease of understanding the measurement.”

As shown above, the reasoning in the evaluation proceeded roughly as follows: given the high degree of competition in the shipbuilding industry, it seemed reasonable to assume that with or without the Meyer shipyard the ships would be built anyhow. Losses to employees and management were assessed/estimated through the value added changes and ‘inconvenience costs’ compensating for movement and extra travel. Changes in consumer surplus or more general changes in consumer preferences (both intermediate and final) for the ships being built in Papenburg or – say – Italy, were thought to be unlikely.²¹

But the policy debate was not around the issue of consumer preferences, and it was not restricted to the consumers of either the ships or the cruise trips. The debate was mostly at the Niedersachsen or German level and was about keeping a given amount of employment at a given location or region, and thereby affecting nature values at a given location. It seems clear then that the approach of focusing on the location of the value added remained close to the perception of the economic issue in the eyes of many stakeholders. In other words, many would agree that the balance of costs and the impact on the growth of Meyer’s turnover (direct and indirect value added) was an important criterion and natural to monetarize.

Again here, as in chapter 7, it seems that this type of consensus cannot be claimed for the strict reasoning that was followed and the various assumptions and estimates that had to be made along the way. These latter areas seem far more the domain of the analyst, and (depending on the specific assumption or estimate) of one or two specialised stakeholders. Generally, the point can then be made that consensus based aggregation of monetary impacts mainly requires that the items included are perceived by most stakeholders as easily monetarized. The criterion for judgement in the CBA part might be classified as economic efficiency.²²

Specific elements of the evaluation seem to require too much specific expertise for a general consensus to exist at the detailed level. This fact did even prove relevant within the research team. Although the economic analysis was not complicated as such, it turned out that to a well-

²¹ Obviously the assumption of consumers being indifferent to the cruise ship being built in Germany or Italy or France could be challenged and market imperfections – government support for shipyards and labour market imperfections – could have been analysed in more depth and this might have led to different choices concerning the correct point of measurement of impacts (compare Oosterhaven and Elhorst, 2003 versus Rouwendal and Verhoef, 2003, for a detailed debate on the correct point of measurement for a specific case). Furthermore, the technique of building and designing cruise ships is developing rapidly, and the market is not one of many small suppliers who mainly compete on price. Different suppliers continuously develop competitive advantages that consumers (or shipping companies) can appreciate. In recent years shipyards have increasingly competed on the time needed for building: due to sophisticated logistics, building ships in sections and preventing the intervention of different production processes on different parts of the ship with each other (such as painting and sandblasting), the time needed for building has been greatly reduced at Meyer. So the assumption of leaving aside the consumer-side of the impacts can be challenged, but all these considerations have little bearing on the applicability of MCCBA in general.

²² That is, the debate in the economic part centered on the (investment) costs needed for securing or realising the production value in a certain place.

trained engineer and environmental sciences researcher, it was difficult enough – just as (conversely) the assessment of which species would suffer or benefit from different alternatives was not within the knowledge realm of the more economically-focused current author. The availability of a multi-disciplinary research team – one way or another – will therefore often be a prerequisite for evaluation in a sustainability context.²³

8.3.7 Stage six: Aggregate non-monetary scores consensus based

When reviewing the outcomes it may be clear that the CBA part of the analysis has covered a lot of ground; thus for the MCA part, not much remained.

The safety impact for example, could not be quantified. Or, to put it a little stronger, it could not be determined whether real changes compared to the nul-alternative would occur. In the end results not much attention was given to this impact, which follows the MCCBA approach of focusing on the most important criteria – and disregarding impacts that prove to be unimportant.

The nature score was therefore the most prominent non-monetized score; its importance was limited mostly to the renaturalisation alternative, although the result of no nature impact of the Emssperrwerk itself was also not beyond significance. The importance of assessing the impact on nature seems to have been beyond dispute. However, this fact does not imply a consensus about the particular way of measurement in this evaluation.

The aggregation of the nature value indicators followed the approach taken in chapter 7. Germany does not have a comparable system in its nature policy and the evaluation process did not actively involve stakeholders. Thus, the aggregation cannot be claimed to have had (or gained) wide acceptance. In a later research (Futh, 2001) a first attempt was made to improve on this state of affairs. An estimate was made of nature score totals for Niedersachsen and Germany using only German data. Such an approach will obviously be easier to understand to German stakeholders.

8.3.8 Stage seven: Interpret trade-offs

For a correct interpretation of the outcomes, project impacts would have to be interpreted at the relevant levels of Niedersachsen, Germany and the world. For the CBA outcomes the total Gross Domestic/Regional Product seems to be a logical benchmark in a ratio analysis. However, this possibility was used in the discussion following the evaluation but not in the main report.

With regard to the nature values the calculation of approximate total amounts of nature value for the relevant scale levels proved to be possible only for Niedersachsen and Germany. The approach here was based on rough approximation on the basis of Dutch basic data.²⁴ Clearly however, this type of data gains in strength when more projects are evaluated in a more or less standardised way. The stakeholders and analysts are then able to develop experience as to whether a certain impact is large or small.

A very interesting element was used in the presentation of the end results. The end result really

²³ Needless to say, perhaps that although no broad consensus can be achieved about all evaluation details, the analyst should still strive for as much consensus as possible. He or she may start with consensus among experts. This (latter) peer-review procedure is, however, standard scientific practice.

²⁴ As said above, the study of Futh (2001) used German data.

consisted of a table with 42 numbers, as there were seven alternatives, three scale levels and two criteria scores, one for the net CBA result and one for the impact on nature value. Such a problem can be rather typical for an approach that resists aggregating without consensus. A simple MCA-like reduction was performed into plusses and minuses showing whether an alternative was desirable, to be considered or undesirable. No real weighing took place; the procedure simply identified ‘win-win’, ‘loss-loss’ and mixed situations for two criteria. This procedure reduced the number of end result observations to 21.

Obviously this table could not replace the non-aggregated results, but from a presentation and communication perspective, it was thought to be worthwhile. In hindsight this table does not seem to have given much added value. The main drawback is that it involves a rather serious information loss. Using more ratio analysis – as mentioned above with the interpretation of size using national totals – seems to provide many more options for extra judgement while simultaneously retaining much of the information content of the end figures of stages 5 and 6.

Using a stakeholder’s perspective analysis²⁵ might also be more fruitful. Information loss may be inevitable but the reduction of end figures can be much stronger.

8.3.9 Stage eight: Perform sensitivity analysis and reconsider project alternatives

Sensitivity analysis

Directly following the main results at the German level a sensitivity analysis was presented as part of the summary of results. The reason for this course of action was because four of the major assumptions/estimates were uncertain. Further questions were analysed:²⁶

- What if the Meyer shipyard can have 20 years of extra growth instead of 10?
- What if the yearly extra growth of Meyer was 0.25% instead of 0.5%?
- What if in the Eemshaven renaturalisation variant instead of 15%, a much larger percentage (85%) of the employees becomes Dutch?
- What if the once-only compensation (inconvenience-costs) are not €15000 but rather €50000?

This analysis clearly showed that the estimate on extra growth and the extra growth period that Meyer might achieve were the most crucial variables. The substantial costs involved in building the Emssperrwerk require a prolonged period of growth of the Meyer shipyard.

No sensitivity analysis was performed for the other levels. Splitting the analysis in three spatial levels has clearly shown the sensitivity of the outcomes for choosing an impact area or population. This treatment showed that at too-low levels the project is easily beneficial.

Furthermore, the analysis indicated that there is a wide range in which renaturalisation alternatives may be serious alternatives.

Reconsidering project alternatives

The question of reconsidering project alternatives was not explicitly addressed in the evaluation, as it was not part of the MCCBA approach at that time.

²⁵ Say, using a nature conservation perspective of a Weser-Ems entrepreneurial perspective.

²⁶ See appendix for details.

8.3.10 *Epilogue on uncertainty*

This type of evaluation is typically an ex-ante activity with a tight time-constraint. In this epilogue a complete ex-post evaluation and interpretation is not made of what happened since the evaluation was finished in 1998. A few major points concerning uncertainty will be touched on that may be useful for the assessment of the applicability of the MCCBA approach.

The costs of the Emssperrwerk

The Emssperrwerk was built in the years after the evaluation – *despite* the outcomes of the evaluation. It was built without a lock. Although some extra millions proved necessary beyond the foreseen (cheap) estimates, the project seems to have done well as far as the budget for investment costs is concerned.²⁷ This result may be somewhat unexpected to many, and is contrary to many larger infrastructure projects. However, it is too early yet to assess the maintenance expenditures.

Meyer investment and the business cycle

The research assumed that the large dry dock construction hall – especially its contents – could be moved to Emden or Eemshaven. Because of the Emssperrwerk, this operation was not necessary. However, Meyer did not simply keep its existing dock at Papenburg, as was assumed in the evaluation. Meyer instead built a new and larger dry dock construction hall as well as a technology hall for welding automatically with a laser technique. The size and volume of the buildings on the shipyard have as a result more than doubled, and Meyer invested approximately €200 million. Meyer is now able to build larger cruise ships than before and work on more ships at a time. Interestingly this proved that the estimated costs of relocation of the shipyard seemed to be fairly accurate.

The decisions to invest and the investments (at least partly) were made when the business cycle was at its peak but it has since gone down. While at the top of the cycle, speed in building was a great advantage; Meyer was asked by its shipping company customer in 2003 to postpone delivery of a new cruise ship – a procedure used in order to prevent financial trouble for the shipping company. Although this implied high costs to Meyer, the request was granted.

Meyer's direct employment had grown substantially after 1998 (from 2000 employees to some 2500). By mid-2004 direct employment seemed to have returned to its 1998 level. Since in the cruise ship segment of the shipbuilding market there were no orders, Meyer had accepted the building of other types of ships (gas container ships). In this way it could keep its direct employment (welders) at a reasonable level. The most severe impact of the drop in cruise ship orders is apparently on indirect employment. Despite building Emssperrwerk, many suppliers have suffered.

It is telling to observe the tremendous dynamics in the economic system – even in a segment generally regarded as a growth market – the market for cruise tourism. Meyer builds a new dock to facilitate the building of two ships at one time; a couple of years later it has to artificially slow down the building of the one cruise ship order remaining, and has to take

²⁷ Other smaller investments of the local government were necessary too: like a larger lock door close to the shipyard.

refuge in building gas container ships. Even bankruptcy is a real risk. Among these myriad dynamics the MCCBA evaluation is trying to predict the extra growth of the Meyer shipyard over 10 or 20 years. If one thing is again shown in this example, it is that 10 years in future is a very long time in the current economic system.

Appendix 1 to chapter 8: Sensitivity analysis for four crucial parameters

Share Dutch employees

What if, in the Eemshaven renaturalisation variant, instead of 15% a much larger percentage (85%) of the employees becomes Dutch?

Relocation of the shipyard to Eemshaven becomes far less attractive from a German perspective. The net benefits of Euro +69 mln turn into net costs of Euro -105 mln. With this share of Dutch employees this alternative is about as attractive to Germany as building the Emssperrwerk without a lock – according to the cheapest cost estimates (Euro -98 mln).

Table App.-8.1 Sensitivity analysis of the share of Dutch employees (level of Germany)

Germany 85% Dutch employees in Eemshaven (instead of 15%)		
	Benefits minus costs (Euro mln)	Change nature value (%)
<i>Renaturalisation</i> Alternative 3b Eemshaven	-105	+ 0.009 %

An extra growth period of 20 years

What if the Meyer shipyard can have 20 years of extra growth instead of 10?

This extra growth period implies that all cost-benefit results of the Emssperrwerk become over €100 mln more profitable. In most alternatives the Emssperrwerk alternatives still yield negative results. At the German level the cheapest cost estimate now gives a modest positive balance of €10 mln.

Table App.-8.2 Sensitivity analysis of 20 year extra growth period (level of Germany)

Germany Meyer can grow an extra 20 years because of the Emssperrwerk (instead of 10 years)		
	Benefits minus costs (Euro mln)	Change nature value (%)
<i>Emssperrwerk</i>		
Alternative 2a	-90	0%
Alternative 2b	+10	0%
Alternative 2c	-533	0%
Alternative 2d	-381	0%

Lower relative growth of Meyer

What if the yearly extra growth of Meyer compared to other sectors in the economy was 0.25% instead of 0.5%?

This difference leads to changes in outcomes for 6 alternatives: the benefits become lower, which yields more negative results. The Emssperrwerk with the cheapest cost estimate (Alternative 2b) now yields a negative result of €200 mln.

The only alternative with a positive result is the Emden alternative. The costs of relocation still outweigh the extra benefits.

Table App.-8.3 Sensitivity analysis of lower relative growth of the Meyer ship yard (level of Germany)

Germany Lower extra growth (0,25% instead of 0,5%)		
	Benefits minus costs (Euro mln)	Change nature value (%)
<i>Sperrwerk</i>		
Alternative 2a (W.-Ems + lock)	-285	0%
Alternative 2b (W.-Ems minus lock)	-185	0%
Alternative 2c (EZ NS + lock)	-727	0%
Alternative 2d (EZ NS minus lock)	-576	0%
<i>Renaturalisation</i>		
Alternative 3a Emden	+55	+ 0.009 %
Alternative 3b Eemshaven	-113	+ 0.009 %

Higher inconvenience compensation

What if the once-only inconvenience compensation to every employee is not €15000 but €50000?

Even if the costs of the Eemshaven alternative increase (by the higher inconvenience compensation) the result is still positive. The rank of the alternatives does not change. The net result of the Eemshaven alternative with these higher costs is quite a bit lower than the Emden alternative.

Table App.-8.4 Sensitivity analysis of higher inconvenience compensation (level of Germany)

Germany Once-only inconvenience compensation not Euro 15.000 but Euro 50.000 per employee		
	Benefits minus costs (Euro mln)	Change nature value (%)
<i>Renaturalisation</i>		
Alternative 3a Emden	288	+ 0.009 %
Alternative 3b Eemshaven	10	+ 0.009 %

Appendix 2 to chapter 8: The results of the total nature score approximation for Niedersachsen and Germany

Table App.-8.5 Approximation of nature score total for Niedersachsen and Germany

Area	Area size in 1.000 ha (1)	Nature score per ha (2)	Area times nature score (*1000) (3)
Niedersachsen			
agriculture	2888.7	1.0	2888.7
nature	1212.8	7.0	8489.6
other land use	633.7	0.1	63.4
total	4735.2	2.4	11441.7
Germany			
agriculture	19414.1	1.0	19414.1
nature	11817.2	7.0	82720.4
other land use	4465.7	0.1	446.6
Total	35697.0	2.9	102581.1

CHAPTER 9: DESIGN AND EVALUATION OF SUSTAINABLE CORPORATE PERFORMANCE (SCP) POLICY: THE VALUE OF CAUSAL ANALYSIS

9.1 Introduction

In recent years initiatives concerning Corporate Social Responsibility (CSR) of businesses have grown strongly and a connection has been established between Corporate Social Responsibility and sustainability (Elkington, 1997; SER, 2000; WBCSD, 1999, 2000; Consumentenbond, 2001). Debate on the evaluation of CSR in the sustainability context has intensified as well (Ranganathan, 1998; Figge et al., 2002; GRI, 2002).¹ This chapter discusses CSR and sustainability, the combination of which will henceforth be called Sustainable Corporate Performance (SCP).

The Ahold Company, one of the leading Dutch multinationals, prior to 2000 was in the process of formulating its SCP policy. Two closely connected aspects were particularly challenging to this food retail firm: first was the confusion about what SCP was or should be, and second was the lack of standardisation in measurement of SCP.² Therefore, Ahold supported a research project in 2000 at the University of Groningen entrusted to develop a 'golden standard' for SCP.³ This golden standard would define what SCP should be, how it could be managed, evaluated and reported.⁴ The results presented in this chapter are derived from this research project, although they should not be seen as the major outcomes of the team effort.⁵

The main objective of this chapter is to describe how MCCBA can be useful for evaluating SCP for business. In order to reach this goal, two main questions are addressed:

1 (How) can MCCBA be helpful in designing SCP policies? (section 9.2)

2 How would a concrete evaluation of such an SCP policy example – implementing EurepGAP⁶ – look? (section 9.3)

¹ The debate on evaluation of SCP may clearly be understood as a debate on the (im)possibilities of CBA and MCA alike to measure the sustainable performance of firms. Much attention has been dedicated to proving that, in the end, a financial CBA-type of evaluation could sufficiently capture all broad societal concerns (see e.g., Waddock and Graves, 1997 for a discussion); meanwhile several attempts have been made to depend exclusively on MCA types of evaluation (e.g., Bernadette et al., 1998).

² Business performance is traditionally evaluated using different forms of cost-benefit analysis, in which financially evaluated efficiency and return on investment are the essentials. However, several non-financial indicators have gained in prominence too. See Ittner and Larcker, 1998 and 1998b for discussion of financial CBA measures in relation to other non-financial measures.

³ Later on Aegon and ABN-AMRO joined Ahold in funding the project.

⁴ A large multidisciplinary team was formed of which the current author was a member.

⁵ See Vlek et al., 2002 for a first outcome of the team effort.

⁶ The EurepGAP protocol is a standard for safe, environmental and socially responsible production of agricultural fresh products. Ahold was one of the initiators of this standard.

9.2 Designing SCP policies following the MCCBA approach⁷

9.2.1 Introduction

To answer the main question of the design of business SCP policy the following topics are addressed. First, what (type of) activities are generally regarded as part of SCP? (9.2.2). Second, in order to identify and appreciate the added value of MCCBA, existing judgement-oriented evaluation tools have to be discussed (9.2.3). Finally, the question is raised about whether and how the MCCBA approach is helpful in designing SCP policies (9.2.4).

9.2.2 SCP in general

9.2.2.1 Definitions

As referred to above, the notions of SCP and CSR have become closely interwoven. CSR has been the topic of scientific debate in the field of business and society for at least half a century.⁸ Bowen in 1953 defined CSR: “It refers to the obligations of businessmen to pursue those policies, to make those decisions, or to follow those lines of action which are desirable in terms of the objectives and values of our society” (Carroll, 1999). Obviously this definition does not clarify what any specific company should do precisely, but it does make clear that it should somehow become attuned to societal values and aims. Davis (1960) puts further emphasis on the fact that CSR is about: “Businessmen’s decisions and actions taken for reasons at least partially *beyond* the firm’s *direct economic* and *technical interest*” (our italics, quoted from Carroll, 1999). This view of CSR derives from defining or assuming an everyday economic view of ‘economics’. The phrase ‘at least partially’ reveals an often-recurring ambivalence; CSR should be connected with the core business of a company, but it should also at the same time be more than that. Or shouldn’t it?

Over time the concept of CSR has become increasingly refined. Donna Wood made an important integrative step. Corporate Social Performance, as she called it, is “a *business organisation’s* configuration of *principles* of social responsibility, *processes* of social responsiveness, and policies, programs, and *outcomes* as they relate to the firm’s societal relationships” (Wood 1991, p. 693, our italics). Apart from the division in principles, processes and outcomes, Wood introduced a division in scale levels: the institutional, company and individual level. At different levels, different questions are relevant.⁹ An important element from earlier literature integrated in Wood’s model was the issue of ‘stakeholder management’. Wood considers this a relevant process at the business level. Various authors have more recently defined the quality of stakeholder relations as crucial to CSR (Ranganathan, 1998; Waddock and Graves, 1997; compare also McWilliams and Siegel, 2001).

Since the 1990s CSR has increasingly been linked to sustainability, which is probably best shown by Elkington’s so-called Triple Ps (or the Triple Bottom Line); CSR linked to

⁷ This section is largely based on three articles: Sijtsma and Broekhof (2002), Broekhof and Sijtsma (2003), Sijtsma, Broekhof and Nijboer (2003).

⁸ For an overview, see Carroll, 1999; see also Broadhurst (2000) and Sison (2000) for CSR and business ethics.

⁹ For instance, at the institutional level the role of public sector versus private sector is an important issue. At the company level it is important what an individual company can do. At the individual level – that of manager or employee – other questions are relevant.

sustainability is about people, planet and profit (Elkington, 1997; WBCSD, 1999 and 2000; SER 2000; Pope et al., 2004). As stated earlier, the combination of CSR and sustainability will be called Sustainable Corporate Performance (SCP).

The subsection below scrutinises the practical content of SCP.

9.2.2.2 SCP business practices

To gain more insight into what is generally regarded as SCP activities, this section analyses a collection of 250 best practices of SCP, published by the organisation *Business for Social Responsibility* (BSR) as so-called *leadership examples*.¹⁰ These leadership examples were grouped under 68 headings.¹¹ To facilitate closer inspection of the examples without being submerged in minor details, the collection of 68 groups of leadership examples are categorised according to stakeholder group to which they *primarily* relate. The economic-analysis-based stakeholder categories of Henriques and Sadorsky (1999) were used; the natural environment was added as a separate stakeholder, resulting in a total of 10 stakeholder groups. Activities can relate to one stakeholder group, to several groups and to all stakeholder groups.

Table 9-1 illustrates the division of the 68 categories of activities according to stakeholder groups. A total of 13 categories relate to all stakeholders. The final columns of Table 9-1 show the division without these 13.

Table 9-1 Division of SCP activities among stakeholder groups

	Number of times an activity relates to a stakeholder group.			
	All activities		Excluding 13 activities that relate to all stakeholder groups	
Stakeholder groups	absolute	% (of 68)	absolute	% (of 55)
Employees	48	71%	35	64%
Community representatives and activists	36	53%	23	42%
Natural environment	27	40%	14	25%
Suppliers	25	37%	12	22%
Government	22	32%	9	16%
Customers	21	31%	8	15%
Competition	20	29%	7	13%
Professional organisations	18	26%	5	9%
Shareholders	15	22%	2	4%
Media	15	22%	2	4%
All stakeholder groups	13	19%	0	0%

Source: www.bsr.org (Oct-2000); own calculations.

In our discussion the practices relating to all stakeholders will be considered first. Table 9-1 shows a clear Top 2 of stakeholders to which CSR activities are related: employees and

¹⁰ Companies could present these cases of their own accord; there was only a marginal test from the part of BSR. In October 2000 on the website of BSR, about 250 leadership examples of CSR were presented.

¹¹ Two empty categories were disregarded. In later years BSR has regrouped these leadership examples under different headings (www.bsr.org).

community representatives and activists. The discussion will continue with activities towards these stakeholder groups and then proceed down the table in the order presented, until the stakeholder group of customers has been analysed. In the remaining categories no activities are found that have not previously been discussed under earlier categories.

In the discussion of the practices, relevant evaluation issues will be addressed briefly.

9.2.2.2a Practices related to all stakeholders

The first group of practices under this category of the BSR database can be called working on values within the company and focuses on discussing and institutionalising values, providing training and making codes of conduct: “What do we stand for?” “Where do we want to go?” When making values more explicit, there is a natural mix of ‘economic’ or business values with other values. A illustrative example of this mix in the BSR database is the set of seven values that the management and employees of McMurtry Publishing developed: 1) do the right thing; 2) help one another; 3) deliver raving customer service; 4) produce quality always; 5) embrace change; 6) accept social responsibility and 7) earn a reasonable profit. This list above all shows that SCP can easily be about every ‘good thing’ in life, including various aspects of sustainability as defined in chapter 2. At this level it also shows that concerns or values are very general and might be valid for every company. Finally, one may notice that, without a view on the economic specifics that limit the *actual possibilities* for realising these values, there appears to be little to evaluate.

The second group of practices concerning all stakeholders is what is known as increased accountability. Businesses have several obligations – often enforced by law – to provide account of their financial performance. SCP invokes many companies to also account voluntarily for the environmental and social aspects of their functioning. Verification (external) and transparency are key words for the practices under this heading. That the issue of accountability to a broad range of stakeholders is closely linked to evaluation in the sustainability context was discussed earlier.

A final group in the BSR database is called organising SCP. The organising process can for example relate to the number of women or minorities on the board, but also to new ways of formalising the organisation of stakeholder involvement: an ombudsman for stakeholders (Coca-Cola), regular consultation through surveys or panels (British Telecom), or plant visits (Novo Nordisk). The stakeholder consultation is sometimes limited to ethical issues or environmental issues and sometimes limited to a specific stakeholder group. More often it is about a whole range of issues and a broad range of stakeholders: customers and suppliers but also representatives of the local communities. The aim is to gain insight into the views and considerations of different parties.

A major dilemma for an evaluation is obvious and was already discussed in chapter 2: involvement of stakeholders may be desirable but what about representation and how to avoid the pitfall of evaluating performance against all the concerns of all stakeholders? In other words, is every concern equally binding or relevant, given the economic possibilities of the company?

9.2.2.2b Practices relating to employees and the community

Employees

The most important practices relating to employees in the BSR database can be combined under the heading of improving labour conditions and three types may then be distinguished: 1) improving the worst labour conditions, 2) improving bad conditions, and 3) further

improving rather good conditions.

The first group of practices relates to child labour, forced labour and abuse of human rights. These practices most often involve other stakeholder groups too. The second group relates to paying minimum wages, allowing labour union influence and providing opportunity for prayer at work. The third group is about optimising fairly good labour conditions: more opportunities for maternity leave, child care, fitness facilities at work, and more part-time job offers. Clearly especially the first (and perhaps also the second group) is linked to the sustainability issue of extreme poverty.

The community

In the BSR database an important part of the practices in which community representatives and activists are involved concerns improvement of labour conditions – these have already been discussed.¹²

Furthermore, practices in this category relate to philanthropy. At least three forms of philanthropy can be found. Giving away products: 1) Microsoft giving away 5000 computers or Merck giving away drugs for river blindness. 2) Companies also give away 1% or 5% or 10% of their profit to charities. A strategic use of philanthropy can often be seen when companies aim their philanthropy efforts at neighbourhoods where most of their customers are. 3) A third form of philanthropy is to arrange cheap credit; for instance Nike setting up micro-credit institutions in Vietnam, Indonesia and Thailand, or companies lending money for cheaper housing.

Another practice under this heading is support for disadvantaged groups including ethnic minorities, long-term unemployed or ex-convicts by giving training, hiring personnel, setting up shops. This type of activity is often undertaken in the neighbourhood, city or region where the company produces or sells.

9.2.2.2c Practices relating to other stakeholder groups

Environment

Practices relating to the environment are manifold in the BSR database and the topics might look familiar: clean water from textile factories, green design of buildings, new recycling-friendly design of automobiles, certification of wood, lower energy use. However, what may come as a surprise are the high aspirations one may discover; for instance, Fetzer vineyards aim at zero-waste factories.

More importantly, there is often a close connection between environment and common business practices, which is why from an evaluation perspective there may be room for a balance of monetarized costs and benefits. An example from the database can illustrate this point.

Dundee textile factories set up a comprehensive water management system with intensive monitoring and analysis of water streams and training of employees; all cost items to the firm. This system has led to more re-use of water and recycling of textile paint, which are at least partly cost savings or benefits. Inspired by this process, Dundee established a new product line of natural towels, either without dye or with recycled textile dye. Provided these product lines are a success, they imply greater benefits.

¹² Environmental concerns are found less often here, which may show an American bias in the leadership examples, as this situation is unrealistic for the European setting.

Suppliers

Suppliers naturally determine a segment of the quality of the product or service a company achieves. A part of the SCP practices found in the BSR database is therefore aimed at suppliers. The relevant issues are the environment, improvement of the worst labour conditions and transparency about labour conditions. Boeing invites its suppliers to participate in groups of 40-50 persons in a week-long forum on how to tackle environmental issues. Ford asks its 5000 suppliers to comply with ISO 14001, which includes independent monitoring of environmental performance.

For evaluation purposes an understanding of the market situation and market-related power relations is crucial here. What type of suppliers are invited? Some suppliers may be delivering exclusively to (e.g.) Boeing or Ford, some have numerous other customers, some have long-term contracts, others only incidentally deliver. Some have one product, others have a product range. It seems that evaluative judgement should depend on specifics, as they determine to a large extent the realistic options of a firm.

Government

Government or governments are not mentioned often in the BSR database as an important partner in the best practices. If they are involved in SCP practices other community representatives are also always involved. However, through its regulation and laws, government is often present in the background. If firms do not meet current or future regulations or laws, this fact implies potential cost items. SCP projects are easily motivated by this 'cost-threat'.

Consumers

The practices relating to the consumer are interesting as (again) one can see the strong connection of CSR with both sustainability issues and the core business of companies. A popular activity in this category is cause-related marketing: marketing where a company is actively promoting a good cause together with its own product. The link between the product and the cause is generally not that strong, although there is often a link at a more abstract level.

An example from the BSR database of such an abstract connection is the American producer of gourmet sets that supports anti-hunger organisation SOS by donating a small share of revenues of a specific pan. Since the SOS logo is on that specific pan, its sales have increased tenfold.

Another illustrative example in the sustainability context –directly linked to poverty – is aimed at consumers and comes from India. One of the world's largest blue jeans producers (Arvind Mills) invented a new formula to sell the jeans to some of the poorest people in the world. Blue jeans normally would cost somewhere between \$20 and \$40 and the distribution system does not reach the smallest rural villages of India. Arvind introduced Ruf and Tuf jeans: a do-it-yourself-package with cloth and zipper distributed to 4000 small tailors around the country, tailors who had not ever sold blue jeans. At a cost of only \$6, Ruf and Tuf jeans, according to Arvind, have become the highest selling jeans in India.

From an evaluation perspective it is interesting to observe that in these two examples, without economic analysis or cost-benefit analysis the evaluator might be tempted to appreciate only the aspect of raising money for an anti-hunger organisation or the provision of clothes to the poorest people; two positive scores on two sustainability-linked criteria. Against the background of economic analysis, the evaluator may see that business opportunities are created and grasped here. Sustainability may be served by that, but a correct perspective on the

effort a company puts into SCP seems also to involve costs and benefits for the company as well as at the higher community level.

9.2.2.2d Conclusion

If one remembers the four key items of sustainability¹³ and compares them with the initiatives discussed above, then the following may be observed:

- Generally speaking, stakeholder involvement is characteristic of the SCP examples above.
- Global and long-term aspects of sustainability are not always specifically addressed in SCP examples, although they may not be completely absent.
- Triple E concerns can be seen as part of these SCP examples, but the examples are often much broader than Triple E concerns alone,¹⁴ and the importance of the relation of the Triple E elements (especially links with economic performance) is hardly addressed.

9.2.3 Existing SCP evaluation tools

9.2.3.1 Introduction

This section discusses several existing SCP evaluation methods. The goal of this discussion is to provide a background for the merits of the MCCBA approach.

In the last decade several SCP evaluation methods have been developed. These initiatives can be categorised according to their practical aim. Table 9-2 gives an overview of the aims and some examples of evaluation methods.

Although all these evaluation methods differ substantially from one another, most share the characteristic of covering a wide range of concerns for evaluation. This section focuses on the most judgement and measurement-oriented evaluation systems; those that a priori seem to fit in best with MCCBA. It will therefore focus on the methods developed for investors' advice and screening.

¹³ Compare figure 2.4.

¹⁴ Using the Triple P terminology and comparing it with the Triple E concerns, the 'People' bottom-line is much broader than Extreme poverty by itself.

Table 9-2 Examples of CSR evaluation methods according to user aim

User aim	Codes of conduct for business	(Annual) Reporting of business	Strategy development and advice of/for companies	Investors advice and screening
Initiators	NGOs, International organisations, Governments	Governments, Partnerships	Businesses, Consultants	Research companies
Examples	SA 8000, ¹ ILO Conventions & Recommendations ² AA 1000, ³ OECD Guidelines ⁴	GRI, ⁵ Environmental regulation/laws Annual reporting standards	SSC, ⁶ Additional reports	DJSI, ⁷ FTSE4Good ⁸ DSI ⁹ BE100 ¹⁰
<p>1 Social Accountability Standard (www.sa-intl.org), 2 International Labour Organisation (www.ilo.org), 3 Institute of Social and Ethical Accountability's AccountAbility Standard (www.accountability.org.uk), 4 OECD Guidelines for Multinational Enterprises (OECD Guidelines; www.oecd.org), 5 Global Reporting Initiative (GRI; www.globalreporting.org), 6 Sustainability Score Card (SSC) of DHV consultants (www.dhv.nl), 7 Dow Jones Sustainability Index (DJSI, www.sustainability-index.com), 8 FTSE4Good (www.ftse.com/ftse4good/), 9 Domini 400 Social Index from Kinder, Lydenberg and Domini (www.domini.com), 10 Business Ethics 100 Best Corporate Citizens (www.business-ethics.com).</p>				

These investor indices were developed to inform institutional and private investors about sustainability of their investors' 'universe', that is, all the possible shares in which they could invest – given certain criteria.¹⁵ The indices mentioned in the table all have operational models to measure sustainability in a broad sense.¹⁶ The models that are transparent about their methods are MCA-like.

Table 9-3 shows some of the most important characteristics of the sustainability indices.

¹⁵ Compare Plantinga and Scholtens (2002) for a comparison of the performance of mutual funds with stated social responsibility objectives and others.

¹⁶ Some publishers of these indices publish other – regular – indices along with their sustainability indices (Dow Jones and FTSE). Examples of regular indices are the American Dow Jones Industrial Average that gives the market value of 30 large American companies; Standard and Poor's 500 (an index of 500 companies) and in the Netherlands for instance, the AEX with 24 Dutch funds. Apart from these general and well-known indices, there are many more, for instance indices that aim at realised growth (E.g., FTSE350 – with 350 companies with above average growth); indices aimed at the European market only (Dow Jones Stoxx, Euronext) or indices covering the real estate market (S&P Real Estate Investment Trusts – with 100 real estate funds). Instead of meeting the standards of size or growth etc., the companies enclosed in the sustainability indices have to meet the standards for sustainability.

Table 9-3 Overview of examples of measurement aspects of different sustainability indices

	Dow Jones Sustainability Index	DSI/ Business Ethics 100 ¹	FTSE4Good
Economic	Quality of management strategic planning Corporate governance Risk management	Exclusion of companies with severe financial problems/ Return to shareholders	None
Social	Stakeholder involvement Social policy Standards for suppliers Social reporting	Philanthropy Minorities policies Employee relations	International human rights Code of ethical conduct
Environmental	Environmental charters Environmental management system Additional reporting	Recycling Waste prevention Innovation	Use of targets Additional reporting Responsible management
Other	Sector-specific risks and opportunities	Contribution to reducing poor-rich gap / West - 3rd world exclusionary screening (alcohol, tobacco, nuclear, gambling & weapons)	High impact companies with extra responsibility
1 DSI (Domini Social Index from Kinder Lydenburg and Domini) and BE 100 are strongly related as BE100 is based on the DSI for the social and environmental part and adds the financial performance – return to shareholders – to it.			

How do these indices measure and evaluate sustainability? Of the sustainability indices mentioned in table 9-3, the Dow Jones Sustainability Index and the Business Ethics 100 will be discussed, as these indices are rather transparent. Furthermore, they cover all three economic, environmental and social performance areas and they seem to be representative for the general approach.

9.2.3.2 The Dow Jones Sustainability Index (DJSI)

In 1999 Dow Jones (DJ) initiated a *sustainability index* to assess companies that pioneer in the area of corporate sustainability.¹⁷ The Dow Jones Sustainability Index has a universe of stocks in which investments can potentially be made. These stocks or rather, companies are evaluated on their sustainability performance. Performance is assessed per sector; the best 10% of companies are included in the index.¹⁸ How then is the performance measured?

The measurement takes the form of a hierarchical MCA. Criteria scores are determined on the basis of a range of sources, of which a questionnaire to the company is key. If senior management has signed the questionnaire, the weight given to the questionnaire answers as a whole increases from 50% to 70%. Questions are divided into the following categories: (in brackets the percentage weight is shown):

¹⁷ After introducing the first worldwide index, Dow Jones introduced several other specialised indices focusing on the Eurozone or excluding certain sectors (weapons, tobacco, etc.).

¹⁸ The best 20% for the Stoxx index.

- economic (23.6 %)
- environmental (20.8 %)
- social (25.7 %)
- risk and opportunity related criteria for each industry group (30 %)

The questionnaire has 91 questions. Table 9-4 gives three examples.

Table 9-4 Examples of questions of the DJSI

Economic	Environmental	Social
9. Which of the following tools does your company use in strategic planning of highly uncertain issues? -Strategic scenario planning -Game theory -System dynamics -Agent-based models -Real options -Others: _____ <i>Comment:</i>	51. Does your company publicly provide information on its environmental activities? -Yes, for units generating % of total revenues -No -Not applicable -Not known <i>Comment:</i>	73. Does your company require specific health tests before hiring, also in cases where this is not necessary for specific working conditions? -AIDS-test -Pregnancy tests -Genetic screening -Others, please specify: -No <i>Comment:</i> <i>Documents (pages):</i>

Even this small random selection of three questions exemplify the broad range of issues tackled and the detail in the questions asked: down to the tools (!) used in strategic planning. Interestingly, the DJSI questionnaire contains a cross-link with some of the other SCP evaluation tools seen in table 9-2; in one of its questions (number 45) it asks about the commitment to principles of sustainability councils/coalitions and environmental charters. Commitment to more than three (!) gives the company the highest score on this point. Apart from questionnaire answers, the quality and public availability of information is assessed (and given a 20% or 33.3% weight¹⁹) and finally, the truthfulness of the information and the cooperation with the assessment process is given a 10% or 16.6 % weight.

9.2.3.3 Business Ethics 100

The Business Ethics 100 has many of the same elements that characterise the Dow Jones Sustainability Index: a hierarchical MCA with many criteria. For the purposes of this section having a closer look at the final evaluation results of the BE100 may suffice. Table 9-5 to the left shows the Top 5 and the lowest 5 companies in the Top 100 and their total scores, which range from 0.342 to 1.586. This score is the un-weighted aggregate of 7 sub-scores. Scores basically indicate above (or below) average performance²⁰. The first sub-score is a financial return score, expressed as a percentage return on investment²¹. The other six criteria (Communities and so on) typically involve 5-10 qualitative subtopic scores (strengths and concerns on various sub-topics) which are aggregated to show above or below average performance. Since all seven variables have different scales they are standardised. For all

¹⁹ The latter in case of a signed questionnaire.

²⁰ Positive scores are above average (in the set of approximately 650 companies), negative scores are below average.

²¹ However, - oddly - 'companies with losses in 2000 were eliminated [from the database]' (www.business-ethics.com/whats_new/2002_100_best_corporate_citizens.html → Getting there.)

seven criteria the mean score for all companies is subtracted from the score of a particular company. The difference is then divided by the standard deviation (typical difference from the mean score) for all companies, which puts all seven stakeholder ratings on a common scale (the standard deviation from the mean). In the final step an un-weighted average of the seven measures is calculated to yield a single score for each company.²²

Table 9-5 in the columns to the right, sub-aggregated and standardised scores are shown – the return to shareholders only shows the non-standardised score.

Table 9-5 Top 5 and lowest 5 of the Business Ethics 100 Best Corporate Citizens - 2002.

Rank	Overall Score	Company	Total Return to Shareholders	Community	Minorities & Women	Employees	Environment	Stakeholders	Non-U.S. Customers
			1998-2000 Average	1998-2000 Standardised values 3-year DSI averages					
1	1.586	IBM	24.73%	2.561	3.804	1.137	1.626	0.206	1.517
2	1.233	HewlettPackard	15.95%	1.095	2.989	1.524	1.251	0.206	1.517
3	1.230	Fannie Mae	19.62%	2.072	3.804	0.75	0.126	0.206	1.517
4	1.179	St Paul Cos	17.56%	5.494	1.36	0.75	0.126	0.206	0.226
5	1.154	Procter & Gamble	3.48%	1.095	1.36	1.524	0.126	3.547	0.656
.....									
96	0.346	Adobe Systems	91.81%	-0.372	-0.268	0.75	0.126	0.206	0.226
97	0.345	Wendys International	5.23%	-0.372	0.546	0.75	1.251	0.206	0.226
98	0.345	Tribune Company	17.67%	1.583	-0.268	-0.412	0.126	0.206	1.087
99	0.344	Kroger	23.47%	0.606	0.275	0.75	0.126	0.206	0.226
100	0.342	Applied Materials	66.25%	-0.372	0.275	0.75	0.126	0.206	0.226

Source: www.business-ethics.com. Score chart 2002. Starting from 2003 *Business-Ethics* has followed a slightly different methodology, including more companies, including a separate score for governance, and using only single year scores.

The table shows that IBM ranks one in the list, followed by Hewlett Packard and Fenny Mae. IBM and Fenny Mae have very high scores on 'Minorities and women'. Adobe (at number 96) seems to have had bad community relations, which, by the way, may come as a surprise to the internet community - considering the massive use of its freely provided Acrobat Reader.

²² See www.business-ethics.com/whats_new/2002_100_best_corporate_citizens.html for further details on the methodology. Some 650 companies are assessed, including the S&P 500 plus another 150 firms chosen for social strengths.

Wendys International (at 97) easily outcompetes the numbers 3, 4 and 5 on environmental performance, while the Tribune Company (at 98) scores rather high on Non-US customers (better than St Paul Cos and Procter & Gamble). The 'Non-US' makes one wonder: would this criterion be relevant for all companies to the same extent, and why would these customers need a separate criterion?

9.2.3.4 Limitations of existing evaluation tools

The main limitations of the existing evaluation tools seem to be twofold.

- 1) large number of issues,
- 2) weak relation to the economic context in which the companies operate.

Ad 1) The SCP evaluation tools mirror the discussion of SCP practice in section 9.2. The aspects evaluated only partly relate to the sustainability context defined in chapter 2. Generally or as a whole, the issues are much broader than this context; especially in as much as they concern the people and planet side. As a result, the total 'sustainability' scores become summations of many small aspects with low weights.²³ As shown in chapter 4, this fact alone may lead to an overall score that is hard to understand²⁴ and a judgement concept that becomes vague. Examining the processes involved – e.g., open stakeholder consultations about possible issues, focus on broad values – it is easily understood why the aspects evaluated become so broad, but is this a necessity?

Ad 2) As in regular internal and external evaluation of business activities financial cost-benefit analysis in one form or another ranks high, one may be struck by the limited attention to the easily monetarized costs and benefits in the existing SCP evaluation tools.²⁵ DJSI does address the strategic management side but ignores financial performance. BE100 chooses one financial parameter, which only covers the return to the shareholder.²⁶

More generally the evaluation tools seem to show a lack of awareness about the close causal connections of different SCP aspects to financial economic indicators.²⁷ This fact may be related to the common conviction, implicitly present in the MCA approaches to evaluation of SCP, that the main purpose of business is to profit, and that helping solve problems in society is a new 'ball game'.²⁸ This perspective can easily be misleading as in reality any business

²³ See Bernadette (1998) for more insight into the sensitivity of MCA measurement (partly) based on ranking.

²⁴ See Ittner and Larcker, 1998 for a discussion of performance measurements in which attention is given to limits to understanding by personnel.

²⁵ Compare also McWilliams and Siegel (2001, p. 124) who at the end of their (theoretical) paper conclude: "The ideal level of CSR can be determined by cost-benefit analysis. [...] [M]anagers should treat decisions regarding CSR precisely as they treat all other investment decisions. [...] Managers can ...make decisions regarding CSR investment by employing the same analytical tools used to make other investment decisions."

²⁶ See Vlek et al., 2002 for a first attempt to improve upon this situation by using a broad many-aspect multi-criteria approach.

²⁷ Compare McGuire et al. (1988) who when analysing Fortune's list of companies and their performance, see high levels of corporate social responsibility after years of high level financial performance, which suggest that CSR is a 'luxury' policy. Furthermore, noteworthy is that because of this lack of connection, the scores say little about the effort a company has put into achieving these results and the benefits it may receive; as long as a company is somewhere along the track from unsustainable to sustainable, this effort will be important in judging progress.

²⁸ As seen in the examples above, implicit – and sometimes explicit – in a great part of the SCP debate is the notion that entrepreneurs should adopt a better motivation; and with the help of improved motivation come to

activity is geared towards solving some kind of problem in society;²⁹ sustainability evaluation cannot do without a clear view on the reasons why the sustainability problems occur and why they may be hard to solve.

9.2.4 MCCBA and the design of SCP policies

9.2.4.1 Four MCCBA elements fruitful for SCP design

The discussion of the SCP policies and the existing evaluation tools above has shown that many of the pitfalls that were discussed in part two are prominent here too. As the MCCBA approach was designed to help evaluation in the sustainability context, the question arises: could MCCBA be useful in this setting too? To be more specific, the first question to address is whether and how the MCCBA approach to evaluation may help in giving a more specific direction to SCP policies than the broad range of possible activities seen above; could MCCBA help in designing more focused SCP policies?³⁰

The MCCBA evaluation approach is thus used for *designing* the SCP policy of business. At first glance this may seem somewhat awkward. However, thinking it through more thoroughly it may be clear that there are many links between design and evaluation that may warrant this procedure. First of all, the initial stages of an evaluation involve design questions, for instance when it is analysed what would be proper project alternatives. Furthermore, evaluation as we have seen, is preferably a reiterative process which involves reconsideration of project alternatives when evaluation outcomes have been analysed (and outcomes were already discussed when we looked at DJSI and other outcomes). Finally, the discussion above was one in which many evaluation results were overseen. This (meta) type of analysis is fitting and often used for improving policy design.

In the context of evaluating the sustainable performance of businesses four elements of the MCCBA approach to evaluation may be seen as distinguishing it most from other approaches – and so seem to be most fruitful for a better SCP policy design. These elements are:

- the importance of analysis of economic causality (preventing sustainability)
- the stakeholder involvement aiming at broadness of perspectives

better conduct. They no longer should focus only on their own profit but should develop a different attitude with a keen eye for poor people and on degradation of the planet. However, the view taken in this thesis is such that motivation is not crucial; more important is the pressure of the situation. An illustration from a study in the 1970s is worth mentioning. An experiment was set up with 40 seminary students – “with the highest expressed standards of moral behaviour” (Brady, 1990, p. 14). They were put into a situation resembling that of the Good Samaritan. Students were asked to prepare a lecture where half was brought under pressure: they were told to hurry; the other half was told that they had ample time. En route to the lecture room, they encountered a man in distress. Only 16 students stopped to help the man, most of whom were from the group with plenty of time. (Brady, 1990; Jeurissen, 2000).

²⁹ As Drucker so eloquently phrased it: “...the proper ‘social responsibility’ of business is to [...] turn a social problem into economic opportunity and economic benefit, into productive capacity, into human competence, into well-paid jobs, and into wealth.” (Drucker, 1984).

³⁰ As aforementioned, this chapter will not show a real MCCBA evaluation, but rather explore further use of MCCBA, and therefore focus on the first stages of the approach. To some extent, stage four is relevant too, as the argument rests partly on probable physical impacts under a competitive market regime. Finally, the last stage – reconsidering project alternatives in the light of evaluation outcomes – is also relevant, as the analysis is focused on (reiterating until) finding alternatives that give substantial outcomes.

- the focus on the most important – core business related – criteria
- the necessity and logic of the global evaluation level

The importance of analysis of economic causality (preventing sustainability)

In its drive for profits, efficiency and growth, private businesses support economic development, but why does the situation prevent business from giving more attention to extreme poverty and environmental degradation? If one looks at the main pressure on business enterprise, we can identify that it is competition.

An ever-advancing division of labour and specialisation characterises the current economic system. This specialisation leads to many narrowly-defined core businesses. Business is becoming increasingly large scale³¹ and competes in a worldwide setting, which makes communication with stakeholders necessarily distant, that is, non-personal. Shareholders can easily switch to other shares; consumers may easily switch to other brands. Small differences in stock return or product price may be decisive for the choice of both groups, as these choices are commonly anonymous, their choice possibilities are usually plentiful, and the knowledge of production process and business details is often rather thin. Communication of a company with its stakeholders is by necessity severely limited in information content. Company image or brand image often plays an important role.

This point should not be taken as rejection of competition as such, but it simply points to the fact that a ‘benevolent’ manager of a product or company will have a difficult job if he or she wants to follow a more sustainable course of action within a context of sharp competition and distant communication.³²

Stakeholder involvement

Could communicative improvement of stakeholder relations settle this problem of distant communication? It is no coincidence that in the SCP debate the communicative improvement in stakeholder relations is regarded as crucial to success (Waddock and Graves, 1997b; Bendheim et al., 1998; Waddock and Smith, 2000). Above in section 9.2.2 it was noted that many SCP initiatives initiate ‘social audits’: a structured communication process with stakeholders about SCP.

However, in reality the possibilities of communicative enforcement of stakeholder relations at the enterprise level are very limited. At this low individual company level, one may wonder how much expertise the (indirect) stakeholders may have to add. Second, it is difficult to imagine how such social audits could ever become widespread. Most multinationals cannot draw all their individual stakeholders (repeatedly) into intensive communication, as their number is far too large.³³ This fact makes it logical to involve representatives of stakeholder groups, preferably with a well-known name to add authority to the accountability exercise.³⁴ For one particular company this process might be arranged, but if all large companies were to

³¹ Compare Sillanpää (1998, p. 97) discussing an analysis of the Worldwatch Institute from 1994, which showed that the world’s top 500 companies control 70% of world resources.

³² Which is not to say that in certain cases competition may not be completely absent – monopolistic markets – or competition may not have a non-price character. However, it may also be observed that in many cases where competition appears not to be on prices, it is because price is not a unique selling point; which is rather different than the price being non-relevant to competition.

³³ One might wonder whether real-life behaviour of stakeholders will change. Involvement in the stakeholder consultation process is involvement as citizen not as consumer or shareholder. As chapters 3 and 4 showed when discussing the dual preference structure, there may be a large gap between how people act in a concrete market choice situation and what people say the world – and that market within it – should look like.

³⁴ For the indirect stakeholders, consider Amnesty International, Greenpeace, Friends of the Earth, etc.

organise such processes, relevant NGOs and other stakeholder representatives would face enormous manpower problems.

In part 2 general problems of fruitful stakeholder involvement were discussed. In answer to these problems MCCBA settled for a minimal involvement, assuring broadness of views only. This approach seems to be useful in this specific situation too. More details of the form of stakeholder involvement will follow below.

Focus on the most important - core business related - criteria

Generally, the real possibilities³⁵ for making progress towards greater sustainability for a specific company will be closely connected, not to making profits as such, but to its core business, i.e., to what it is good at, and to its main activities. It is therefore extremely important that SCP evaluation be related to that fact; if not, it is problematical to prevent high overall company scores resultant of unsatisfactory performance on important – core business related – issues and satisfactory performance on a collection of relatively unimportant aspects.³⁶

The necessity and logic of the global evaluation level

Better communication with stakeholders can be crucial, but will only make a difference if it works on the key aspect preventing large steps towards sustainability: competitive pressure. Changing this pressure is not possible at the individual company level (see Box 9-1).

Box 9-1: Individual or market/sector action?

Using the analysis of economic causality it may become clear that many companies did not become involved in bad labour conditions or low environmental standards by accident. Quite often competition drove them to seek low or least cost production facilities. This same competition will now limit the possibilities for serious improvement of these conditions: an individual company driven to very low wage countries by competitive pressures will not be able to change much in the direction of higher wage without being out-competed by other firms; the same holds for environmental standards. This situation can only change if the market situation changes. With fixed market conditions the only other route open for achieving more than very small changes is apparently working on actions at the sector or market level.³⁷

For non-coordinated individual firm action higher performance on environmental or extreme poverty issues may lead to a gain in profitability (recall the SOS gourmet pan of section 9.2.2.2c) or to a loss of profitability. Generally, at the individual firm level the relation is: the higher the performance the lower the profitability becomes; but there is wide variance

³⁵ And ... it might be added: the expectations of stakeholders.

³⁶ Compare Doane, 2000 and 2002 on reporting on social and environmental issues. For instance, (Doane, 2000, p. 9) quotes from The Global Reporters (Sustainability, Nov. 2000): "...the Top 50 reports fail to address what we consider to be the biggest sustainability issues associated with a company's activities." The reader may also turn to the discussion of the results of the 100 Best Corporate Citizens of 2004 (at www.business-ethics.com) which – interestingly – does not highlight the companies with a high overall score, but instead begins by discussing companies with a top performance on sub-scores (their overall ranking was 'only' 6, 32 and 58 respectively).

³⁷ Because the macro level route in which a global government could easily arrange preconditions is, as we have said –as yet – no option.

possible, depending on the specifics of the market and the position of different companies therein.³⁸ As long as a substantial share of the companies is not involved in environmental improvement and poverty reduction, competition will limit the overall improvement. For collective/coordinated action the relationship is different; higher performance on improving environmental degradation and extreme poverty will lead to substantial costs, but there is then no competition – within the sector – on price; consumer prices may simply rise. The overall outcome on sector profitability and turnover is difficult to predict. There might still be a somewhat lower profitability but the potential gains – for the environmental improvement and poverty reduction – are much higher than in the case of individual behaviour.³⁹

Directly linked to this point is the fact that the existing SCP evaluation tools give a (limited) perspective on the performance of individual companies, but they seem to give little perspective on sustainable performance of the whole group of businesses within a sector or market.⁴⁰ Sustainability or lack of it occurs as a system outcome; it is a judgement about overall performance at a high spatial/organisational level. Many sustainability problems only manifest on a higher scale level than that of individual organisations. If all individual businesses are actively friendly to their customers and try to satisfy both their active and latent wishes, then this may be seen as very sustainable at an individual level, but the consumption level at a market can simply be too high. One example is the ever-increasing energy use of automobiles in the face of drastically ever-more efficient engines.⁴¹ Useful sustainability measurement at the individual level then seems to require a clear link to overall measurement and evaluation outcomes. As Sutton (1998, p. 129) expressed it: “Even for very well-run companies, we cannot possibly know if they are ecologically or socially sustainable until we know something about the society or environment in which they and their products are embedded.”

Government intervention could arrange changing the rules, and it is commonly seen as the traditional role of government to define and guard the preconditions for the market system and competition. However, in the globalised market system there is – as yet – no government that can play this role. The ‘good news’ of the current globalised system and the large size of multinationals seems to be that there may be another route to changing competition rules besides having a globally effective government. The business sector could arrange this task itself; it could be arranged/initiated by the (largest) companies active in a market or within a sector.

³⁸ As we observed above, *in certain cases* competition may be completely absent or competition may have a non-price character.

³⁹ For instance, in the light of a to-be-expected factor 10 increase in energy and material flows, an improvement of eco-efficiency by 90% would be necessary, only to accommodate material welfare and population growth up to 2050 (Sillanpää, 1998).

⁴⁰ Compare Hibbit (2004) whose main critique on existing sustainability evaluation and accounting is seemingly that the real system-level questions ignored. Compare also Murray et al., 2000 (p. 236) who argue for the importance of overall/sectoral assessments in the area of population health. Barbier et al. (1990) also support this point as in their analysis a clear view on the portfolio of possible actions is crucial. Compare further Hill (1999) for arguments to include macro-economic considerations in CBA-based project appraisal.

⁴¹ Reasons being: larger average size of automobiles, more frequent use, greater number of cars, more standardised air-conditioning, etc.

What this analysis of causes preventing sustainability shows is that SCP policies should concentrate on changing the rules of the game at the (global) market or sectoral level. Changing the rules of the game is about changing the *preconditions* to competition and not competition altogether. At the individual company level SCP policy should be about contributing to this aim.

Changing the rules of the game of global competition may sound somewhat utopian but in fact it is not; reorganising competition at the sectoral or market level is an everyday reality. A historically unique feature of the contemporary economic system is that it has already globalised, certainly at a sectoral level. From a company perspective the markets and sectors have become global villages with recognisable actors. Whether one looks at the sector/markets for oil, audio systems or automobiles, the number of players is limited. Usually with a number of companies of somewhere between 3 and 40 (own estimate⁴²), the largest share of the market will be covered. Moreover, these companies know each other. Very often there is a branch organisation and the likelihood is that the companies meet regularly to debate issues of mutual interest. This consultation may lead to joint ventures, strategic alliances, branch-wide action, etc.

The financial-economic pages of newspapers in print at the time of the study (end 2001) allowed one to make a small list of cooperation initiatives between competitors in the global village:

- Air France formed the SkyTeam Cargo Alliance together with Delta Air Lines and Korean Air – worldwide one of the largest alliances in the field of air-cargo (mark the geographical spread of the partners).
- The biggest job site on the internet Monsterboard is a company of TMP Worldwide that has an alliance with many different companies in every country. In the Netherlands the alliance included among others: BNN, Lycos, IDG Webwereld, Ilse, Freemail, CompuServe, WorldOnline (mark the number of competitors cooperating).
- Philips and Sony cooperate on fundamental research for the new generation of audio media. This partnership prevents the conducting of very expensive research to a standard that (for one of the partners) eventually will fail (mark the benefit to society, or the prevention of loss through cooperation).
- Automobile producers Ford, DaimlerChrysler, General Motors, Renault, Nissan, and Peugeot-Citroën initiate a business-to-business electronic marketplace to achieve more efficiency in procurement (mark the combined market power and again the geographical spread).⁴³

A last example of sectoral cooperation worth mentioning, because it influences the amount of

⁴² Giving more exact estimates, and even giving supporting evidence may be rather difficult however. First of all due to a lack of global data and second to definition aspects; how do sectors relate to markets? Calculating concentration ratio's (like the Hirschman-Herfindahl Index) basically requires defining a product market and a geographical market first. Assessment of concentration for a specific market requires a mixed qualitative and quantitative assessment (Federal Trade Commission, 1996).

Still a combination of facts points in the direction of the figures given. As mentioned above in Sillanpää (1998) the world's top 500 companies control 70% of world resources. Furthermore Franko (2004) sees a strong focussing trend – away from diversification - in industrial firms in a great many sectors. Finally, sectors are and can easily be so defined as to be small in numbers; markets can be easily defined in infinitely small parts.

⁴³ This type of agreement is the area for anti-trust authorities (e.g. EU regulation is published in EC 1979/C 372/03 and EC 2004/C 31/03). Discussions of this aspect would require an inappropriate digression. Generally speaking, it is clear that agreements about innovations are often allowed and that 'hard core' agreements –on prices or market shares are forbidden (see <http://europa.eu.int/comm/compe.pdf>) for a report on European judgements on this type of agreement. See further Federal Trade Commission (1996) for a discussion of standards (and networks) and their position in securing in fair competition.

competition directly, is the soccer sector. Through the FIFA-organisation, professional soccer clubs agree on the best form and intensity of competition; play schemes, division of TV rights, etc. Individual clubs have many aspects to attend to and there is enough competition, but the preconditions are explicitly determined among themselves.

The precise content of these agreements is not very important here. The point is that in the global village of a specific sector or market the competitors know each other and agree to direct or limit competition. This state of affairs can also spread to sustainability issues (see the appendix to this chapter for examples). Initiatives for more sustainability aiming at improvements at the sectoral – or market – level are termed sector-SCP.

9.2.4.2 More important elements of SCP policy design

We have seen that sustainability can be improved to some extent within the normal market performance, but that the larger sustainability gains within a reasonably short time span – which several stakeholders deem necessary in the areas of poverty and environment – will only be reached if initiatives try to diminish the pressure of competition.

SCP will therefore require sector initiatives; action has to be organised by the (leading) companies in a sector.⁴⁴ This higher level process is required to avoid the counteractive pressure of competition; and so high ambitions – with high cost levels – are not insurmountable from the beginning, as competitors encounter these costs too (compare box 9-1). The SCP evaluation of an individual company could then focus on monitoring individual contributions towards these sector initiatives.

By taking inspiration directly from stage 2 and stage 6 of the MCCBA approach, it is clear that sector initiatives or initiatives to ‘change the rules of the competitive game’ are processes which require two further elements for success.⁴⁵

- set up a stakeholder consultation at the sectoral level and
- a limitation of the number of issues.

A ‘sustainability audit’ at the sectoral level

As stated above, the sector should organise a process to change certain rules of the game and arrange sufficient commitment within the sector or market.

Following an MCCBA approach this process should be framed within a stakeholder consultation at the sector/market level. Such a sustainability audit with stakeholder involvement can co-determine the issues and the pace in which they are realised.⁴⁶ Concerning the involved stakeholders here again representation is less important than broadness of perspectives.

An audit at the sectoral level has important advantages compared to an audit at the individual business level, as the manpower problem of the NGOs and other stakeholders is solved, and a serious role concerning the content of the matters at hand is indeed possible. Beyond that, the process is simply efficient, because many individual companies in one sector or market face the same sustainability problems and have comparable possibilities.

⁴⁴ The appendix to this chapter provides examples.

⁴⁵ Noteworthy is that the first is explicitly linked to the stakeholder involvement stage of MCCBA evaluation (stage 2) and the second is closely related to specification of the functional limitations of the evaluation (stage 1).

⁴⁶ Compare Sillanpää (1998) for an outline of a stakeholder audit process – at the company level.

Limitation of the number of issues

The identification of issues is a process of negotiation and analysis in which several forces are active. The process should address the urgency of the problems, the possibilities and expertise of the companies to contribute something concrete within a reasonably short time span, and the connection of the sector or market with the problems (Jeurissen, 2000). The pace at which the issues are tackled, the way they are tackled and how to monitor progress is part of this ongoing process of attuning to the aims and wishes of society as a whole.

An important further element of designing sector SCP with the help of MCCBA is to sharpen the results of the stakeholder audit in advance. The issues to be potentially tackled have to be (very) limited in number and have to relate to the core business of the companies involved.⁴⁷

These demands are relevant for two reasons. First, if the point is to try and use the power of private enterprise, one has to stay close to what they know. No actions should be requested that are better asked of consumers or shareholders, governments or other companies. Second, companies do not accidentally have a 'core business.' Businesses are aimed at a small number of things; they have a core competence, which is narrow and specialised, concrete and specific.⁴⁸

9.3 Exploring a concrete example: Evaluating EurepGAP⁴⁹

9.3.1 Introduction

At the beginning of this chapter we observed that two closely connected aspects were challenging to the Ahold food retail Holding concerning their SCP policy. First was the confusion about what SCP was, and second was the lack of standardisation in evaluative measurement of SCP. Evaluation of SCP using MCCBA would not follow a DJSI or BSE100 approach of assessing many different aspects at an individual company level; instead it would focus on the sector level sustainability projects.

In the appendix to this chapter three examples are identified of sector-SCP projects that change the rules of the game of competition in the direction towards more sustainable behaviour; EurepGAP is one of these.

EurepGAP was launched in 1997 by a group of 26 leading European food retailers. It was designed to assure product safety, reduced agrochemical use, environmental protection and animal- and worker welfare. EurepGAP is an initiative aimed at the sector level, although the entire sector did not participate from the start. EurepGAP explicitly aims at greater sustainability and food safety. The Euro-retailer Produce Working Group (EUREP) wanted a standard for food safety that was closely connected to sustainability and which had the following characteristics:

- no double standard, which might give differences in safety around the world;

⁴⁷ Compare Sutton (1998) who argues that sustainability must refer to maintaining or restoring something specific.

⁴⁸ Given the scope of its activities one company may be involved in a number of narrow sector/market SCP initiatives.

⁴⁹ The results of this section are based on research on EurepGAP and the apple market, which students Anna Kussyk (2003) and Krista Koeleman (2004) conducted as internship research and master thesis research respectively – under the supervision of the current author. Furthermore, some results are used from Sybren de Jong's master thesis, written in 2005 on the position of Kenya and Senegal under EurepGAP.

- taking into account the Triple Bottom Line of people, planet and profit, which may reassure investors;
- creating supply chain reliability, which guarantees food safety in all parts of the supply chains.

EurepGAP in use now is mainly a framework for Good Agricultural Practice; it is not a consumer brand, but is instead a number of rules between supplier and retailer. In the EurepGAP documents several topics are treated: traceability, soil management, pesticides, health and safety of employees, etc. and for all topics ‘musts’ and ‘shoulds’ are defined. At first EurepGAP aimed for a superior quality standard, but many suppliers would not at that time have been able to fulfill the obligations. The weakened result of ‘musts and shoulds’ reveals the final concessions. Although the ‘musts’ are hard demands, they are often a repetition of (existing) lawful demands, and the ‘shoulds’ are mere encouragements. This section will focus on EurepGAP and is not aimed at showing results of a real MCCBA evaluation. The aim is to give a more concrete illustration of evaluating an SCP project following all the stages of MCCBA. The guideline in this illustration will be the eight stages of MCCBA. Logically, given the explorative and qualitative style of the discussion, not all the stages will be ‘filled’ to the same extent.

9.3.2 Stage one: Identify function, project alternatives and scale of the evaluation

The function of an evaluation of EurepGAP in the setting of a sustainability audit would naturally focus on efforts made and results achieved by the sector and one or several companies within it. It stands to reason that some agreement is achieved about either the yearly efforts or the yearly results.

As for the project alternatives, MCCBA would, like any serious evaluation, try to evaluate EurepGAP against the option of not having it: a ‘with’ and ‘without’ situation. The evaluation might consider more alternatives that specify different contents of EurepGAP: more or less environmental demands, more or less worker conditions demands, etc.

MCCBA would use different scale levels in the evaluation, one of which is the global level.

As European retailers set up the EurepGAP initiative, a logical level of a non-MCCBA approach would seem to be to choose the EU, as this is the primary target region where EurepGAP should be working, at least from a consumer perspective. However, the global level is standard in MCCBA, and in this particular case this level is very explicitly relevant as well, since many of the EurepGAP suppliers are outside the EU and EurepGAP is trying to develop a global standard. In the evaluation the EU level might then still be chosen as a level of specific interest.

Furthermore, this SCP evaluation project began as a project for the Ahold Company. In such a case the evaluation would have a specific interest in the performance of Ahold (and its main offspring operating in the Dutch market: Albert Heijn) and so this should be a scale level too. Such a low level can also be seen as part of a distribution analysis because many of the benefits logically fall outside this level: but it may give a clear insight into the SCP effort of Ahold.

9.3.3 Stage two: Involve a broad group of stakeholders

MCCBA would require broadness of views in the stakeholder audit. In the EurepGAP activities the stakeholders most involved – apart from retailers – are suppliers. Given the many

close economic ties between retailers and suppliers this involvement cannot be considered as broad. Determining which stakeholder groups to involve further seems to require a notion about the impacts and working of the EurepGAP scheme.

Concerning the indirect stakeholders it might be logical to at least involve an environmental NGO, as EurepGAP claims to have a positive impact on the environment. Furthermore, a development organisation such as Oxfam or perhaps the World Bank (compare EurepGAP, Newsletters 2003, 2004) might be considered to assure broadness of views concerning poverty impacts – or impacts on poor primary producers. Since EurepGAP partly overlaps with other food safety and quality assurance schemes, stakeholders from ‘competing’ schemes might be involved as well. Some EU representative might be involved because the relation to EU regulation on, for instance, minimum residue levels is very relevant, or a representative from the Codex Alimentarius⁵⁰ could also be considered.

9.3.4 Stage three: Organise judgement criteria on Triple E impacts

In this illustrative exercise it is difficult to generate a list of possible impacts and organise them properly: applying the standard criteria structure seems fair enough. EurepGAP involves many impacts closely related to economic market transactions and which can probably be covered under the heading of economic development impacts. Monetarized costs and benefits will at least partly be feasible for these impacts. Furthermore, considering the formal aims of EurepGAP there may be some impact on diminishing environmental degradation and improving worker safety; the place of the worker safety criterion in the criteria structure below is not a priori clear. Finally, one may find some impact on extreme poverty, as growers and suppliers come from all over the world, including low income countries.

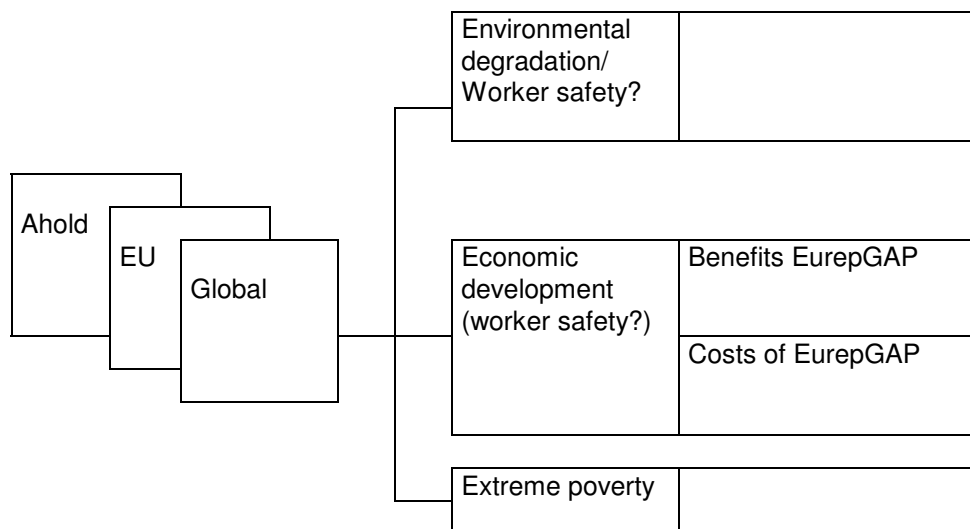


Figure 9-1: The preliminary criteria structure of the EurepGAP evaluation

⁵⁰ This is an international program initiated by the FAO and WHO to protect consumers and promote fair trade in food products.

9.3.5 Stage four: *Quantify impacts physically*

As shown in earlier chapters, stage four of MCCBA is often the core of the evaluation, in the sense that new information is gathered about the impacts during this stage. Although nothing will be presented below that matches the real evaluation results expected in an actual MCCBA application, some tentative information will be given.⁵¹ To some extent the value of this information lies in increased understanding of the economic logic within and underpinning EurepGAP, and the costs and benefits to different parties.

The basis of this analysis is comprised of three explorative studies; two mostly concern the apple market (Kusyk, 2003 and Koeleman, 2004)⁵² while the third concerns the position of less developed countries (De Jong, 2006).

Global impacts

A survey was conducted in the first study among apple growers, wholesalers and branch organisations worldwide. Respondents were asked to give their views on the possible impacts of EurepGAP, and both EurepGAP-certified and non-certified were surveyed. Table 9-6 gives a short summary of the results.

First, given EurepGAP's Triple P ambitions, it is remarkable that the environmental impact according to most respondents is absent or – according to some – quite limited.

Second, the table shows something about the division of costs and benefits. Most respondents see no gain in market share due to the adoption of EurepGAP: it mainly functions to keep their existing market share. Neither do they expect benefits in terms of higher prices or returns. Note that the initiators were not surveyed, but rather only the suppliers and the food retailers. Despite their lack of extra benefits, suppliers do make costs; basically there are two cost components to suppliers: extra investments and extra time costs for registration.⁵³

⁵¹ Clearly EurepGAP will need time to try and reach its aims and various degrees of success are feasible. For instance, the demands now labeled as 'shoulds' to the EurepGAP certified suppliers could, over time, be turned into 'musts'. MCCBA evaluation might in a real evaluation use scenarios in which the future requirements of EurepGAP and its implementation rate would vary among retailers and producers.

⁵² The apple market, it should be noted, is only one of the markets covered by EurepGAP. The Kusyk study focused more on worldwide impacts, the Koeleman study more on economic relations and motivations.

⁵³ The investments can be made for pesticide storage and warehouses, equipment for pesticides and soil analysis, restrooms in the fields, safety equipment and information and warning signs. Some of the respondents say no investments are needed, while others estimate it to be some 0.1% of their turnover up to 0.8%.

Table 9-6 Summary of views on the impact of EurepGAP

Responding group -Replies from	Major motives for implementation	Major views on economic impact	Major views on environmental impact
Growers (4) -Italy (2) -New Zealand (1) -United States (1)	- Protect existing market share (fulfilling customers' demands) - Raise the standards of professionalism among market partners - Make production environmentally safe	- Cost increase (new investments in equipment and materials, extra auditing, time for the first registration) - No increase in benefit	- No impact
Wholesalers (6) -France (3) -New Zealand (1) -South Africa (1) -United States (1)	- Protect existing market share - Achieve growth in a market share - Raise the standards of professionalism among market partners	- Cost increase (new investments in equipment and materials, extra auditing, time for the first registration) - No significant increase in benefit	- No impact on environmental protection - Small impact on pesticide use
Branch organisations (12) -Chile (3) -France (2) -New Zealand (1) -South Africa (2) -United Kingdom (2) -United States (2)	- Protect existing market share - Comply with retailer demands	- Cost increase for equipment (pesticides storage and warehouses, pesticides and soil analysis, restrooms in the fields, signs e.g., information, warning, danger, etc.) - Extra auditing, time for the first registration - No increase in benefit	- No impact

Source: Kusyk 2003

First, given EurepGAP's Triple P ambitions, it is remarkable that the environmental impact according to most respondents is absent or – according to some – quite limited.

Second, the table shows something about the division of costs and benefits. Most respondents see no gain in market share due to the adoption of EurepGAP: it mainly functions to keep their existing market share. Neither do they expect benefits in terms of higher prices or returns. Note that the initiators were not surveyed, but rather only the suppliers and the food retailers. Despite their lack of extra benefits, suppliers do make costs; basically there are two cost components to suppliers: extra investments and extra time costs for registration.⁵⁴

The size of the costs that have been made – or that need to be made – differ between respondents and seems to depend on the state of affairs in their company before EurepGAP. If the suppliers had a well-developed and registered quality scheme, if they were already keeping reasonable environmental standards and good working conditions, then costs are sometimes quite small or even negligible. If the opposite is true they may incur substantial costs. Whether

⁵⁴ The investments can be made for pesticide storage and warehouses, equipment for pesticides and soil analysis, restrooms in the fields, safety equipment and information and warning signs. Some of the respondents say no investments are needed, while others estimate it to be some 0.1% of their turnover up to 0.8%.

they are willing to pay those costs and become certified under EurepGAP depends on the demands of their main customers. The closer they are to EurepGAP initiators, the more they simply conform. For those suppliers that are economically more distant or to those that only consider entering the market, EurepGAP presents an entry barrier to – part of – the market.

This point is supported by the research of De Jong (2006), which shows that the chances of EurepGAP being beneficial to less (or least) developed countries are small. These chances were estimated at a country level by using a relative – in comparison with competitors – competitiveness index. This index consists of proxy indicators for the demands of EurepGAP (e.g., literacy rate as one indicator for tracking and tracing demands, general pesticide use as indicator for the environmental standards, etc.). The analysis included an investigation of investment capabilities to overcome deficiencies. The empirical analysis focused on two case studies: first, Kenya and tea, second, Senegal and ground nuts. Overall, the analysis showed that EurepGAP can easily have detrimental impacts on less developed countries.

The Kusiak results showed another important aspect of EurepGAP, namely its relation and possible overlap with other certification schemes. In the sector several other schemes exist including Integrated Fruit Production (IFP), Hazard Analysis and Critical Control Point (HACCP), ISO standards, etc. To the farmer they often strongly overlap in their concrete demands, although they might come from different sources. The question naturally arises: why introduce a new standard? Why new ‘paper work’ while there is hardly any real change in the production process?

The position of the food retailers

This question brings us to the position of the initiators of EurepGAP: food retailers. What were their motives for initiating the EurepGAP standard? The second study on EurepGAP conducted interviews with one of the initiators of EurepGAP, the Dutch Albert Heijn supermarket chain and a number of its apple suppliers (Koeleman, 2004). The study showed that EUREP retailers were less interested in sustainability than in increased traceability. On occasion, fresh products do not meet minimum food safety standards. For instance, in recent years in the Netherlands some pears and grapes contained too many pesticides. The crisis led to recalls of the product – often among all retailers – as it is difficult to identify exactly which pears or grapes are contaminated. Recalls are commonly costly operations and difficulties arise about who should pay these costs: retailer, supplier, wholesaler?

For the food retailers extra traceability of fresh produce caused by the EurepGAP protocol has two advantages. First, they may not have to recall their products, because due to their EurepGAP administration they can show the authorities that their fresh products are free from contamination or remain below a minimum residue level. Second, the traceability within EurepGAP gives far better possibilities for defining liability in the event of a genuine problem.

The EurepGAP protocol is preferable above other schemes from the perspective of food retailers. EurepGAP is their own standard dedicated to their own demands, and is allegedly more convincing to consumers than farmer-initiated quality schemes (Koeleman, 2004).

A critical point for EurepGAP’s success is the adoption of the standard by the different suppliers of fresh products. The food retailers as initiators of the standard are in a strong position as they are mostly market leaders. Because of the dependence of many of their suppliers, the retailers managed to introduce EurepGAP, while the main part of the costs was

borne by suppliers and the main benefits fall to retailers.⁵⁵ The interviews showed that ultimately it was the consumer demand for safe food that triggered EurepGAP (Koeleman, 2004). The food-safety part of the consumer preferences seems much stronger than the environmental and social issue part. Consumers expect their food to be safe. The retailers experienced this – especially during a crisis – and through EurepGAP they devised a better way to guarantee it. Environmental NGOs and the media strengthened this consumer demand.

Exploration of overall results

To maintain the EurepGAP standard, wide adoption among producers is preferred. In recent years EurepGAP has increasingly applied benchmarking of other quality schemes against EurepGAP. If another scheme meets the same standards it can acquire EurepGAP recognition. This policy has strongly promoted EurepGAP's position, which by 2004 includes 10,000 suppliers of fresh produce from 35 countries (EurepGAP, Newsletter, 2004).

Much more could be said about the details of EurepGAP and the positions of various parties, but for illustrative purposes here, this short analysis of economic causality and the costs and benefits to different parties may suffice. Table 9.2 shows these results.

Table 9-7 Explorative results of the impacts of EurepGAP

Physical impacts of EurepGAP – explorative assessment				
	Global level	EU level	Ahold	Distribution matters:
Tracking and tracing	+++	+++	+++	Benefits of more tracking and tracing fall to retailers, costs fall to suppliers
Environmental performance	0 (0/+)	0	0	Zero or small benefits to everyone Costs fall to suppliers
Working conditions	0/+	0	0	Benefits and costs to suppliers
Poverty reduction	-	0	0	Producers from less developed countries have a higher chance of not being able to fulfill EurepGAP demands – New producers face a higher entry barrier

Source: Own estimations.

Table 9-7 and the analysis thus far show that sustainability is not very relevant to the main actors. The food safety issue mainly triggers the elements of environmental concern in EurepGAP. The sustainability results in terms of increased environmental performance, for instance, are limited. Given the fact that there was an initial intention on the part of the food retailers to work towards a higher sustainability standard, it is worth asking whether more formal involvement of NGOs in a sustainability audit (as preferred in MCCBA) might have triggered higher end results. Undoubtedly however, the acceptance rate by suppliers would have been lower or slower.

Be that as it may, without understanding market related causality, the point that this separate standard overlaps with many existing initiatives might have received far less attention. The

⁵⁵ The EurepGAP protocol also requires time for registration. For the first time, registration between 15 and 30 days seems to be rather common. Once certified, the annual auditing may then require a couple of days per year.

remarkable fact that initiators of the protocol benefit most, and others bear the costs, would also not have been discovered. Instead of EurepGAP being a burden – or posing a substantial effort – to the initiators it seems directly beneficial to them. Without such an understanding, one may – like the EU Commissioner – easily praise EurepGAP members for their “voluntary measures assisting safe and sustainable agriculture” (EurepGAP press statement, 28–10-2003). In a real MCCBA evaluation focusing on division of costs and benefits, this economic causality-based outcome would almost certainly prove to be invaluable judgement information to stakeholders. The lack of improvement in environmental performance thus far – especially relative to existing standards – might be important too.

9.3.6 Stage five: Aggregate monetary scores consensus based

Given the general character of this illustration, the main point to be highlighted at this stage is that a CBA component is helpful in capturing some of the impacts.

At the level of Ahold (or Albert Heijn) costs and benefits of increased traceability seem best captured in a CBA. Historical estimates of the costs of crisis can be made and improved possibilities for less damage⁵⁶ in the next, say, 15 years can be estimated – compared to a specified nul-alternative. This procedure would give a reasonable estimate of the benefits to Ahold (or Albert Heijn) of its participation. On the basis of time and money spent on the initiative, the costs to Albert Heijn can also be estimated.⁵⁷ Of these cost and benefits a net present value can easily be calculated. In terms of judgement-oriented information the use of CBA here seems to be quite ideal.

At the global and EU levels the matter becomes complicated, in that there is more choice and judgement involved as to whether impacts are best treated within the monetary CBA part or in the MCA part. Naturally these choices depend on data availability. On the cost side the investment costs, registration time etc. are relevant. Monetary valuation seems to be suitable for this. Other costs may be increased poverty due to the extra entry barriers hampering growth of farming activity in underdeveloped regions.⁵⁸ The impacts on the benefit side concern possible environmental improvements and more importantly increased food safety, either in terms of objectively reduced risk, or in terms of perceived reduced risk. For assessing the more subjective feelings about extra food safety provided by EurepGAP, the size of this impact depends heavily on the extra safety of EurepGAP compared to other standards. For the measurement of the impact on food safety there are several options; one obvious way is to directly assess the impact of safer food for human health. Less death, illness or sickness of consumers may theoretically be the impacts to be assessed. Monetary valuation of these health effects may, however, be problematical; treatment in the MCA part might then be a better option.

The actual assessment of these health impacts may nevertheless also pose difficulties. The causal link between a certain residue level and its health impact is often far from straightforward, let alone measurable. One could take a more indirect approach and rather than estimate the health damage as such, estimate the efforts of the food inspection services. Obviously this approach would be second-best, but it would lend itself more easily to

⁵⁶ Even increased turnover is a possibility when the adoption rate by others is low, and trust in Albert Heijn's quality has increased due to its lower vulnerability in case of crisis.

⁵⁷ On the basis of the preliminary results of stage 5 it seems clear that the monetized net benefits to Ahold/ (Albert Heijn) would be substantially positive.

⁵⁸ See Newsletter, 2004 for an initiative to try and facilitate Senegalese growers to comply.

monetary measurement.

9.3.7 Stage six: Aggregate non-monetary scores consensus based

An important benefit at a global or EU level would be in the area of increased health level of consumers; either objective changes or more subjective changes in feelings about this level. Non-monetized measurement of objective changes might be represented by natural units such as number of deaths, or number of hospital days. One might also opt for an MCA-based construct such as Quality Adjusted Life Years (QALY) that has some intuitive appeal and which aggregates a broad set of health impacts. With regard to the more subjective changes, several techniques are available.

9.3.8 Stage seven/Stage eight: Interpret trade-offs/perform sensitivity analysis and reconsider project alternatives

These stages are logically somewhat superfluous in the current exercise, which has mainly focused on the structure of the evaluation and hardly given real results.

9.4 Conclusions

Ahold's challenges

At the beginning of this chapter the discussion centered on two closely related aspects, which posed challenges to the Ahold food retail Holding concerning their SCP policy. First, the confusion about what SCP was, and second, the lack of standardisation in evaluative measurement of SCP.

The MCCBA approach to evaluation was used here, both for answering the design question of what SCP policy should be, and for exploring what a concrete and somewhat standardised SCP evaluation might look like.

MCCBA evaluation elements were used in order to define SCP and detect where existing SCP evaluation tools fail. SCP policy should focus on one or two sector-SCP issues, such as the introduction of EurepGAP, as part of a sustainability audit with stakeholders. However, both type of project and stakeholder audit setting will differ among different companies (unless they are in the same sector or market).

As a standard approach to any evaluation of an SCP policy, MCCBA can also be applied. The example of EurepGAP showed how such an evaluation could turn out; and generally for the cost side, monetized (discounted) costs will be helpful. For benefits, both monetary and non-monetary criteria may prove to be the best choice, but choice will depend on the specifics of each evaluation.

Gains from the MCCBA approach

Making use of the insight into (market related) economic causality and the need for higher level evaluation of sustainability has led to a simpler picture of the evaluation of SCP performance. Evaluation of SCP is now refined such that project evaluation can be straightforward. A gain from the MCCBA approach was that understanding economic

causality helped identify the sectoral-level as an important level for stakeholder involvement. This use of causal economic analysis and the emphasis on high level evaluation has overcome a main weakness of MCA-like evaluation in the sustainability context: long lists of criteria with little weight that try to cover the widest possible number of issues. The MCCBA approach is much more focused.

Finally, the analysis in this chapter has attempted to show the usefulness of taking monetarized costs and benefits into account, since a correct perspective of the effort of both a firm and a sector is problematical without this element.

Appendix to chapter 9: Some examples of sector-level SCP projects

To illustrate the type of projects that can be evaluated when designing sector level SCP policies using MCCBA evaluation insights, three initiatives are discussed below: EurepGAP, Global Alliance and Dutch covenants.

EurepGAP

EurepGAP was launched in 1997 by a group of leading European food retailers. It was designed to assure product safety, reduced agrochemical use, environmental protection and animal- and worker welfare. EurepGAP is an initiative aimed at the sector level. EurepGAP explicitly aims at greater sustainability and food safety. (see section 9.3 for more details)

*Global Alliance*⁵⁹

Global Alliance for Workers and Communities was initiated by the companies Nike and Gap, by the International Youth Foundation and the World Bank, by St. John's University, and by the government of the state of Pennsylvania. The mission of the Global Alliance is to give workers worldwide more chances to develop their potential. How does this work in practice? Global Alliance works as a bottom-up process. Which issues are important for which stakeholder groups and what can be improved? Instead of working with a compliance standard,⁶⁰ Global Alliance works with initiatives that can take various forms.

Global Alliance works in two phases. In the first phase employees are asked what they think of their work climate, what they expect of the future and what they think is needed for their environment. Specially trained project teams hold the interviews in close cooperation with local unions, active NGOs⁶¹ and scientific experts (from one of the Global Alliance partner universities). After a first analysis of context, a round with face-to-face interviews and a round of working groups, a report is made of key issues. The second phase sets up specific help programs and a list of six priorities is made. A training program is formulated with clear time-lines. Efficiency, scale, sustainability and positive long-term impacts are the important components of the programs.

Although Global Alliance works as a bottom-up process and yields rather open outcomes, it has several links to compliance standards. Before a production site can opt for an intensive Global Alliance program, it has to comply to national legislation, a code of conduct and conventions of the ILO and UN. Furthermore, the bottom-up process often leads to initiatives of living up to other compliance standards.

*Dutch covenant on energy efficiency*⁶²

A special form of partnership much used in the Netherlands in recent years is known as a covenant. Private enterprises 'voluntarily' sign a contract – often to prevent new requirements being forced by new laws. Generally, these agreements are signed by branch organisations and governments at different levels (local, regional and state) but many variants exist. Sometimes NGOs are co-signers. Several covenants could be mentioned that relate to sustainability, including the covenants on sustainable construction and waste reduction. Another example is the Covenant Benchmarking Energy Efficiency, which aims at making Dutch industry join the world's top in energy efficiency in 2012. Many partners signed, among those were several

⁵⁹ www.theglobalalliance.org

⁶⁰ See Broadhurst (2000).

⁶¹ See for instance: www.oxfam.org.au/campaigns/nike/dialogue/

⁶² See website: www.benchmarking-energie.nl

industrial branch organisations and two ministries. In 2002 104 companies joined representing 234 industrial installations. Their goal is to reduce their energy use in 2012 by 82.000 TJ, which gives a 5,7 million reduction of CO₂ emissions.

PART FOUR – CONCLUSIONS AND FINAL REFLECTIONS

THE MERITS OF THE MCCBA APPROACH

CHAPTER 10: CONCLUSIONS AND FINAL REFLECTIONS

10.1 Introduction

In chapter 5 the MCCBA approach was formulated; the formulation at that point was predictably general and somewhat abstract. After having assessed the specific evaluation cases of chapters 7, 8 and 9, the merits and limitations of MCCBA as a tool for accountability enhancing evaluation in the sustainability context can be highlighted from a more comprehensive perspective. This chapter draws conclusions from this more complete perspective (10.2) and will subsequently give final reflections on important issues and identify topics for further research (10.3).

10.2 Additional conclusions (additional to conclusions of chapter 5)

Below we will discuss the most significant conclusions from the case study per stage. It should be emphasised that what is judged here as important are only those conclusions or aspects that *add to* or *deviate from* the description of MCCBA in chapter 5.

Stage one: Identify function, project alternatives and scale of the evaluation

In stage one the most important key word is acceptance, acceptance of a change in function - from decision-making to decision-aiding, and acceptance of the different – often higher geographical or organizational - scales of the evaluation as being relevant. This acceptance need not be ‘automatic’; stakeholders may have to be persuaded to accept certain points and success is not guaranteed. In the EHS evaluation, the change of function was accepted as it still covered the original function, while the new elements had a chance of being beneficial to initiator. In the Emssperrwerk evaluation the global evaluation level was accepted because of its underlying rationale and the global economic scale of the market.

Stage two: Involve a broad group of stakeholders

As for the involvement of stakeholders in *stage two*, despite the ‘solution’ proposed in chapter 5 of merely focusing on a broad group of stakeholders - incorporating a spectrum of views - , the case studies show that even this type of involvement may remain difficult for various reasons. The EHS evaluation showed that stakeholder involvement was beneficial for measurement consensus and that their involvement may create its own dynamics as well as generate unforeseen impacts on the evaluation’s content. Involvement of the main losing stakeholder in the evaluation (agriculture) proved no guarantee for commitment to results. Results from the Emssperrwerk evaluation showed that stakeholders with nothing to gain from the evaluation may be less willing to participate. Furthermore, this case showed that cultural and language barriers – even in neighbouring regions – are serious barriers. The SCP evaluation mainly illustrated that stakeholder involvement is only useful at the correct level (in this case the sectoral level).

Stage three: Organise judgement criteria on Triple E impacts

The Triple E structure, proposed as a standardised element of MCCBA in *stage three*, is only preliminary. If no significant impacts can be found, the structure should be changed and criteria skipped. In the EHS evaluation extreme poverty was irrelevant, and the criterion was therefore skipped. Furthermore, despite the fact that the financial burden to the government

would normally be seen as ‘double counting’ within a CBA, this important criterion was added.

With regard to the global level assessment needed in the MCCBA, the Emssperrwerk evaluation indicated that global level analysis is often not the primary focus. It therefore, requires existing knowledge and expertise of the evaluation team; otherwise, too much time is needed for impacts deemed relatively unimportant at the decision-making level.

Stage four: Quantify impacts physically

The case studies clearly acknowledge that *stage four* is indeed critical in many respects. Chapter 5 already signaled this, but the practical cases have shown it concretely. One especially noteworthy aspect is the role of stakeholders in facilitating access to information and data. The EHS and Emssperrwerk evaluations both showed that physical quantification of the impacts takes the most time. Furthermore, it is interesting that, in the EHS case, stakeholder involvement was crucial in the physical quantification stage in acquiring and generating data.

Stage five: Aggregate monetary scores consensus based and, Stage six: Aggregate non-monetary scores consensus based

Much has been learnt from the case studies about the aggregation of impacts and how it can be called ‘consensus based’. The case studies clearly show that the possible consensus only concerns the minimum importance of criteria and ease of understanding of measurement. This conclusion relates to *stages five and six*, to both monetary and non-monetary aggregation. From *stage six* alone it became clear that simplicity of the analysis should/may overrule the demand of completeness.

In the EHS study aggregating various impacts to a single monetary impact in a CBA worked well. Such a consensus based aggregation of monetary impacts mainly requires that included items are perceived by most stakeholders as easily monetarized because they belong to the everyday economic realm. The SCP evaluation demonstrated that at the individual company level there is ample opportunity for consensus based monetary aggregation. In practice, however, consensus cannot be expected to be reached on every detail of reasoning, or on all the different estimates and assumptions.

The EHS evaluation made clear that trying to be complete leads to a criteria structure and a measurement method which is harder to grasp. Applying the consensus based perspective of MCCBA more strictly will lead to greater attention to simpler indicators.

Stage seven: Interpret trade-offs

Stage seven proved to be very rewarding in the EHS and Emssperrwerk case studies – its value is difficult to overestimate. It may turn out that the main results of the evaluation are generated at this stage, as it can potentially strongly increase the judgement value of the results from stages five and six. In the EHS study for many stakeholders the interpretations of trade-offs using various ratios were the main results of the evaluation. It may substantially increase the value of the judgement information to relate the outcomes to relevant benchmarks. Options for benchmarking and interpretation increase markedly in the case of standardisation of impact measurement.

Stage eight: Perform sensitivity analysis and reconsider project alternatives

The main conclusion from stage eight is that although the sensitivity analysis can show increased problem understanding – a stumbling block may be the attention it receives in presenting main results. The explicit demand for reconsideration of project alternatives has the potential to remove this stumbling block, but was not yet practiced in the case studies.

In the EHS evaluation the sensitivity analysis did focus on sharing increased problem understanding, at least in as much as it concerned the CBA part of the analysis; the results were not however, given prominence. In the Emssperrwerk case this stage clearly showed the crucial variables (extra growth and extra growth period that Meyer might achieve). Furthermore, the sensitivity of the outcomes for choosing an impact area or impact population was clarified: at a very low level the project is easily beneficial.

Although no real evaluation results were presented in the SCP evaluation, the reconsideration of project alternatives in the light of evaluation outcomes (from other evaluation methods) was relevant and the analysis aimed at (reiterating until) project alternatives that might give substantial outcomes were found. The implication here was that the focus in the evaluation should shift from company to sector.

Having discussed these outcomes per stage the main building blocks of the MCCBA approach as identified in chapter 5 may be restated: 1) stakeholder involvement is useful for broadness of views and for checking on consensus; 2) standardisation of a global and long-term Triple E criteria structure is necessary; 3) judgement and measurement are closely related and in an MCCBA approach both should be understandable to broad groups of stakeholders. Now that the case studies in part three have been discussed, a general conclusion is that number 1) needs ‘qualification’: stakeholder involvement may also be important for acquiring relevant data, but difficulties remain around fruitful involvement. The other building blocks have proven to be valid in practice.

10.3 Final reflections

Decision aid versus decision-making?

Chapter 5 showed that MCCBA aims at a decision-aiding style of performing project evaluation in the sustainability context. This implies that MCCBA should be used for structuring judgement information and framing the problem situation. The structuring and framing can be pursued aiming for more and more or fuller and fuller judgement, as long as broad groups of stakeholders understand and agree upon the partial judgement elements. Now the question may be raised, to what extent did the MCCBA case studies of the chapters 7,8 and 9 follow the decision-aid style; or were they performed in a decision-making style – presenting results as being ‘best choices’? Pondering this question raises two topics for our attention.

The first is the position of the analyst, a naturally dominant factor concerning the style of the evaluation. It was noted earlier that the evaluator in a sustainability context should have myriad qualities in his or her search for a balance between analytical rigour and acceptability and ease of understanding to a broad group of stakeholders.

Second, whether an evaluation is decision aiding in style also depends on the perception of the recipient stakeholders; what may seem to be ‘firm advice’ to one, may feel like ‘a nice suggestion’ to someone else. The perception of results as a topic for further research may be a rewarding and, perhaps, necessary endeavour to give a satisfying answer to the question raised.

Testing for consensus

Consensus among a broad group of stakeholders concerning the importance of judgement concepts and the understanding of measurement issues is a crucial element in the improvement of accountability. In the case studies incidental testing was used but little explicit testing took place. Such testing should be undertaken, but it obviously places an extra workload on the evaluation and may therefore not always be addressed. Separate research

conducted on this topic may be one way to proceed. An attractive option seems to redo important evaluations that have aroused fierce public debate and to test different measurement options and result presentations among the stakeholders.¹

Compared with CBA and MCA: Improvement or more problems than before?

Formulated loosely, MCCBA may be seen as an improvement upon the strictness-in-reasoning and poverty-in-measurement of CBA; it simultaneously tries to improve upon the over-flexibility and over-richness of MCA. This approach, as seen above, results in more measurements in – or close to – natural dimensions, more attention to causality among criteria, and greater focus on most important criteria. This should ultimately lead to a gain in persuasive power of an evaluation.

However, clearly from case studies, in practice one remains with the problems of both CBA and MCA *to some extent*; even in an MCCBA approach the CBA problem of strictness in reasoning will linger, more complete MCA criteria will still suffer from a difficulty in understanding, etc. A cynical analyst could argue that instead of problems associated with one technique, the problem is now doubled.

This thesis has argued that MCCBA has important added value because the extent to which these problems occur can be actively managed within the evaluation, and thus the potential ‘problematicness’ to broad groups of stakeholders can be reduced.

The ‘hidden’ power of CBA

The power of CBA is often associated with its monetary valuation options. For determining the proper limits of CBA application in the sustainability context, the citizen-consumer dual preference structure proved very useful in this thesis. The more the judgement values involved approach the citizen spectrum (e.g. equity, non-use, long-term), the less useful monetary assessment seems to be.²

Be that as it may, this thesis has shown a somewhat ‘hidden’ power of CBA: its strict causal reasoning. Although the MCA literature has shown that the notion of tracking causality is not completely absent, this notion is far more dominant in any practical CBA. In the sustainability context one may even appreciate this CBA power as much as the monetary and discounting measurement possibilities with which it is most associated.

Sustainability: Global level more important than future generations

Chapter 2 briefly addressed sustainability notions and in chapter 9 we returned to sustainability. The challenge of evaluation in the sustainability context and the merits of MCCBA raises the following point: should global level evaluation be more prominent than long-term evaluation?

The notion of sustainability is closely connected with ‘future generations’ and the long-term perspective. But any assessment of future impacts is necessarily uncertain, regardless of the amount of research effort spent on the issue. The rate of uncertainty may differ, but generally speaking, the longer the time scale, the higher the uncertainty.

Apart from the long-term perspective, MCCBA has focused largely on the global level of evaluation. Although decision-makers may have limited interest in this level of evaluation, it is crucial in the assessment of sustainability; and in many practical cases its perceived

¹ Because of its redo character, relatively little time is required to gather data, and consensus issues can be focused on.

² Although there may be promising directions around the provision point mechanism CVM in trying to combine both citizen and consumer spectra.

relevance by decision-makers and stakeholders is increasing. A positive aspect of global level evaluation is that there need not be much uncertainty left; given sufficient research effort.

Long-term and global level are both important for sustainability evaluation, but it might be more fruitful from a research standpoint to exert effort at the global level: on gathering missing global data, and meta-analyses to find benchmarking options that may facilitate global assessments in individual project evaluation.³

Obviously, due to the weakness of global political power, no overly optimistic expectations should be held about the impact of such a new emphasis.

The uncertainty reduction

Chapter 2 discussed the possible reduction by evaluation of the unawareness and uncertainty spaces. Before an evaluation decision-makers and stakeholders will be uncertain about the particular impacts of various alternatives and they may be unaware of others. The evaluation changes this situation; generally uncertainty and unawareness will decrease due to the evaluation. The extent to which uncertainty and unawareness are decreased may vary. Throughout the thesis it was argued that when aiming for accountable evaluation in the sustainability context the combination of CBA and MCA in the MCCBA approach gives better judgement-oriented evaluation results than either MCA or CBA alone would do. The MCCBA may provide more stakeholders with information they can grasp and therefore help to reduce uncertainty and unawareness of impacts. In this sense the MCCBA approach reduces the uncertainty space more than either MCA or CBA alone.⁴

Still, the end result of the evaluation may also be that remaining uncertainty is large.⁵ If this situation occurs in a specific setting, the entire exercise of combining techniques may then be futile; the possibilities for decision aid are simply not big enough.

The novelty of the approach

Although MCCBA developed here can be classified as new, several authors have combined CBA and MCA (elements) in practice to an extent (Oosterhaven, 1982; Cortenraad, et al., 1986; Lichfield, 1996; Liu et al., 1998; Janssen and Padilla, 1999). On closer inspection some of these approaches may differ seriously from the approach elaborated here (e.g., Liu et al., 1998), while other combinations come rather close (e.g., Janssen and Padilla, 1999). The main contribution of this thesis has been to clarify the relative advantages of the different CBA and MCA elements and outline the common ground for integrating both approaches.

The thesis showed that, although CBA and MCA are often perceived as being ‘worlds apart’, in the reality of the sustainability context they can be put to work together easily and legitimately.

Dominance of practice over theory

This thesis is part of ‘the active to-and-fro between theory and practice’. At several points it was noted that the development of evaluation techniques has only been loosely connected with evaluation theory. Welfare economics is often thought to provide a basis for CBA but in practice CBA is used more or less independently from it; and sophisticated MCA techniques are scarcely used, let alone understood, in practice. The practice of evaluation dominates the direction of the field, be it towards CVM in CBA, or towards large value trees in MCA.

³ The Costanza et al. (1997) paper is praiseworthy for its attempts in this direction.

⁴ Referring to figure 2.5 of chapter 2 – and probably exaggerating the claims visually – instead of finding oneself in situation B when using either CBA or MCA – one moves to situation C when using MCCBA.

⁵ As in situation A in figure 2.5.

Perhaps unexpectedly, this thesis has also shown that even CBA and MCA theory acknowledge that as a final basis of the practices ‘ethical consensus’ and ‘jointly held objective conceptions’ rank high.

The success of MCCBA is therefore hardly a theoretical matter. Nevertheless, the confusion surrounding the correct applicability of different techniques in the context of accountable sustainability evaluation seems imposing enough to warrant further elaboration of theoretical reasoning in order to underpin the use of MCCBA – as this thesis has done. Hopefully, it will trigger fruitful and widespread use in practice.

LITERATURE CITED

- Abelson P., 1979. Cost-Benefit Analysis and environmental problems, Saxon House, Westmead.
- Adler M. and E. A. Posner, 1999. Rethinking cost-benefit analysis. *The Yale law journal*, 109:165-247.
- Adriaanse A., 1993. Environmental policy performance indicators - a study on the development of indicators for environmental policy in the Netherlands, Sdu Uitgeverij Koninginnegracht, Den Haag.
- Aguilera-Klink F., E. Pérez-Moriana, and J. Sánchez-Garcia, 2000. The social construction of scarcity. The case of water in Tenerife (Canary Islands). *Ecological Economics*, 2000:233-235.
- Ahlenius I.-B., 1997. Auditing and evaluation in Sweden. In: E Chelimsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 80-85.
- Aivazian V. and Callen J.L., 2003. The core, transaction costs, and the Coase theorem. *Constitutional political economy*, 14:287.
- Aldred J. and M. Jacobs, 2000. Citizens and wetlands: evaluating the Ely's citizens' jury. *Ecological Economics*, 2000:217-232.
- Aldred J., 2002. Cost-benefit analysis, incommensurability and rough equality. *Environmental Values*, 11:27-48.
- Alexander E. R., 1998. Where do we go from here? In: N Lichfield, A Barbanente, D Borri, A Khakee, and A Pratt (Editors), *Evaluation in planning - Facing the challenge of complexity*. Kluwer Academic Publishers, Dordrecht, pp. 355-374.
- Alexander E. R., 2001. Unvaluing evaluation: sensitivity analysis in MODM applications. In: H Voogd (Editor), *Recent developments in evaluation in spatial, infrastructure and environmental planning*. Geo Press, Groningen, pp. 319-340.
- Alkin M. C., M. Q. Patton, C. H. Weiss, and R. Connor, 1990. *Debates on evaluation*, Sage, Newbury Park.
- Alston J. M. and D. M. Larson, 1993. Hicksian versus Marshallian Welfare measures: why do we do what we do? *American Journal of Agricultural Economics*, 75:764-770.
- Amelung S. B., M. B. A. Van Asselt, S. C. H. Greeuw et al., 2001. *Evaluatie van OEEI (Onderzoeksprogramma Economische Effecten Infrastructuur)*. ICIS - International Centre for Integrative Studies, Maastricht.
- Arrow K. J. and H. Raynaud, 1986. *Social choice and multicriterion decision-making*, MIT Press, Cambridge Massachusetts.
- Ashby E., 1980. What worth the furbish lousewort? *Environmental Science and Technology*, 14:1176-1181.
- Atherton E. and S. French, 1998. Valuing the future: A MADA example involving nuclear waste storage. *Journal of multi-criteria decision analysis*, 7:304-321.

- Ayres R. U., 1998. The price-value paradox. *Ecological Economics*, 17-19.
- Ayres R. U., 2001. The need for a new growth paradigm. In: CL Cleveland, DI Stern, and R Costanza (Editors), *The economics of nature and the nature of economics*. Edward Elgar, Cheltenham, pp. 111-133.
- Azar C., J. Holmberg, and K. Lindgren, 1996. Socio-ecological indicators for sustainability. *Ecological Economics*, 18:89-112.
- Azar C., 1999. Weight factors in Cost-Benefit Analysis of climate change. *Environmental and Resource Economics*, 13:249-268.
- Bal D., H. M. Beije, Y. R. Hoogeveen, S. R. J. Jansen, and P. J. v. d. Reest, 1995. *Handboek Natuurdoeltypen in Nederland (Target Nature Types Manual)*, IKC Natuurbeheer, Ministerie van Landbouw, Natuurbeheer en Visserij, Wageningen.
- Balbus I. D., 1971. The concept of interest in pluralist and Marxist analysis. *Politics and Society*, 1:151-177.
- Bana e Costa C. A. and (ed.), 1990. *Readings in multiple criteria decision aid*, Springer-Verlag, Berlin, Heidelberg.
- Bana e Costa C. A. and M. Pirlot, 1997. Thoughts on the Future of the Multicriteria Field: Basic Convictions and Outline for a General Methodology. In: J Clímaco (Editor), *Multicriteria Analysis*. Proceedings of the XIth International Conference on MCDM, 1-6 August, 1994, Coimbra, Portugal.
- Bana e Costa C. A., T. J. Stewart, and J.-C. Vansnick, 1997. Multicriteria decision analysis: some thoughts based on the tutorial and discussion session of the ESIGMA meetings. *European Journal of Operational Research*, 28-37.
- Bana e Costa C. A., 2001. The use of multi-criteria decision analysis to support the search for less conflicting options in a mult-actor context: case study. *Journal of multi-criteria decision analysis*, 10:111-125.
- Barbanente A., D. Borri, and G. Concilio, 2001. Escapable dilemmas in planning: decisions vs transactions. In: H Voogd (Editor), *Recent developments in evaluation in spatial, infrastructure and environmental planning*. Geo Press, Groningen.
- Barbier E. B., A. Markandya, and D. W. Pearce, 1990. Environmental sustainability and cost-benefit analysis. *Environment-and-planning-A*, 22:1259-1266.
- Barde J.-P., D. Pearce, and (eds), 1991. *Valuing the environment - six case studies*, Earthscan Publications, London.
- Bateman I., 1994. Research methods for valuing environmental benefits. In: A Dubgaard, I Bateman, and M Merlo (Editors), *Economic valuation of benefits from countryside stewardship*. Wissenschaftsverlag Vauk, Kiel, pp. 47-82.
- Bateman I. J., A. A. Lovett, and J. S. Brainard, 2002. Developing a GIS-based methodology for benefit transfer. In: RJGM Florax, P Nijkamp, and KG Willis (Editors), *Comparative environmental economic assessment.*, pp. 154-176.
- Bazerman M. H., D. M. Messick, A. E. Tenbrunsel, K. A. Wade-Benzoni, and (eds), 1997. *Environment, ethics and behavior - The psychology of environmental valuation and degradation*, The

New Lexington Press, San Fransisco.

Becker N. and D. Soloveitchik, 1998. Dynamic Multiple-objective optimization of environmental regulation policies. *Journal of multi-criteria decision analysis*, 7:3-19.

Beinat E., 2001. Multi-Criteria Analysis for environmental management. *Journal of multi-criteria decision analysis*, 10:51.

Beladi H. and R. Frasca, 1996. Regional pollution and multinational firms. *Ecological Economics*, 17:117-125.

Bell M. L., B. F. Hobbs, E. M. Elliot, H. Ellis, and Z. Robinson, 2001. An evaluation of multi-criteria methods in Integrated Assessment of climate policy. *Journal of multi-criteria decision analysis*, 229-256.

Bendheim C. L., S. Waddock, and S. B. Graves, 1998. Determining best practice in corporate-stakeholder relations using data envelopment analysis. *Business and Society*, 37:306-339.

Bentkover J. D., V. T. Covello, J. Mumpower, and (eds), 1986. *Benefits Assessment - The state of the art*, D.Reidel Publishing Company, Dordrecht.

Berenschot (C.W.J.Schrijver and J.D.Wiersma), 1994. *Onderzoek naar de financiële gevolgen van het relatienota-beleid*, Berenschot Interim Management, Utrecht.

Berg A. E. v. d., 1999. Individual differences in the aesthetic evaluation of natural landscapes, Rijksuniversiteit Groningen - PhD Thesis, Groningen.

Bergh C. J. M. v. d., 2000. Themes, approaches, and differences with environmental economics, Tinbergen Institute Discussion paper 2000-080/3 edition. Tinbergen Institute, Rotterdam/Amsterdam.

Bergh C. J. M. v. d., 2002. Economie van klimaatbeleid: kritiek en alternatief. *Economisch Statistische Berichten*, 184-187.

Bergh C. J. M. v. d., 2004. Optimal climate policy is a utopia: from quantitative to qualitative cost-benefit analysis. *Ecological Economics*, 2004:385-393.

Bergland O., K. Magnussen, and S. Navrud, 2002. Benefit transfer: testing for accuracy and reliability. In: RJGM Florax, P Nijkamp, and KG Willis (Editors), *Comparative environmental economic assessment.*, pp. 117-132.

Bernadette M., Krishnamurty, and K. Paul, 1998. The development of a systematic aggregate measure of corporate social performance. *Journal of management*, 24:119-138.

Beroggi G. E. G., 1999. *Decision modeling in policy management - An introduction to the analytic concepts*, Kluwer Academic Publishers, Boston, Dordrecht, London.

Bettels K. and F. J. Sijtsma, 1998. *Het Emssperrwerk - Een evaluatie op duurzaamheid van een waterkering in de rivier de EMS*, Wetenschapswinkel voor Economie, Groningen.

Blair A., J. Debenham, and J. Edwards, 1997. A comparative study of methodologies for designing IDSSs. *European Journal of Operational Research*, 277-295.

Blamey R., M. Common, and J. Quiggin, 1995. Respondents to contingent valuation surveys: consumers or citizens? *Australian Journal of Agricultural Economics*, 39:263-288.

- Blaug M., 1985. *Economic theory in retrospect*, Cambridge University Press, Cambridge.
- Boadway R. W., 1974. The welfare foundations of cost-benefit analysis. *The Economic Journal*, 926-939.
- Boadway R. W. and D. E. Wildasin, 1984. *Public sector economics*, Little, Brown and Company, Boston, Toronto.
- Boardman A. E., D. Greenberg, A. Vining, and D. Weimer, 2001. *Cost-benefit analysis: concepts and practice*, Prentice Hall, Upper Saddle River, NJ.
- Boer B. d., P. R. Bosch, R. Brouwer, F. Duijnhouwer, and (eds), 1997. *Monetarisering van milieuverliezen - Eindrapport van het discussie platform*, Centraal Bureau voor de Statistiek, Voorburg.
- Boisvert V., N. Holec, and F.-D. Vivien, 1998. Economic and environmental information for sustainability. In: S Faucheux and M O'Connor (Editors), *Valuation for Sustainable Development - Methods and policy indicators*. Edward Elgar, Cheltenham, pp. 99-119.
- Bouyssou D., 1990. Building criteria: a prerequisite for MCDA. In: CA Bana e Costa (Editor), *Readings in Multiple criteria Decision Aid*. Springer-Verlag, Berlin, Heidelberg, pp. 58-80.
- Bowen H. R., 1953. *Social responsibilities of the businessman*, Harper & Row, New York.
- Brady F. N., 1990. *Ethical managing - Rules and results*, MacMillan Publishing Company, New York.
- Brans J. P., 2003. Ethics and decision. *European Journal of Operational Research*, 340-352.
- Brekke K. A., H. Luraos, and K. Nyborg, 1996. Allowing disagreement in evaluation of social welfare. *Journal of Economics*, 63:303-324.
- Brekke K. A., 1997. The numéraire matters in cost-benefit analysis. *Journal of Public Economics*, 64:117-123.
- Brekke K. A., 1998. Reply to J.Drèze and P.-O. Johansson. *Journal of Public Economics*, 70:495-496.
- Brent R. J., 1996. *Applied cost-benefit analysis*, Edward Elgar, Cheltenham.
- Brent R. J., 1998. *Cost-benefit analysis for developing countries*, Edward Elgar, Cheltenham.
- Brink B. t., 2002. *Biodiversity indicators for the OECD environmental outlook and strategy - a feasibility study*. RIVM - National Institute of Public Health and the Environment, Bilthoven.
- Broadhurst A. I., 2000. Corporations and the ethics of social responsibility: an emerging regime of expansion and compliance. *Business Ethics: A European review*, 9:86-98.
- Broekhof M. and F. J. Sijtsma, 2002. *Maatschappelijk Verantwoord Ondernemen: hoe meet men dat?* SPIL, 23-28.
- Bromley D. W., 2001. Searching for sustainability: the poverty of spontaneous order. In: CL Cleveland, DI Stern, and R Costanza (Editors), *The economics of nature and the nature of economics*. Edward Elgar, Cheltenham, pp. 74-88.

- Brouwer R., I. H. Langford, I. J. Bateman, and R. K. Turner, 1999. A meta-analysis of wetland contingent valuation studies. *Regional Environmental Change*, 1:47-57.
- Brouwer R., 2000. Environmental value transfer: state of the art and future prospects. *Ecological Economics*, 137-152.
- Bröcker J., 2003. Indirect effects: do they matter for Cost-Benefit analysis? Twee jaar ervaring met OEEI. Centraal Planbureau, The Hague, pp. 31-39.
- Bruce N. and R. G. Harris, 2001. Cost-benefit criteria and the compensation principle in evaluating small projects. *Journal of Political Economy*, 90:755-776.
- Bulte E. H. and G. C. v. Kooten, 1999. Marginal valuation of charismatic species: implications for conservation. *Environmental and Resource Economics*, 119-130.
- Campen J. T., 1986. Benefit, cost, and beyond - The political economy of benefit-cost analysis. Ballinger Publishing Company, Cambridge, Massachusetts.
- Carlsson C. and R. Fuller, 1997. Interdependence in multiple criteria decision making. Springer Verlag, Berlin, pp. 25-36.
- Carroll A. B., 1999. Corporate social responsibility - Evaluation of a definitional construct. *Business and Society*, 38:268-295.
- Cernea M. M., 1993. The sociologist's approach to sustainable development. *Finance & Development*, 11-13.
- Chelimsky E., 1997. The coming transformations in evaluation. In: E Chelimsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, pp. 1-26.
- Chelimsky E., W. R. Shadish, and (ed.), 1997. *Evaluation for the 21st century - a handbook*, Sage, Thousand Oaks, London, New Delhi.
- Chelimsky E., 1997. The political environment of evaluation and what it means for the development of the field. In: E Chelimsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 53-68.
- Clarke Ph. M., 2002. Testing the convergent validity of the contingent valuation and travel cost methods in valuing the benefits of health care. *Health Economics*, 117-127.
- Clayton A. M. H. and N. J. Radcliffe, 1996. *Sustainability: a systems approach*, Earthscan, London.
- Cleveland C. J., D. I. Stern, R. Costanza, and (eds), 2001. *The economics of nature and the nature of economics*, Edward Elgar, Cheltenham.
- Clinch J. P. and A. Murphy, 2001. Modelling winners and losers in contingent valuation of public goods: appropriate welfare measures and econometric analysis. *The Economic Journal*, 420-443.
- Clímaco J. and (ed.), 1997. *Multicriteria Analysis*, Springer Verlag, Berlin.
- CLM (J.C.Buys), 1995. *Naar een natuurmeetlat voor landbouwbedrijven*, Centrum voor Landbouw en Milieu, Utrecht.

Cohon J. L., 1997. MCDM and sustainable development: the case of water resource planning in India. In: MH Karwan, J Spronk, J Wallenius, and (eds) (Editors), *Essays in decision making - A volume in honour of Stanley Zionts*. Springer-Verlag, Berlin, pp. 294-312.

Coker A. and C. Richards, 1992. *Valuing the environment - Economic approaches to environmental evaluation*, Belhaven Press, London.

Commissie Financiering Natuurbeleid, 1992. *Naar een sluitende groenfinanciering - Een bittere noodzaak om Nederland ook voor de toekomstige generaties leefbaar te houden*, The Hague.

Common M., Reid, I., and R. Blamey, 1997. Do existence values for Cost-Benefit Analysis exist? *Environmental and Resource Economics*, 9:225-238.

Connor H., 1999. Taking Non-Monetizable Impacts (NMIs) into Account in an Eco-Development Strategy. In: M O'Connor and CL Spash (Editors), *Valuation and the Environment: theory, method and practice*. Edward Elgar, Cheltenham.

Consumentenbond, 2001. *Eerlijk duurt het langst.*, pp. 24-31.

Cook T. D., 1997. Lessons learned in evaluation over the past 25 years. In: E Chelmsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 30-52.

Cortenraad W. M. H., K. J. Moning, and J. H. Jonk, 1986. *Allocation of priorities for national highway projects in the Netherlands*. Rijkswaterstaat, The Hague.

Costanza R., R. D'Arge, R. De Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. Van den Belt, 1997. The value of the world's ecosystem services and natural capital. *Nature*, 253-260.

Council of the European Union, 2006. *Renewed EU sustainable development strategy*.

Covello V. T. and M. W. Merkhofer, 1993. *Risk assessment methods - Approaches for assessing health and environmental risks*, Plenum Press, New York.

Curtis F., 2003. Eco-localism and sustainability. *Ecological Economics*, 46:83-102.

Dabbert S., A. Dubgaard, L. Slangen, and M. E. Whitby, 2001. *The Economics of Landscape and Wildlife Conservation*, CAB International, Oxon, UK.

Daly H. E. and J. B. Cobb (jr.), 1989. *For the common good - redirecting the economy toward community, the environment, and a sustainable future*, Beacon Press, Boston.

Daly H. E., 1998. The return of Lauderdale's paradox. *Ecological Economics*, 21-23.

Dasgupta P., A. Sen, and S. Marglin, 1972. *Guidelines for project evaluation*, United Nations, New York.

De Marchi B., S. O. Funtowicz, S. Lo Cascio, and G. Munda, 2000. Combining participative and institutional approaches with multicriteria evaluation. An empirical study for water issues in Troina, Sicily. *Ecological Economics*, 2000:267-282.

Deegan C., 2000. *Financial accounting theory*, McGraw-Hill Australia, Roseville.

- Diamond P., 1996. Testing the internal consistency of contingent valuation surveys. *Journal of Environmental Economics and Management*, 337-347.
- Diaz-Balteiro L. and C. Romero, 2004. In search of a natural systems sustainability index. *Ecological Economics*, 2004:401-405.
- Diggelen R. v., F. J. Sijtsma, D. Strijker, and J. v. d. Burg, 2005. Relating land use intensity and biodiversity at the regional scale. *Basic and Applied Ecology*, 6:145-159.
- DJSI - Dow Jones Sustainability Indexes, 2001. Dow Jones STOXX Sustainability Indexes Guide - version 1.0, www.sustainability-index.com.
- Doane D., 2000. The corporate spin - the troubled teenage years of social reporting. New Economics Foundation, London.
- Doane D., 2002. Market failure: the case for mandatory social and environmental reporting. New Economics Foundation, London.
- Drèze J. and N. Stern, 1994. Shadow prices and markets: policy reform, shadow prices and market prices. In: R Layard and S Glaister (Editors), *Cost-Benefit Analysis*. Cambridge University Press, Cambridge, pp. 59-99.
- Drèze J., 1998. Distribution matters in cost-benefit analysis - Comment on K.A.Brekke. *Journal of Public Economics*, 70:485-488.
- Drucker P. F., 1984. The new meaning of corporate social responsibility. *California Management Review*, Vol. XXVI:53-63.
- DTLG, 2001. Multi criteria analysis: a manual. Department for transport, local government and the regions.
- Dubgaard A., I. Bateman, M. Merlo, and (eds), 1994. Economic valuation of benefits from countryside stewardship, Wissenschaftsverlag Vauk, Kiel.
- Eding G. J., T. M. Stelder, and J. Oosterhaven, 1995. Structuur en interactie, een interregionale structuurschets van Groningen, Friesland en Drenthe voor 1990, Stichting Ruimtelijke Economie Groningen, Groningen.
- Edwards-Jones G., B. Davies, and S. Hussain, 2000. *Ecological economics - an introduction*, Blackwell Science, Oxford.
- Edwards W. and J. R. Newman, 1982. *Multiattribute Evaluation*, Sage, Beverly Hills.
- Edwards W. and D. Von Winterfeldt, 1986. On cognitive illusions and their implications. In: HR Arkes and KR Hammond (Editors), *Judgment and Decision making*. Cambridge University Press, Cambridge, pp. 642-679.
- Eijgenraam C. J. J., C. C. Koopmans, P. J. G. Tang, and A. C. P. Verster, 2000. Evaluatie van infrastructuurprojecten - Leidraad voor kosten-baten analyse, Centraal Planbureau / Nederlands Economisch Instituut, Den Haag / Rotterdam.
- Eijgenraam C. J. J., C. C. Koopmans, P. J. G. Tang, and A. C. P. Verster, 2003. Evaluatie van infrastructuurprojecten - Leidraad voor kosten-batenanalyse (tweede editie). Centraal Planbureau /

Nederlands Economisch Instituut, Den Haag / Rotterdam.

El Serafy S., 1998. Pricing the invaluable: the value of the world's ecosystems services and natural capital. *Ecological Economics*, 25-27.

Elhorst J. P., J. Oosterhaven, F. J. Sijtsma, and T. M. Stelder, 1999. Welfare effects of spatial deconcentration: a scenario for the Netherlands. *TESG (Tijdschrift voor Economische en Sociale Geografie)*, 90:17-31.

Elhorst J. P., A. Heyma, C. C. Koopmans, and J. Oosterhaven, 2004. Indirecte effecten infrastructuurprojecten - Aanvulling op de Leidraad OEI, Ministerie van Verkeer en Waterstaat en Ministerie van Economische Zaken, Den Haag.

Elkington J., 1997. *Cannibals with forks: the triple bottom line of 21st century business*, Capstone publishing, Oxford.

Engel S., 2002. Benefit function transfer versus meta-analysis as policy-making tools: a comparison. In: RJGM Florax, P Nijkamp, and KG Willis (Editors), *Comparative environmental economic assessment.*, pp. 133-153.

Eom S. B. and H. Min, 1999. The contributions of multi-criteria decision making to the development of decision support systems subspecialties: an empirical investigation. *Journal of multi-criteria decision analysis*, 1999:239-255.

Espinasse B., G. Picolet, and E. Chouraqui, 1997. Negotiation support systems: a multi-criteria and multi-agent approach. *European Journal of Operational Research*, 1997:389-409.

Fabré P., 1997. *Manual – Financial and Economic Analysis of projects*, European Commission, Luxembourg.

Farrow S. and M. Toman, 1999. Using cost-benefit analysis to improve environmental regulations. *Environment*, 41:12-38.

Faucheux S., F. Laroui, and J. W. E. Velthuijsen, 1993. *Energy, economy and the environment: traditional models and alternative approaches*, SEO, Foundation for Economic Research University of Amsterdam, Amsterdam.

Faucheux S., G. Froger, and J.-F. Noël, 1993. Towards a reconsideration of rationality hypotheses for sustainable development decision-making. In: S Faucheux, F Laroui, and JWE Velthuijsen (Editors), *Energy, economy and the environment: traditional models and alternative approaches*. SEO, Foundation for Economic Research University of Amsterdam, Amsterdam, pp. 45-91.

Faucheux S. and M. E. O'Connor, 1998. *Valuation for Sustainable Development - Methods and policy indicators*, Edward Elgar, Cheltenham.

Faucheux S., G. Froger, and G. Munda, 1998. Multicriteria decision aid and the sustainability tree. In: S Faucheux, M O'Connor, and ' (Editors), *Valuation for Sustainable Development - Methods and policy indicators*. Edward Elgar, Cheltenham, pp. 187-214.

Federal Trade Commission, 1996. *Anticipating the 21st century - Competition policy in the new high-tech, global marketplace*, Federal Trade Commission (US).

Feyerabend P., 1975. *Against method*, Humanities Press, London.

- Figge F., T. Hahn, S. Schaltegger, and M. Wagner, 2002. The sustainability balanced scorecard - linking sustainability management to business strategy. *Business Strategy and the Environment*, 11:269-284.
- Finco A. and P. Nijkamp, 2001. Pathways to urban sustainability. *Journal of Environmental Policy and Planning*, 289-302.
- Fischhoff B., 1977. Cost-Benefit Analysis and the art of motorcycle maintenance. *Policy Sciences*, 8:177-202.
- Fischhoff B., 1991. Value elicitation: is there anything in there? *American Psychology*, 835-847.
- Fishburn P. C. and I. H. Lavalley, 1999. MCDA: Theory, practice and the future. *Journal of multi-criteria decision analysis*, 8:1-2.
- Florax R. J. G. M., 2002. Methodological pitfalls in meta-analysis: publication bias. In: RJGM Florax, P Nijkamp, and KG Willis (Editors), *Comparative environmental economic assessment.*, pp. 177-207.
- Folke C., C. S. Holling, and C. Perrings, 1996. Biological diversity, ecosystems, and the human scale. *Ecological Applications*, 6:1018-1024.
- Forte F., P. Nijkamp, and F. Torrieri, 2001. Shared choices on local sustainability projects: a decision support framework. *Vrije Universiteit Amsterdam, Faculteit der Economische wetenschappen en econometrie, Amsterdam.*
- Foxon T. J., G. McIlkenny, D. Gilmour, C. Oltean-Dumbrava, N. Souter, R. Ashley, D. Butler, P. Pearson, P. Jowitt, and J. Moir, 2002. Sustainability criteria for decision support in the UK water industry. *Journal of environmental planning and management*, 45:285-301.
- Franco L. G., 2004. The death of diversification? The focussing of the world's industrial firms, 1980-2000. *Business Horizons*, 47:41-50.
- French S., 1998. Decision making not decision theory. *Journal of multi-criteria decision analysis*, 7:303.
- Froger G. and G. Munda, 1998. Methodology for environmental decision support. In: S Faucheux, M O'Connor, and ' (Editors), *Valuation for Sustainable Development - Methods and policy indicators.* Edward Elgar, Cheltenham, pp. 165-186.
- Fromm O., 2000. Ecological structure and functions of biodiversity as elements of total economic value. *Environmental and Resource Economics*, 2000:303-328.
- Fusco Girard L. and P. De Toro, 2001. Human sustainable development and planning: towards integrated evaluations. In: H Voogd (Editor), *Recent developments in evaluation in spatial, infrastructure and environmental planning.* Geo Press, Groningen.
- Futh U., 2001. *Meten van natuurwaarden in Duitsland, Wetenschapswinkel voor Economie, Groningen.*
- Garrod G. and K. G. Willis, 1999. *Economic valuation of the environment - Methods and case studies,* Edward Elgar, Cheltenham, UK.
- Geneletti D., 2002. *Ecological evaluation for environmental impact assessment, PhD Thesis, Vrije*

Universiteit Amsterdam, Amsterdam.

Gollier C., 2001. Should we be aware of the precautionary principle? *Economic Policy*, 2001:303-327.

Gordon J. S., 1998. The tragedy of the commons - the fate of the codfish suggests that collectivism on the ownership level is as destructive as anywhere else. *American Heritage*, 49:22-23.

Goulder L. H. and R. N. Stavins, 2002. An eye on the future. *Nature*, 419:673-674.

Granovetter M., 1992. Economic action and social structure: the problem of embeddedness. In: M Zey and (ed.) (Editors), *Decision making - Alternatives to rational choice models*. Sage publications, Newbury Park, California, pp. 304-333.

Grasso P. G., S. S. Wasty, and R. V. Weaving, 2005. *World Bank operations evaluation department - the first 30 years*, The World Bank, Washington.

Gregory A., 2000. Problematizing participation - A critical review of approaches to participation in evaluation theory. *Evaluation*, 6:179-199.

GRI: Global Reporting Initiative, 2004. *Sustainability Reporting Guidelines*, GRI, www.globalreporting.org.

Groot R. S. d., 1992. *Functions of nature - Evaluation of nature in environmental planning, management and decision making*, Wolters-Noordhoff.

Guba E. G. and Y. S. Lincoln, 1989. *Fourth Generation Evaluation*, Sage Publications, Newbury Park.

Hammond J. S., R. L. Keeney, and H. Raiffa, 1999. *Smart choices - A practical guide to making better decisions*, Harvard Business School Press, Boston, Massachusetts.

Hanley N., 1992. Are There Environmental Limits to Cost-Benefit Analysis? *Environmental and Resource Economics*, 2:33-59.

Hanley N. and C. Spash, 1993. *Cost-Benefit Analysis and the environment*. Edward Elgar, Aldershot - Hants.

Hanley N., C. L. Spash, and L. Walker, 1995. Problems in Valuing the Benefits of Biodiversity Protection. *Environmental and Resource Economics*, 5:249-272.

Hanley N., R. E. Wright, and V. Adamowicz, 1998. Using choice experiments to value the environment - Design issues, current experience and future prospects. *Environmental and Resource Economics*, 11:413-428.

Hanley N., 2000. Macroeconomic measures of 'sustainability'. *Journal of Economic Surveys*, 14:1-30.

Harberger A. C., 1972. *Project Evaluation - Collected papers*, MacMillan, London.

Harrop D. O. and J. A. Nixon, 1999. *Environmental assessment in practice*, Routledge, London.

Hart S. L., 1997. Beyond greening: strategies for a sustainable world. *Harvard Business Review*, 66-76.

Heal G., 1997. *Valuing our future: cost-benefit analysis and sustainability*, United Nations

- Development Programme, Office of Development Studies (UNDP/ODS), New York.
- Heap J. L., 1995. Constructionism in the rhetoric and practice of fourth generation evaluation. *Evaluation and Program Planning*, 51-61.
- Hellendoorn J. C., 2001. *Evaluatiemethoden ex ante - een introductie*, SDU Uitgevers, Den Haag.
- Henriques I. and P. Sadorsky, 1999. The relationship between environmental commitment and managerial perceptions of stakeholder importance. *Academy of Management Journal*, 42:87-99.
- Herendeen R. A., 1998. Monetary-costing environmental services: nothing is lost something is gained. *Ecological Economics*, 29-30.
- Hibbitt Ch. J., 2004. External environmental disclosure and reporting by large European companies: an economic, social and political analysis of managerial behaviour, PhD Thesis Vrije Universiteit Amsterdam, Amsterdam.
- Hill G., 1999. Project appraisal for the Keynesian investment planner. *Economics of Planning*, 1999:153-164.
- Hirokawa R. Y. and M. S. Poole, 1996. *Communication and group decision making*, 2nd edition. Sage, Thousand Oaks.
- Hisschemöller M., R. S. J. Tol, and P. Vellinga, 2001. The relevance of participatory approaches in integrated environmental assessment. *Integrated Assessment*, 2001:57-72.
- Hodgson J. G., G. Montserrat-Marti, J. Tallowin, K. Thompson, S. Díaz, M. Cabido, J. P. Grime, P. J. Wilson, S. R. Band, A. Bogard, R. Cabido, D. Cáceras, P. Castro-Díez, C. Ferrer, M. Maestro-Martínez, M. C. Pérez-Rontomé, M. Charles, H. C. Cornelissen, S. Dabbert, N. Pérez-Harguindeguy, T. Krimly, F. J. Sijtsma, D. Strijker, F. Vendramini, J. Guerrero-Campo, A. Hynd, G. Jones, A. Romo-Díez, L. Torres Espuny, P. Villar-Salvador, and M. R. Zak, 2005. How much will it cost to save grassland diversity? *Biological Conservation*, 2005:263-273.
- Hoevenagel R., 1994. An Assessment of the Contingent Valuation Method. In: R Pethig (Editor), *Valuing the Environment: Methodological and Measurement Issues*. Kluwer, Dordrecht, pp. 195-228.
- Hoevenagel R., 1994. A Comparison of Economic Valuation Methods. In: R Pethig (Editor), *Valuing the Environment: Methodological and Measurement Issues*. Kluwer, Dordrecht, pp. 251-270.
- Hoevenagel R., 1994. *The contingent valuation method: scope and validity*, Vrije Universiteit Amsterdam - PhD thesis, Amsterdam.
- Holmes T. P., J. C. Bergstrom, E. Huszar, S. B. Kask, and F. Orr III, 2004. Contingent valuation, net marginal benefits, and the scale of riparian ecosystem restoration. *Ecological Economics*, 49:19-30.
- House E. R., 2000. The limits to cost-benefit analysis. *Evaluation*, 6:79-86.
- Howarth A., D. W. Pearce, E. Ozdemiroghu et al., 2001. *Valuing the benefits of environmental policy: the Netherlands*. RIVM report 481505 024., RIVM - National Institute of public health and the environment, Bilthoven.
- Howarth R. B., 2001. Sustainability under uncertainty: a deontological approach. *Land Economics*, 71:417-427.

- Hueting R., 1991. The use of the discount rate in a cost-benefit analysis for different uses of a humid tropical forest area. *Ecological Economics*, 3:43-57.
- Hueting R., L. Reijnders, B. d. Boer, J. Lambooy, and H. Jansen, 1998. The concept of environmental function and its valuation. *Ecological Economics*, 31-35.
- Husted B. W. and D. B. Allen, 2000. Is it ethical to use ethics as a strategy? *Journal of Business Ethics*, 21-31.
- Hwang C. L. and M. J. Lin, 1987. *Group decision making under multiple criteria*, Springer Verlag, Berlin.
- Hyman D. and J. Shingler, 1999. The hierarchy of consumer participation and patterns of economic social and political participation. *The Journal of consumer affairs*, 33:380-407.
- Irwin A., 1995. *Citizen Science - A study of people, expertise and sustainable development*, Routledge, London.
- Islam S. M. N. and M. Clarke, 2002. The relationship between economic development and social welfare: a new adjusted GDP measure of welfare. *Social Indicators Research*, 201-228.
- Ittner C. D. and D. F. Larcker, 1998. Innovations in performance measurement: trends and research implications. *Journal of Management Accounting Research*, 10:205-247.
- Ittner C. D. and D. F. Larcker, 1998. Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of Accounting Research*, 36:1-35.
- J.C.J.M.van den Bergh, 2004. Optimal climate policy is a utopia: from quantitative to qualitative cost-benefit analysis. *Ecological Economics*, 2004:385-393.
- Jacobs M., 1991. *The green economy - environment, sustainable development and the politics of the future*, Pluto Press, London.
- Jaffry S., H. Pickering, Y. Ghulam, D. Whitmarsh, and P. Wattage, 2004. Consumer choices for quality and sustainability labelled seafood products in the UK. *Food Policy*, 2004:215-228.
- Janssen R., 1994. *Multiobjective decision support for environmental management*. Kluwer Academic Publishers, Dordrecht.
- Janssen R. and J. E. Padilla, 1999. Preservation or conversion? Valuation and evaluation of a mangrove forest in the Philippines. *Environmental and Resource Economics*, 1999:297-331.
- Janssen R., M. v. Herwijnen, and E. Beinat, 2000. *BeslissingsOndersteunend Systeem voor Discrete Alternatieven (BOSDA)*, Instituut voor Milieuvraagstukken - Vrije Universiteit Amsterdam, Amsterdam.
- Janssen R., 2001. On the use of multi-criteria analysis in environmental impact assesment in the Netherlands. *Journal of multi-criteria decision analysis*, 10:101-109.
- Jepma C. J., 1995. *Tropical Deforestation - A socio-economic approach*, Earthscan Publications Ltd, London.

- Jepma C. J., H. Jager, and E. Kamphuis, 1996. Introduction to international economics, Dutch Open University.
- Jepma C. J. and M. Munasinghe, 1998. Climate change policy - Facts, issues, and analyses, Cambridge University Press, Cambridge.
- Jeurissen R. J. M. and (ed.), 2000. Bedrijfsethiek: een goede zaak, Van Gorcum, Assen.
- Johansson-Stenman O., 2000. On the value of life in rich and poor countries and distributional weights beyond utilitarianism. *Environmental and Resource Economics*, 17:299-310.
- Johansson P.-O., 1993. Cost-benefit analysis of environmental change. Cambridge University Press, Cambridge.
- Johansson P.-O., 1998. Does the choice of numéraire matter in cost-benefit analysis? *Journal of Public Economics*, 70:489-493.
- Jones Ph. and J. Cullis, 1996. Legitimate and illegitimate transfers: dealing with "political" cost-benefit analysis. *International review of law and economics*, 16:247-258.
- Jong S. H. d., 2006. EurepGAP en ontwikkelingslanden - een onderzoek naar het effect van maatschappelijk verantwoorde landbouw in Kenia and Senegal, Faculteit Ruimtelijke Wetenschappen (master thesis).
- Jordan G. J. and M. J. Fortin, 2002. Scale and topology in the ecological economics sustainability paradigm. *Ecological Economics*, 2002:361-366.
- Jorgenson D. W., 1993. Applied welfare economics: discussion. *American Journal of Agricultural Economics*, 75:782-783.
- Joubert A. R., A. Leiman, H. M. de Klerk, S. Katau, and J.C. Aggenbach, 1997. Fynbos (fine bush) vegetation and the supply of water: a comparison of multi-criteria decision analysis and cost-benefit analysis. *Ecological Economics*, 123-140.
- Junger P. D., 1979. The inapplicability of Cost-Benefit Analysis to environmental policies. *Ekistics*, 46:184-194.
- Kacprzyk J., M. Fedrizzi, and (eds), 1990. Multiperson decision making models using fuzzy sets and possibility theory, Kluwer Academic Publishers, Dordrecht.
- Kahneman D. and A. e. Tversky, 2000. Choices, values and frames, Cambridge University Press, Cambridge.
- Kahneman D. and A. Tversky, 2000. Choices, values and frames. In: D Kahneman and Ae Tversky (Editors), Choices, values and frames. Cambridge University Press, Cambridge, pp. 1-16.
- Kahneman D. and A. Tversky, 2000. Prospect theory - an analysis of decision under risk. In: D Kahneman and Ae Tversky (Editors), Choices, values and frames. Cambridge University Press, Cambridge, pp. 17-43.
- Kahneman D., I. Ritov, and D. Schkade, 2000. Economic preferences or attitude expressions - An analysis of dollar responses to public issues. In: D Kahneman and Ae Tversky (Editors), Choices, values and frames. Cambridge University Press, Cambridge, pp. 642-671.

Kahneman D., 2000. New challenges to the rationality assumption. In: D Kahneman and Ae Tversky (Editors), Choices, values and frames. Cambridge University Press, Cambridge, pp. 758-774.

Kamminga K. J., 2001. Steering sustainability - On the potential of a radical energy tax-subsidy scheme to move the Dutch society towards sustainable development in the energy domain, University of Groningen - PhD thesis, Groningen.

Kamp-Roelands N. and C. Hibbit, 2000. Maatschappelijke verslaggeving: de richtlijnen van Global Reporting Initiative. De Accountant, 184-187.

Kapteyn A., 1985. Utility and economics. De Economist, 133 :1-20.

Karwan M. H., J. Spronk, J. Wallenius, and (eds), 1997. Essays in decision making - A volume in honour of Stanley Zionts. Springer-Verlag, Berlin.

Keeney R. L. and H. Raiffa, 1976. Decisions with multiple objectives: preferences and value tradeoffs, John Wiley and Sons, New York.

Keeney R. L., D. V. Winterfeldt, and T. Eppel, 1990. Eliciting public values for complex policy decisions. Management Science, 36:1011-1030.

Keeney R. L., 1992. Value-focussed thinking - A path to creative decisionmaking, Harvard University Press, Cambridge, Massachusetts.

Keeney R. L., 1996. Valuing billions of dollars. In: RJ Zeckhauser, RL Keeney, JK Sebenius, and (eds) (Editors), Wise choices - Decisions, games and negotiations. Harvard Business School Press, Boston, Massachusetts, pp. 63-80.

Kingma B. R., 1996. The economics of information - A guide to economic and Cost-Benefit Analysis for information professionals, Libraries unlimited, Englewood Colorado.

Kitson A. and R. Campbell, 1996. The ethical organisation - ethical theory and corporate behavior, MacMillan Business, London.

Kleindorfer P. R., H. C. Kunreuther, and P. J. H. Schoemaker, 1993. Decision sciences - An integrative perspective, Cambridge University Press, Cambridge.

Klose Th., 2002. A utility-theoretic model for QALY's and willingness to pay. Health Economics, 12 (1): 17-31.

Knaap T. and J. Oosterhaven, 2001. Het eerste ruimtelijke algemene evenwichtsmodel voor Nederland, met resultaten voor de magneetweefbaan Schiphol-Groningen. Maandschrift Economie, 65:89-107.

Kopp R. J., 1992. Why existence value should be used in Cost-Benefit Analysis. Journal of Policy Analysis and Management, 11:123-130.

Korhonen P. and J. Wallenius, 1994. Behavioral Issues in MCDM: Neglected Research Questions. In: J Climaco (Editor), Multicriteria Analysis.

Krabbe J. J., 2001. A "physiocratic" approach to sustainability. International Journal of Social Economics, 20:15-21.

- Krutilla J. V., 1981. Reflections of an applied welfare economist. *Journal of Environmental Economics and Management*, 8:1-10.
- Lahdelma R., P. Salminen, and M. Kuula, 2005. Testing the efficiency of two pairwise comparison methods in discrete criteria problems. *European Journal of Operational Research*, 2005:496-508.
- Lakshmanan T. R., P. Nijkamp, P. Rietveld, and E. T. Verhoef, 2001. Benefits and costs of transport. Classification, methodologies, and policies. *Papers in regional science: Journal of the regional science association international*, 80:139-164.
- Langhelle O., 2000. Sustainable development and social justice: expanding the Rawlsian framework of global justice. *Environmental Values*, 9:295-323.
- Layard R., S. Glaister, and (eds), 1994. *Cost-Benefit Analysis*, 2nd edition. Cambridge University Press, Cambridge.
- Layard R. and A. A. Walters, 1994. Income distribution: allowing for income distribution. In: R Layard and S Glaister (Editors), *Cost-Benefit Analysis*. Cambridge University Press, Cambridge, pp. 179-198.
- Lefley F., 2004. An assessment of various approaches for evaluating project strategic benefits: Recommending the strategic index. *Management decision*, 42:850.
- Lempert R. J., M. E. Schlesinger, and S. C. Bankes, 1996. When we don't know the costs or the benefits: adaptive strategies for abating climate change. *Climatic Change*, 1996:235-274.
- Lempert R. J., 1999. Book Review of F.L.Tóth - *Cost-Benefit analysis of climate change: the broader perspective*. *Climatic Change*, 635-640.
- Levin H. M. and P. J. McEwan, 2001. *Cost-effectiveness analysis - methods and applications*, Sage Publications, Thousand Oaks.
- Lichfield N., 1996. *Community impact evaluation*, UCL Press, London.
- Lichfield N., A. Barbanente, D. Borri, A. Khakee, and A. Pratt, 1998. *Evaluation in planning - Facing the challenge of complexity*, Kluwer academic publishers, Dordrecht.
- Lichfield N., 2001. Where do we go from here? In: H Voogd (Editor), *Recent developments in evaluation in spatial, infrastructure and environmental planning*. Geo Press, Groningen, pp. 7-15.
- Lind R. C., 1995. Intergenerational equity, discounting, and the role of cost-benefit analysis in evaluating global climate policy. *Energy policy*, 23:379-389.
- Lind R. C., 1997. Intertemporal equity, discounting, and economic efficiency in water policy evaluation. *Climatic Change*, 41-62.
- Little I. M. D. and J. A. Mirlees, 1994. The costs and benefits of analysis. In: R Layard and S Glaister (Editors), *Cost-Benefit Analysis*. Cambridge University Press, Cambridge, pp. 199-231.
- Liu A., A. Collins, and S. Yao, 1998. A multi-objective and multi-design evaluation procedure for environmental protection forestry. *Environmental and Resource Economics*, 1998:225-240.
- LNV (Ministry of Agriculture Nature and Fisheries., 1990. *Natuurbeleidsplan - regeringsbeslissing*,

SDU publishers, The Hague.

LNV/VROM (Ministry of Agriculture Nature and Fisheries; Ministry of Housing, Spatial Planning and the Environment), 1993. Structuurschema Groene Ruimte, LNV, The Hague.

Loewenstein G. and D. Prelec, 2000. Anomalies in intertemporal choice - evidence and an interpretation. In: D Kahneman and A. Tversky (Editors), Choices, values and frames. Cambridge University Press, Cambridge, pp. 578-596.

Luchini S., 2002. De la singularité de la méthode d'évaluation contingente - The singularity of the contingent valuation method. *Economie et Statistique*, 2002:141-152.

Luini L. e., 1999. Uncertain decisions - Bridging theory and experiments, Kluwer Academic Publishers, Boston, Dordrecht. London.

Lundgren L. and P. Burnet, 1999. The tragedy of the commons revisited. *Environment*, 41:4-5.

MacCrimmon K. R., 1999. Decision theory and real decisions: a research agenda. In: Le Luini (Editor), Uncertain decisions - Bridging theory and experiments. Kluwer Academic Publishers, Boston, Dordrecht, London, pp. 263-291.

Maitre D. L., C. Gelderblom, L. Maphasa, S. Yssel, M. v. d. Belt, and T. Manuel, 1997. Communicating the value of fynbos: results of a survey of stakeholders. *Ecological Economics*, 1997:105-121.

Mann Borgese E., 1998. The oceanic circle: governing the seas as a global resource, United Nations University Press, Tokyo.

Martinez-Alier J., G. Munda, and J. O'Neill, 1998. Weak comparability of values as a foundation for ecological economics. *Ecological Economics*, 26:277-286.

McGuire J. B., A. Sundgren, and T. Schneeweis, 1988. Corporate social responsibility and firm financial performance. *Academy of Management Journal*, 31:854-872.

McWilliams A. and D. Siegel, 2001. Corporate social responsibility: a theory of the firm perspective. *Academy of Management Review*, 26:117-127.

Meadows D. L., 1972. Rapport van de Club van Rome (The limits to growth - a report for the Club of Rome project on the predicament of mankind), Het Spectrum, Utrecht.

Meppem T. and S. Bourke, 1999. Different ways of knowing: a communicative turn toward sustainability. *Ecological Economics*, 30 (1999):389-404.

Merkhofer M. W., 1987. Decision science and social risk management - A comparative evaluation of Cost-Benefit Analysis, decision analysis and other formal decision aiding approaches, D. Reidel Publishing Company, Dordrecht, Boston, Lancaster, Tokyo.

Michalos A. C., 1995. A pragmatic approach to business ethics, Sage Publications, Thousand Oaks.

Miller D., 2001. Evaluating environmental justice for planning. In: H Voogd (Editor), Recent developments in evaluation in spatial, infrastructure and environmental planning. Geo Press, Groningen, pp. 17-35.

Mishan E. J., 1975. Cost-benefit analysis - an informal introduction. George Allen & Unwin, London.

Mishan E. J., 1980. How valid are economic evaluations of allocative changes? *Journal of Economic Issues*, XIV:143-161.

Mishan E. J., 1981. Do economic evaluations of allocative changes have any validity in the West today? In: EJ Mishan (Editor), *Economic efficiency and social welfare - Selected essays on fundamental aspects of the economic theory of social welfare*. George Allen & Unwin, London, pp. 256-265.

Mishan E. J., 1981. *Economic efficiency and social welfare - selected essays on fundamnetal aspects of the economic theory of social welfare*, George Allen & Unwin, Boston.

Mishan E. J., 1982. The new controversy about the rationale of economic evaluation. *Journal of Economic Issues*, XVI:29-47.

Mishan E. J., 2001. What monsters technology hath wrought! *Business and Society Review*, 8.

Mishan E. J., 2002. Invited eminent paper series - on the conceptual underpinning of a cost-benefit analysis. *The Singapore economic review*, 47:1-16.

Moore M. A., A. E. Boardman, A. R. Vining, D. L. Weimer, and D. H. Greenberg, 2004. "Just give me a number!" Practical values for the social discount rate. *Journal of Policy Analysis and Management*, 23:789-812.

Morris P., R. Therivel, and (Eds.), 2001. *Methods of environmental impact assessment*, Spon Press, New York.

Morgan M. G. and H. Dowlatabadi, 1996. Learning from integrated assessment of climate change. *Climatic Change*, 34:368.

Moss B., D. C. Catt, N. G. Bayfield, and D. D. French, 1996. The application of decision theory to sustainable management of an upland Scottish estate. *Journal of applied statistics*, 23:211-229.

Munasinghe M., 1993. The economist's approach to sustainable development. *Finance & Development*, 16-19.

Munasinghe M., 2001. Towards sustainomics: a transdisciplinary meta-framework for making development and growth more sustainable. In: M Munasinghe, O Sunkel, and C De Miguel (Editors), *The sustainability of long-term growth. Socioeconomic and ecological perspectives*. Edward Elgar, Cheltenham, pp. 19-71.

Munasinghe M., O. Sunkel, C. De Miguel, and (eds), 2001. *The sustainability of long-term growth. Socioeconomic and ecological perspectives*, Edward Elgar, Cheltenham.

Munasinghe M., O. Canziani, O. Davidson et al., 2003. *Integrating sustainable development and climate change in the IPCC fourth assessment report*, Munasinghe Institute for Development (MIND), Colombo, Sri Lanka.

Munda G., 1993. *Fuzzy information in multicriteria environmental evaluation models*, Vrije Universiteit Amsterdam - PhD thesis, Amsterdam.

Munda G., 1995. *Multicriteria evaluation in a fuzzy environment: theory and applications in ecological economics*, Physica, Heidelberg.

- Murray C. J. L., D. B. Evans, A. Acharya, and R. M. P. M. Baltussen, 2000. Development of WHO guidelines on generalized cost-effectiveness analysis. *Health Economics*, 2000:235-251.
- Musters C. J. M., H. J. d. Graaf, and W. J. t. Keurs, 1998. Defining socio-environmental systems for sustainable development. *Ecological Economics*, 243-258.
- Nas T. F., 1996. *Cost-benefit analysis - theory and application*, Sage, Thousand Oaks.
- Nash C. A., D. Pearce, and J. K. Stanley, 1975. An evaluation of cost-benefit analysis criteria. *Scottish Journal of Political Economy*, XXII:121-134.
- Nederveen Pieterse J. and (ed.), 2000. *Global futures: shaping globalization*, Zed Books, London / New York.
- Ng Y.-K., 1983. *Welfare economics - introduction and development of basic concepts*, MacMillan, London.
- Niculae C. and S. French, 2003. Bringing understanding in societal decision making: explaining and communicating analyses? *Journal of multi-criteria decision analysis*, 2003:191-202.
- Niewijk R. K., 1994. Misleading quantification: the contingent valuation of environmental quality. *Regulation - The Cato Review of Business & Government*, 17:60-71.
- Nijkamp P. and A. Delft van, 1977. *Multi-criteria analysis and regional decision-making*, Nijhoff, Leiden.
- Nijkamp P., P. Rietveld, and H. Voogd, 1990. *Multicriteria evaluation in physical planning*, North-Holland, Amsterdam.
- Nijkamp P. and C. J. M. v. d. Bergh, 1997. New advances in economic modelling and evaluation of environmental issues. *European Journal of Operational Research*, 180-196.
- Noël J.-F., M. O'Connor, and J. Tsang King Sang, 2000. The Bouchereau woodland and the transmission of socio-ecological economic value. *Ecological Economics*, 2000:247-266.
- Norgaard R. B., C. Bode, and Values Reading Group, 1998. Next, the value of God, and other reactions. *Ecological Economics*, 37-39.
- Nowotny H., P. Scott, and M. Gibbons, 2001. *Re-thinking science. Knowledge and the public in an age of uncertainty*, Polity Press, Cambridge.
- Nunes P. A. L. D., J. C. J. M. v. d. Bergh, and P. Nijkamp, 2000. *Ecological-Economic Analysis and valuation of biodiversity*, Tinbergen Institute, Rotterdam/Amsterdam.
- Nyborg K., 1997. Some Norwegian politicians' use of cost-benefit analysis. *Public Choice*, 381-401.
- O'Connor M., C. L. Spash, and (eds), 1999. *Valuation and the Environment: theory, method and practice*, Edward Elgar, Cheltenham.
- O'Connor M., 2000. The VALSE project - an introduction. *Ecological Economics*, 2000:165-174.
- O'Connor M., 2000. Pathways for environmental evaluation: a walk in the (Hanging) Gardens of

- Babylon. *Ecological Economics*, 2000:176-193.
- O'Hara S. U., 1996. Discursive ethics in ecosystems, valuation and environmental policy. *Ecological Economics*, 95-107.
- OECD (Organisation for Economic Co-operation and Development), 1983. Product safety - Risk management and cost-benefit analysis, Author, Paris.
- OECD (Organisation for Economic Co-operation and Development), 1997. Investing in biological diversity - The Cairns conference, OECD (Organisation for Economic Co-operation and Development), Paris.
- Olson D. L., 1996. Decision aids for selection problems. Springer-Verlag, New-York.
- Oosterhaven J., 1982. Evaluating land reclamation plans for northern Friesland: an interregional cost-benefit and input-output analysis. Twenty-second European congress of the regional science association.
- Oosterhaven J. and J. P. Elhorst, 2003. Repliek: Indirecte effecten in kosten-batenanalyses van de Zuiderzeelijn - The devil is in the detail. *Tijdschrift Vervoerswetenschap*, 39:12-17.
- Oosterhaven J. and T. Knaap, 2003. Spatial Economic Impacts of Transport Infrastructure Investments. In: A Pearman, P Mackie, J Nellthorp, and (eds) (Editors), *Transport Projects, Programmes and Policies: Evaluation Needs and Capabilities*. Ashgate, Aldershot, pp. 87-105.
- Ophuls W. and A. S. j. Boyan, 1992. Ecology and the politics of scarcity revisited - the unraveling of the American dream, W.H. Freeman and Company, New York.
- Opschoor J. B., 1998. The value of ecosystem services: whose values? *Ecological Economics*, 41-43.
- Oskam A. J., 1989. Decision based economic theory. Wageningen Agricultural University, Wageningen.
- Oskam A. J., 1994. Het landbouw/natuur vraagstuk: economisch gezien., University of Wageningen - Vakgroep Algemene Agrarische Economie, Wageningen.
- Ozernoy V. M., 1997. Some fundamental problems in the selection and justification of discrete alternative MCDM methods. In: J Clímaco (Editor), *Multicriteria Analysis*. Springer, Berlin Heidelberg, pp. 103-112.
- Özelkan E. C. and L. Duckstein, 1996. Analysing water resource alternatives and handling criteria by multi criterion decision techniques. *Journal of Environmental Management*, 1996:69-96.
- Pannell D. J. and S. Schilizzi, 1999. Sustainable Agriculture: A matter of ecology, equity, economic efficiency or expedience? *Journal of Sustainable Agriculture*, 13:57-66.
- Patton M. Q., 1997. Utilization-focused evaluation - the new century text, Sage, Thousand Oaks.
- Pawson R. and N. Tilley, 1997. An introduction to scientific realist evaluation. In: E Chelmsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 405-418.
- Pearce D. and D. Moran, 1994. The economic value of biodiversity, IUCN - the World Conservation

Union and Earthscan Publications Ltd., London.

Pearce D., 1998. Economics and environment - Essays on Ecological Economics and sustainable development, Edward Elgar, Cheltenham.

Pearce D. and G. Atkinson, 1998. Capital theory and the measurement of sustainable development: an indicator of 'weak' sustainability. In: D Pearce (Editor), Economics and environment - Essays on Ecological Economics and sustainable development. Edward Elgar, Cheltenham, pp. 91-95.

Pearce D. W., 1998. Environmental appraisal and environmental policy in the European Union. Environmental and Resource Economics, 11:489-501.

Pelt van M., A. Kuyvenhoven, and P. Nijkamp, 1990. Project Appraisal and Sustainability: the Applicability of Cost-benefit and Multi-criteria Analysis, University of Wageningen, Wageningen.

Pethig R. and (ed.), 1994. Valuing the Environment: Methodological and Measurement Issues, Kluwer, Dordrecht.

Petitti D. B., 2000. Meta-analysis, decision analysis and cost-effectiveness analysis - Methods for quantitative synthesis in medicine, Oxford University Press, New York, Oxford.

Pezzey J., 1992. Sustainable Development Concepts - An economic analysis, World Bank, Washington.

Pierrehumbert R. T., 2003. Counting the cost - can cost-benefit analysis solve the problem of assessing environmental risk? Nature, 422.

Pimentel D., 1998. Economic benefits of natural biota. Ecological Economics, 45-47.

Pires C., 1998. Sustainability and cost-benefit analysis. Environment, 30:2181-2194.

Planco Consulting, 1997. Bewertung des Einbaus einer Schleuse in das geplante Ems-Sperrwerk bei Gandersum, Planco Consulting, Essen.

Plantinga A. and B. Scholtens, 2002. Socially responsible investing and the management style of mutual funds in the Euronext stock markets. Financieel Forum / Bank-en Financiewezen, 2002:248-254.

Poe G. L., J. E. Clark, D. Rondeau, and W. D. Schulze, 2002. Provision point mechanisms and field validity tests of contingent valuation. Environmental and Resource Economics, 105-131.

Polak J. B., A. Heertje, and (eds), 2000. Analytical transport economics, Edward Elgar, Cheltenham.

Pollak R., 1998. Imagined risks and cost-benefit analysis. American economic review, 88:376-380.

Pomerol J.-C. and S. Barba-Romero, 2000. Multicriteria decision in management - principles and practice, Kluwer Academic Publishers, Boston.

Poole M. S. and R. Y. Hirokawa, 1996. Introduction - Communication and group decision making. In: RY Hirokawa and MS Poole (Editors), Communication and group decision making. Sage, Thousand Oaks, pp. 3-18.

Pope J., D. Annandale, and A. Morrison-Saunders, 2004. Conceptualising sustainability assessment.

Environmental Impact Assessment Review, 595-616.

Porter T. M., 1995. Trust in numbers - the pursuit of objectivity in science and public life, Princeton University Press, Princeton, New Jersey.

Potts D., 1999. Forget the weights, who gets the benefits? How to bring a poverty focus to the economic analysis of projects. *Journal of International Development*, 581-595.

Quiggin J., 1997. Discount rates and sustainability. *International Journal of Social Economics*, 24:65-90.

Ranganathan J., 1998. Sustainability rulers: measuring corporate environmental and social performance. World Resources Institute, Washington.

Rauschmayer F., 1999. Decisions in the context of sustainable development: ethics and implementation of multi-criteria analysis. In: I Ring, B Klauer, F Wätzold, and BA Mansson (Editors), *Regional sustainability - Applied ecological economics - Bridging the gap between natural and social sciences*. Physica, Heidelberg, pp. 19-35.

Rauschmayer F., 2001. Reflections on ethics and MCA in environmental decisions. *Journal of multi-criteria decision analysis*, 65-74.

Ray A., 1984. Cost-Benefit Analysis - Issues and methodologies, The World Bank, Washington.

Redclift M., 2003. Sustainability and human security: challenges for the twenty-first century. *Journal of interdisciplinary economics*, 14:403-414.

Rees C., 1993. The ecologist's approach to sustainable development. *Finance & Development*, 14-15.

Rees W. A., 1998. How should a parasite value its host? *Ecological Economics*, 49-52.

Renn O., T. Webler, H. Rakel, P. Dienel, and B. Johnson, 1993. Public participation in decision making; A three-step procedure. *Policy Sciences*, 189-214.

Revesz R., 1999. Environmental regulation, cost-benefit analysis, and the discounting of human lives. *Columbia law review*, 99:941-1017.

Richardson H. S., 2000. The stupidity of the cost-benefit standard. *Journal of Legal Studies*, 2000:971-1002.

Rietveld P., J. Rouwendal, and A. Van der Vlist, 2002. Equity issues in the evaluation of transport policies and transport infrastructure projects., pp. 1-32.

Ring I., B. Klauer, F. Wätzold, B. A. Mansson, and (eds), 1999. *Regional sustainability - Applied ecological economics - Bridging the gap between natural and social sciences*, Physica, Heidelberg.

Rischar J. F., 2002. High Noon - 20 global problems, 20 years to solve them, Basic Books, New York.

Ritov I. and D. Kahneman, 1997. How people value the environment - Attitudes versus economic values. In: MH Bazerman, DM Messick, AE Tenbrunsel, and KA Wade-Benzoni (Editors), *Environment, ethics and behavior - The psychology of environmental valuation and degradation*. The New Lexicon Press, pp. 33-51.

- Robinson J., 1990. Philosophical origins of the social rate of discount in cost-benefit analysis. *The Milbank quarterly*, 68:245.
- Rogers B., 2000. The nature of value and the value of nature: a philosophical overview. *International affairs*, 76:315.
- Rosenhead J., J. Mingers, and (eds), 2001. *Rational analysis for a problematic world - Problem structuring methods for complexity, uncertainty and conflict*, John Wiley and Sons, Chichester.
- Rosenthal D. H. and R. H. Nelson, 1992. Why existence value should not be used in cost-benefit analysis. *Journal of Policy Analysis and Management*, 11:116-122.
- Rossi P. H., H. E. Freeman, and M. W. Lipsey, 1999. *Evaluation - A systematic approach*, 6 edition. Sage, Thousand Oaks.
- Rouwendaal J. and E. Verhoef, 2003. Indirecte effecten in kosten-batenanalyses van de Zuiderzeelijn. *Tijdschrift Vervoerswetenschap*, 39:2-11.
- Roy B., 1990. Decision-aid and decision-making. In: CA Bana e Costa (Editor), *Readings in multiple criteria decision aid*. Springer-Verlag, Berlin, Heidelberg, pp. 17-35.
- Roy B., 1990. The outranking approach and the foundations of the Electre methods. In: CA Bana e Costa (Editor), *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Heidelberg, pp. 155-183.
- Roy B. and D. Vanderpooten, 1996. The European school of MCDA: Emergence, basic features and current works. *Journal of Multi Criteria Decision Aid*, 22-38.
- Ruf B. M., K. Muralidhar, and et al., 1998. The development of a systematic, aggregate measure of corporate social performance. *Journal of management*, 24:119-134.
- Ruijgrok E. C. M., 1999. *Valuation of nature in coastal zones*, Vrije Universiteit Amsterdam - PhD thesis, Amsterdam.
- SACTRA (standing committee on trunk road assessment), 1999. *Transport and the economy*, Department of the Environment, Transport and the Regions; the Stationary Office, London, UK.
- Samuelson P. A., 1980. *Economics*, McGraw-Hill Kogakusha, Tokyo.
- Sarkar A. U., 1997. Sustainable development and technology. *The Environmentalist*, 97-102.
- Sassone P. G. and W. A. Schaffer, 1978. *Cost-Benefit Analysis - A handbook*, Academic press, New York.
- Scheper W. J., 1991. *Group decision support systems - an inquiry into theoretical and philosophical issues*, Katholieke Universiteit Brabant - Thesis, Tilburg.
- Schultz II C. J. and M. B. Holbrook, 1999. Marketing and the tragedy of the common: a synthesis, commentary and analysis for action. *Journal of public policy and marketing*, 18:218-230.
- Schütz J., 2000. Sustainability, Systems and Meaning. *Environmental Values*, 9:373-382.
- Scriven M., 1997. Truth and objectivity in evaluation. In: E Chelimsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 477-500.

- Sen A., 1982. Choice, welfare and measurement, MIT, Cambridge.
- Sen A. K., 1985. Well-being, Agency and Freedom. The Dewey lectures 1984. *Journal of Philosophy*, 169-221.
- SER: Sociaal Economische Raad, 2000. De winst van waarden. Advies over maatschappelijk ondernemen, SER, Den Haag.
- Serageldin I., 1993. Making development sustainable. *Finance & Development*, 6-10.
- Shadish W. R., 1995. Philosophy of science and the quantitative-qualitative debates: 13 common errors. *Evaluation and Program Planning*, 63-75.
- Shaw B., 2001. Economics and the environment: A "Land Ethic" critique of economic policy. *Journal of Business Ethics*, 51-57.
- Sheldrake R., 1996. Zeven experimenten die de wereld kunnen veranderen, Maarten Muntinga / Kosmos-Z&K Uitgeverij, Amsterdam / Utrecht.
- Shi T., 2004. Ecological economics: moving towards a transdisciplinary research on sustainability. *Journal of interdisciplinary economics*, 15:61-82.
- Siemons R. V., 2002. A development perspective for biomass-fuelled electricity generation technologies - Economic technology assessment in view of sustainability, University of Amsterdam - PhD thesis, Amsterdam.
- Sijtsma F. J., D. Strijker, m.m.v. A. v. Berg, and G. J. Rotmensen, 1995. Effect-analyse Ecologische Hoofdstructuur - Deel I: Hoofdrapport, Stichting Ruimtelijke Economie Groningen, Groningen.
- Sijtsma F. J., D. Strijker, m.m.v. A. v. Berg, and G. Rotmensen, 1995. Effect-analyse Ecologische Hoofdstructuur - Deel II: Natuurwaarde, Stichting Ruimtelijke Economie Groningen, Groningen.
- Sijtsma F. J., D. Strijker, and M. L. A. W. Hoefsloot, 1998. Duurzame ontwikkeling in het Waddengebied. Een methode voor het afwegen van economie, natuur, milieu en landschap, Wetenschapswinkel voor Economie, Groningen.
- Sijtsma F. J., M. Broekhof, and M. Nijboer, 2002. Maatschappelijk Verantwoord Ondernemen: de kern van de zaak is de sector. *SPIL*, 25-32.
- Sijtsma F. J. and M. Broekhof, 2002. Maatschappelijk Verantwoord Ondernemen: wat is dat? *SPIL*, 20-24.
- Sillanpää M., 1998. A New Deal for sustainable development in business: taking the social dimension seriously. *Greener Management International*, 93-116.
- Sison A. J. G., 2000. Integrated risk management and global business ethics. *Business Ethics: A European review*, 9:288-295.
- Slangen L. H. G., 1994. De financiële en economische aspecten van het omzetten van landbouwgronden in natuurterreinen. *Agrarisch Recht*, 54:313-318.
- Slocum R., 2004. Consumer citizens and the Cities for Climate Protection campaign. *Environment-*

and-planning-A, 36:763-782.

Smith V. K., 1986. A conceptual overview of the foundations of Benefit-Cost Analysis. In: JD Bentkover, VT Covello, J Mumpower, and (eds) (Editors), *Benefits Assessment - The state of the art*. D.Reidel Publishing Company, Dordrecht, pp. 13-34.

Smith V. K. and S. K. Pattanayak, 2002. Is meta-analysis a Noah's ark for non-market valuation? *Environmental and Resource Economics*, 271-296.

Soest D. v., 1998. *Tropical Deforestation*, University of Groningen, Groningen.

Söderbaum P., 1998. Economics and ecological sustainability: an actor-network approach to evaluation. In: N Lichfield, A Barbanente, D Borri, A Khakee, and A Pratt (Editors), *Evaluation in planning*. Kluwer Academic Publishers, Dordrecht, pp. 51-71.

Söderbaum P., 1999. Valuation as part of a microeconomics for ecological sustainability. In: M O'Connor and C Spash (Editors), *Valuation and the environment. Theory, method and practice*. Edward Elgar, Cheltenham.

Söderbaum P., 2004. Decision processes and decision-making in relation to sustainable development and democracy - where do we stand? *Journal of interdisciplinary economics*, 15:41-60.

Spash C. L., 2000. Ecosystems, contingent valuation and ethics: the case of wetland re-creation. *Ecological Economics*, 2000:195-215.

Squire L. and H. G. Van der Tak, 1975. *Economic analysis of projects*, Washington / Baltimore, The World Bank / John Hopkins University Press.

Stake R. E., 1997. Advocacy in evaluation - a necessary evil? In: E Chelimsky and WR Shadish (Editors), *Evaluation for the 21st century - a handbook*. Sage, Thousand Oaks, London, New Delhi, pp. 470-476.

Steg L., C. Vlek, D. Feenstra et al., 2001. *Towards a comprehensive model of sustainable corporate performance - three dimensional modelling and practical measurement (interim report - revised)*, University of Groningen, Groningen.

Stennek J., 1999. The expected consumer's surplus as a welfare measure. *Journal of Public Economics*, 265-288.

Stewart T. J., 1992. A critical survey on the status of multiple criteria decision making theory and practice. *Omega International Journal of Management Science*, 20:569-586.

Stewart T. J. and F. B. Losa, 2003. Towards reconciling outranking and value measurement practice. *European Journal of Operational Research*, 645-659.

Steyaert S. and H. Lisoir, 2005. *Participatory methods toolkit - a practitioner's manual*, King Baudouin Foundation/Flemish Institute for Science and Technology Assessment (vIwTA).

Stiglitz J. E., 1994. Discount rates: the rate of discount for benefit-cost analysis and the theory of the second best. In: R Layard and S Glaister (Editors), *Cost-Benefit Analysis*. Cambridge University Press, Cambridge, pp. 116-159.

Stockman F. N., M. A. L. M. v. Assen, J. v. d. Knoop, and R. C. H. v. Oosten, 2000. *Strategic*

decision making. *Advances in Group Processes*, 17:131-153.

Strijker D. and T. Deinum, 1992. Relative regional welfare differences in the EC. In: M. Tykkylainen and (ed.) (Editors), *Development Issues and Strategies in the New Europe*. Avebury, Aldershot, pp. 53-64.

Strijker D. and F. J. Sijtsma, 1996. Kosten en baten van het natuurbeleid. *Landeigenaar*, 42:20-23.

Strijker D. and F. J. Sijtsma, 1996. Drie effect-analyses van de ecologische hoofdstructuur. *Tijdschrift voor sociaal-wetenschappelijk onderzoek van de landbouw (TSL)*, 11:309-324.

Strijker D. and F. J. Sijtsma, 1996. Produktiviteit van EHS-grond. *De Landeigenaar*, 42:5-8.

Strijker D., F. J. Sijtsma and D. Wiersma, 2000. Evaluation of nature conservation - An application to the Dutch Ecological Network. *Environmental and Resource Economics*, 16:363-378.

Strijker D., F. J. Sijtsma and K. Bettels, 2000. Evaluating nature Conservation - The case of Meadow birds in the Netherlands. *Agricultural Economics Review*, 1:57-70.

Sugden R. and A. Williams, 1978. *The principles of practical cost-benefit analysis*, Oxford University Press, Oxford.

Sunstein C. R., 2004. Lives, life-years and willingness to pay. *Columbia law review*, 2004:205-252.

Sutton Ph., 1998. The sustainability promoting firm. *Greener Management International*, 127-152.

Svenson O., 1998. Multi-criteria decision aids and human decision making: two worlds? *Journal of multi-criteria decision analysis*, 352-354.

Tavasszy L. and G. Renes, 2003. Overzicht Effecten Infrastructuur (OEI) en Indirecte Effecten: elementen voor een handleiding. *Tijdschrift Vervoerswetenschap*, 39:28-34.

Teisman G. R., 1998. Complexe besluitvorming - Een pluricentrisch perspectief op besluitvorming over ruimtelijke investeringen, Elsevier Bedrijfsinformatie, 's Gravenhage.

Tell B., 1978. Factor analysis - a method of reducing the number of criteria in a Multi-criteria model. *Omega International Journal of Management Science*, 8:451-454.

Templet P. H., 1998. The neglected benefits of protecting ecological services. *Ecological Economics*, 53-55.

Thaler R. H., 2000. Mental accounting matters. In: D Kahneman and Ae Tversky (Editors), *Choices, values and frames*. Cambridge University Press, Cambridge, pp. 241-268.

Therivel R. and M. R. Partidário, 1996. *The practice of strategic environmental assessment*, Earthscan Publications, London.

Tinbergen J., 1970. *Een leefbare aarde*, Agon Elsevier, Amsterdam.

Tiwari D. N., 2000. Sustainability Criteria and Cost-Benefit Analysis: An Analytical Framework for Environmental-Economic Decision Making at the Project Level. *Environment and development economics*, 5:259-288.

- Toman M., 1998. Why not to calculate the value of the world's ecosystem services and natural capital. *Ecological Economics*, 57-60.
- Toman M. A., 1999. Sustainable Decision-Making: The State of The Art from an Economics Perspective. In: M O'Connor and CL Spash (Editors), *Valuation and the Environment: theory, method and practice*. Edward Elgar, Cheltenham.
- Tóth F. L. and (ed.), 1997. *Cost-Benefit Analysis of Climate Change: The Broader Perspective*, Birkhäuser, Basel.
- Tóth F. L., 2000. Intergenerational equity and discounting . *Intergrated Assessment*, 1:127-136.
- Turner R. K., W. N. Adger, and R. Brouwer, 1998. Ecosystem services value, research needs, and policy relevance: a commentary. *Ecological Economics*, 61-65.
- Turner R. K., K. Button, and P. e. Nijkamp, 1999. *Ecosystems and nature - Economics, science and policy*. Edward Elgar, Cheltenham.
- Turner R. K., C. J. M. v. d. Bergh, T. Söderqvist, A. Barendregt, J. v. d. Straaten, E. Maltby, and E. C. v. Ierland, 2000. The values of wetlands: landscape and institutional perspectives - Ecological-economic analysis of wetlands: scientific integration for management and policy. *Ecological Economics*, 35:7-23.
- Turner R. K., C. J. M. v. d. Bergh, T. Söderqvist, A. Barendregt, J. v. d. Straaten, E. Maltby, and E. C. v. Ierland, 2005. Ecological-economic analysis of wetlands: scientific integration for management and policy. *Ecological Economics*, 2000:7-23.
- Tversky A. and D. Kahneman, 2000. Advances in prospect theory. In: D Kahneman and Ae Tversky (Editors), *Choices, values and frames*. Cambridge University Press, Cambridge.
- Tversky A. and I. Simonson, 2000. Context-dependent preferences. In: D Kahneman and Ae Tversky (Editors), *Choices, values and frames*. Cambridge University Press, Cambridge, pp. 518-527.
- Urli B. and R. Nadeau, 1999. Evolution of multi-criteria analysis: a scientometric analysis. *Journal of multi-criteria decision analysis*.
- Vansnick J.-C., 1990. Measurement theory and decision aid. In: CA Bana e Costa (Editor), *Readings in Multiple Criteria Decision Aid*. Springer-Verlag, Berlin, pp. 81-100.
- Vatn A. and D. W. Bromley, 1994. Choices without prices, without apologies. *Journal of Environmental Economics and Management*, 129-148.
- Veen A. v. d. and H. S. Otter, 2002. Scales in space. *Integrated Assessment*, 3:160-166.
- Veisten K., H. F. Hoen, S. Navrud, and J. Strand, 2004. Scope insensitivity in contingent valuation of complex environmental amenities. *Journal of Environmental Management*, 2004:317-331.
- Vercelli A., 1999. The recent advances in decision theory under uncertainty - a non-technical introduction. In: Le Luini (Editor), *Uncertain decisions - Bridging theory and experiments*. Kluwer Academic Publishers, Boston, Dordrecht, London, pp. 237-260.
- Vickers G., 1965. *The art of judgement - A study of policy making*, Harper & Row, London.
- Vincke Ph., 1992. *Multicriteria decision-aid*, John Wiley & Sons, Chichester.

Vlek C. A. J., E. M. Steg, D. W. Feenstra, W. Gerbens-Leenis, S. Lindenberg, H. Moll, T. Schoot Uiterkamp, F. J. Sijtsma, and A. v. Witteloostuijn, 2002. Een praktisch model voor duurzaam bedrijfspresteren. *Economisch Statistische Berichten*, 87:524-527.

Vogtländer J. G., 2001. The model of the Eco-Costs/Value ratio - A new LCA based decision support tool, Technische Universiteit Delft - PhD thesis, Delft.

Vollebergh H. R. J., 1999. Milieu en schaarste - Over draagwijdte en toepassingsmogelijkheden van milieu-economische analyse, Erasmus Universiteit - PhD thesis, Rotterdam.

Voogd H., 1983. Multicriteria evaluation for urban and regional planning, Pion, London.

Voogd H., 2001. Recent developments in evaluation, Geo Press, Groningen.

Voogd H., 2004. Formal evaluation in environmental and infrastructure planning. In: G Linden and H Voogd (Editors), *Environmental and infrastructure planning*. Geo Press, Groningen, pp. 203-230.

Vreeker R. and P. Nijkamp, 2001. Advances in multicriteria decision support - Methods for evaluating development scenarios; An application to Thailand. Vrije Universiteit Amsterdam, Faculteit der Economische wetenschappen en econometrie, Amsterdam.

Vries M. S. d., 1993. Calculeren met beleid - theorie en praktijk van multicriteria-evaluaties, Van Gorcum, Assen.

Wackernagel M., N. B. Schultz, D. Deumling, A. C. Linares, M. Jenkins, V. Kapos, C. Monfreda, J. Loh, N. Myers, R. Norgaard, and J. Randers, 2002. Tracking the ecological overshoot of the human economy. *PNAS*, 99:9266-9271.

Waddock S. and S. B. Graves, 1997. The corporate social performance - financial performance link., pp. 303-319.

Waddock S. and S. B. Graves, 1997. Finding the link between stakeholder relations and quality of management. *Journal of Investing*, 6:20-25.

Waddock S. and N. Smith, 2000. Relationships: the real challenge of corporate global citizenship. *Business and Society Review*, 105:47-62.

Wal v. d. R., 1986. Toegevoegde Waarde. In: J. Dijkma (Editor), *Jaar in - jaar uit Verslag van een empirisch onderzoek van jaarrapporten over de jaren 1983 en 1984*. Wolters-Noordhof, Groningen.

Ward H., A. Norval, T. Landman, and J. Pretty, 2003. Open citizens' juries and the politics of sustainability. *Political Studies*, 51:282-299.

Water directors EU, 2002. EU water framework directive guidance document - guidance document on public participation in relation to the water framework directive.

WBCSD: World Business Council for Sustainable Development, 1999. Meeting changing expectations: corporate social responsibility, WBCSD, Conches-Geneva.

WBCSD: World Business Council for Sustainable Development, 2000. Corporate Social Responsibility: making good business sense, WBCSD, Geneva.

- WCED (World Commission on Environment and Development), 1987. *Our common future*, Oxford University Press, Oxford.
- Weiss C. H., 1988. Evaluation for decisions - Is anybody there? Does anybody care? *Evaluation Practice*, 9:5-19.
- Weiss C. H., 1990. Evaluation Utilization. In: MC Alkin, MQ Patton, CH Weiss, and R Connor (Editors), *Debates on evaluation*. Sage, Newbury Park.
- Weizsäcker E. U. v. and et al., 1997. *Factor four: doubling wealth - halving resource use*, Earthscan, London.
- Welle P. G., 1990. Benefit, cost, and beyond: the political economy of Benefit-Cost Analysis. *Atlantic Economic Journal*, 18:79-90.
- Wenstop F. and K. Seip, 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*, 10:53-64.
- Wenum J. H. v., 2002. *Economic analysis of wildlife conservation in crop farming*, Wageningen University - PhD thesis, Wageningen.
- Wenum J. v., A. Wossink, and J. Buys, 1999. Nature quality indicators in agriculture. In: F Brouwer and JR Crabtree (Editors), *Environmental indicators and agricultural policy*. Oxford University Press, Oxford.
- Wenum J. v., 2002. *Economic Analysis of Wildlife Conservation in Crop Farming.*, PhD thesis Wageningen University, Wageningen.
- Werkgroep Financiële dekking SGR, 1993. *Financiële achtergronden Structuurschema Groene Ruimte*, The Hague.
- Wiedeman P., 1978. Planning with multiple objectives. *Omega International Journal of Management Science*, 6:427-431.
- Wierstra E., P. Geurts, and A. v. d. Veen, 2001. Validity of CVM related to the type of environmental good: an empirical test. *Integrated Assessment*, 1-16.
- Wilbanks T. J., 2002. Geographic scaling issues in integrated assessments of climate change. *Integrated Assessment*, 3:100-114.
- Wildavsky A., 1966. The political economy of efficiency: cost benefit analysis, systems analysis, and program budgeting. *Public Administration Review*, 26:292-310.
- Willis K. G., K. Button, P. Nijkamp, and (eds), 1999. *Environmental Valuation - Volume I: methods and anomalies*, Edward Elgar, Cheltenham.
- Winterfeldt D. v. and W. Edwards, 1986. *Decision analysis and behavioral research*, Cambridge University Press, Cambridge.
- Witteloostuijn A. v., 1999. *De anorexiastrategie - over de gevolgen van saneren*, De Arbeiderspers, Amsterdam - Antwerpen.
- Woltjer J., 2001. *Evaluating complex decision making processes: an assessment of Dutch*

- communicative planning initiatives. In: H Voogd and (ed.) (Editors), Recent developments in evaluation in spatial, infrastructure and environmental planning., pp. 85-109.
- Woltjer J., 2004. Consensus planning in infrastructure and environmental development. In: G Linden and H Voogd (Editors), Environmental and infrastructure planning. Geo Press, Groningen, pp. 37-57.
- Wood D. J., 1991. Corporate social performance revisited. *Academy of Management Review*, 16:691-718.
- Worldwatch Institute, 2005. State of the world 2005 - Redefining global security.
- Wossink A., J. v. Wenum, C. Jurgens, and G. d. Snoo, 1999. Co-ordinating economic, behavioral and spatial aspects of wildlife preservation in agriculture. *European Review of Agricultural Economics*, 26:443-460.
- WRR (Netherlands Scientific Council for Government Policy), 1992. Grond voor keuzen - Vier perspectieven voor de landelijke gebieden in de Europese Gemeenschap, SDU uitgeverij, The Hague.
- Yandle B., 1998. Antitrust and the commons - cooperation or collusion? *Independent review*, 3:37-53.
- Yates J. F., 1990. Judgement and decision making. Prentice Hall, Englewood Cliffs.
- Yew-Kwang Ng, 1983. Welfare Economics, MacMillan, Houndsmill, Basingstoke, Hampshire, London.
- Yoon K. P. and C. L. Hwang, 1995. Multiple attribute decision making, Sage Publications, Thousand Oaks London New Delhi.
- Yoon K. P. and C. L. Hwang, 1995. Multiple attribute decision making, Sage Publications, Thousand Oaks London New Delhi.
- Yu P.-L., 1985. Multiple-criteria decision making - concepts techniques and extensions, Plenum Press, New York.
- Zeckhauser R. J., R. L. Keeney, and J. K. Sebenius, 1996. Wise choices - Decisions, games and negotiations, Harvard Business School Press, Boston, Massachusetts.
- Zerbe R. O. and D. D. Dively, 1994. Benefit-cost analysis in theory and practice, HarperCollinsCollegePublishers, New York.
- Zerbe R. O., 1998. Is cost-benefit analysis legal? Three rules. *Journal of Policy Analysis and Management*, 17:419-456.
- Zerbe R. O. and H. E. McCurdy, 1999. The failure of market failure. *Journal of Policy Analysis and Management*, 18:558-578.
- Zey M. and (ed.), 1992. Decision making - Alternatives to rational choice models, Sage publications, Newbury Park, California.
- Zimmerman H.-J. and L. Gutsche, 1991. Multi-Criteria Analyse, Springer, Berlin.
- Zuidema T., 1982. Een onderzoek naar de alternatieve kosten van overheidsprojecten: theorie en empirie, Rijksuniversiteit Groningen - PhD thesis, Groningen.

SAMENVATTING

Duurzaam, verantwoord

Wat zijn de kosten en baten van het realiseren van een duurzaam nationaal netwerk van natuurgebieden op het land van Nederlandse boeren? Welke gevolgen heeft het om een waterkering in de Duitse getijde rivier de Ems aan te leggen voor de bouw van reusachtige cruise schepen in Papenburg - ver van zee? Kan Ahold duurzamer of meer maatschappelijk verantwoord ondernemen door een nieuwe Europese productie standaard voor duurzame groente en fruit te ontwikkelen?

Dit proefschrift gaat over evalueren, het op waarde schatten, van projecten – zoals bovenstaande - waarbij de duurzaamheid van het project in het geding is (sustainability) en waarbij tevens velerlei groepen betrokken zijn die of meekijken met de beslisser of half meebeslissen (accountability). Kort geformuleerd, noemen we dit ‘evalueren in een duurzaamheidscontext’.

De grote lijn van dit proefschrift

Om een oordeel te vellen over verschillende plannen, projecten en varianten van projecten kunnen evaluatiemethoden helpen. De twee belangrijkste methoden om projecten ‘op waarde te schatten’ zijn de Kosten Baten Analyse (KBA) en de Multi Criteria Analyse (MCA).

Het proefschrift gaat na hoe de KBA en MCA presteren in de duurzaamheidscontext (hoofdstuk 3 en 4). In het proefschrift wordt voor het eerst een combinatie van beide methoden geïntroduceerd: 'Multi Criteria Kosten Baten Analyse' (MCKBA) genoemd¹. Een dergelijke combinatie is wel eens eerder voorgesteld en soms ook wel eens een beetje toegepast, maar in het overgrote deel van de evaluaties is het of het een of het ander. Het is niet overdreven te zeggen dat de werelden van KBA en MCA grotendeels ‘aparte werelden’ zijn. Dit proefschrift gaat – overigens op een niet technische manier - diep in op beide methoden en constateert dat beide methoden in de duurzaamheidscontext heel goed te combineren zijn. Het hoe, wat en waarom van die combinatie staat in hoofdstuk 5, en dat hoofdstuk begint na zo’n 120 bladzijden.

De rest van het proefschrift, na hoofdstuk 5 dus, bespreekt een drietal ‘cases’: concrete evaluaties die zijn uitgevoerd met de nieuw ontwikkelde combinatie van methoden; MCKBA dus. In die cases gaat het om het evalueren van de aanleg van de Ecologische Hoofdstructuur, om de bouw van het EmsSperrwerk en hoe het zou moeten met Ahold en duurzaam – maatschappelijk verantwoord – ondernemen. In dat resterende deel kan men dus zien hoe die MCKBA methode in de praktijk werkt.

De toepassingen waren MCKBA ‘avant la lettre’ en ze vormen daarom tegelijkertijd directe inspiratie voor de algemene MCKBA. Het proefschrift had meer cases kunnen bespreken dan de huidige drie. Deze drie zijn gekozen omdat de meest in het oog springende kenmerken van de verschillende evaluaties de essentie van de MCKBA goed illustreren.

In de Ecologische Hoofdstructuur (hoofdstuk 7) evaluatie is het meest opvallend de combinatie van monetaire en niet-monetaire meetmethoden, waarbij de uitgebreide kwantitatieve evaluatie van de impact op de natuurwaarde in Nederland is gecombineerd met een beperkte KBA van gemakkelijk op geld te waarderen impacts zoals het verlies aan

¹ Het proefschrift is in het Engels geschreven en gebruikt daarom de Engelse afkortingen: CBA, MCA en MCCBA. Omdat de afkorting KBA in Nederland tamelijk ingeburgerd is gebruiken we hier Nederlandse pendanten.

landbouwproductie.

In de Emssperrwerk evaluatie (hoofdstuk 8) is het meest typerend de evaluatie op meerdere ruimtelijke schaalniveaus: Niedersachsen, Duitsland en mondiaal. De economische baten van de bouw van het Emssperrwerk (en dus behoud van de werf in Papenburg) bleken vooral op te wegen tegen de kosten als enkel het laagste schaalniveau (Niedersachsen) werd beschouwd. Op Duits niveau was verplaatsing van de werf naar Emden of de Eemshaven veel wenselijker; economisch maar ook vanwege de kansen voor de natuur. Op globaal niveau is de afruil tussen economie en natuur misschien nog het scherpst zichtbaar; verplaatsing van de activiteit naar een bestaande werf buiten Duitsland of de Eemshaven liet dan in ieder geval de laagste maatschappelijke kosten zien.

In de evaluatie van Ahold en duurzaam ondernemen (hoofdstuk 9) is het meest opvallend het belang van een goede causale analyse voor het scherp vormgeven van de evaluatie. Door een analyse van de (economische) causaliteit kon worden voorkomen dat de nadruk teveel op bedrijfsniveau komt te liggen en dat duurzaamheid een veel te ruim begrip wordt om zinvol te evalueren. De case laat zien dat de kern van duurzaam ondernemen en het evalueren daarvan ligt op sector of markt niveau en op het initiatief nemen om op dat niveau randvoorwaarden voor de competitie op te zetten die duurzaamheid dienen. Een voorbeeld van een dergelijke activiteit is EurepGAP; een wereldwijde productiestandaard c.q. -protocol voor vers producten.

Het zijn deze drie cases die tegelijkertijd testcases en inspiratiebronnen zijn voor de MCKBA waar dit proefschrift om draait.

Dit deel van het proefschrift – met de case studies - eindigt met hoofdstuk 10, waarin specifieke lessen vanuit de praktijk op een rij worden gezet en waarin ook enige afsluitende reflecties ten aanzien van het gehele proefschrift worden gegeven.

De inhoud van dit proefschrift aan de hand van 11 stellingen

De belangrijkste resultaten van dit proefschrift zullen hierna worden weergegeven aan de hand van 11 stellingen. Echter, om de kern van dit proefschrift – en dus de stellingen - te kunnen begrijpen moeten eerst drie zaken kort worden toegelicht: duurzaamheid en verantwoording afleggen, evalueren met KBA en evalueren met MCA.

Duurzaamheid en verantwoording afleggen

Voor evalueren is het idee van verantwoording afleggen, aan betrokkenen of aan een breed publiek, duidelijk genoeg, maar wat is duurzaamheid? In tegenstelling tot vele anderen acht ik in dit proefschrift het definiëren van duurzaamheid niet zo ingewikkeld. Uit de analyse van de kernelementen van duurzaamheid blijkt bovendien dat duurzaamheid en het verantwoording afleggen aan belanghebbenden c.q. betrekken van belanghebbenden bij evaluaties elkaar versterken.

Voor het evalueren binnen de ‘duurzaamheids context’ op een manier die de ‘openheid’ naar belanghebbenden versterkt identificeert dit proefschrift de volgende kernelementen:

- ‘Triple E’ beoordeling² (dat is evaluatie van effecten op economische ontwikkeling, extreme armoede en achteruitgang van het milieu).
- beoordeling op mondiaal schaalniveau en op het niveau van besluitvorming

² Triple E komt van het Engelse ‘Economic development, Extreme poverty and Environmental degradation’.

- beoordeling van effecten in de tijd over een lange periode,
- beoordeling die een ieders begrip vergroot (ook het begrip van de beperkingen die verbonden zijn aan de beoordeling).

Deze kernelementen moeten volgens dit proefschrift standaard onderdelen zijn van een project evaluatie in de duurzaamheidscontext.

>>>Lees meer in H2 en met name in 2.3

Evalueren met Kosten-Baten Analyse

De Kosten-Baten Analyse (KBA) probeert alle effecten van een project op geld te waarderen. Recent is dat in Nederland bijvoorbeeld gebeurd voor de Zuiderzeelijn varianten. Vaak is het op geld waarderen voor een deel van de kosten niet zo'n probleem: die zijn al in geld uitgedrukt. Voor veel batenposten en een ander deel van de kosten ligt dat anders. Hoe waardeer je bij de Zuiderzeelijn varianten bijvoorbeeld milieueffecten of een andere verdeling van de werkgelegenheid binnen Nederland?

Het in geld waarderen van effecten binnen de KBA is gebaseerd op het begrip van markten en hun werking: hoe het een loonsverhoging in Maleisië uiteindelijk als kostenstijging terecht komt bij de koper van auto's met rubberbanden bijvoorbeeld, en hoe de voorkeuren van een hele hoop individuen uiteindelijk tot uitdrukking komen in een prijs op een markt. In een KBA worden niet willekeurig marktprijzen gebruikt, maar op basis van secuur redeneren en het (in)schatten van reacties van consumenten en producenten worden verschillende effecten op geld gewaardeerd – met vaak marktdata (prijzen, hoeveelheden, elasticiteiten) als achtergrond. Kosten en baten die verspreid zijn in de tijd, worden op één noemer gebracht door te 'disconteren': door een rentevoet te hanteren gedurende de beschouwde periode. De KBA bepaalt een rangorde van projectalternatieven door het saldo van (verdisconteerde) kosten en baten te hanteren. Het alternatief met het hoogste saldo is favoriet.

>>>Lees meer in 3.1 en uitgebreider in 3.2

Evalueren met Multicriteria analyse

De multicriteria analyse helpt beslissers bij het maken van complexe keuzeproblemen door de verschillende gevolgen van verschillende opties uiteen te rafelen, in verschillende dimensies of criteria, de effecten op die dimensie te meten en die vervolgens van een waardering te voorzien. MCA werkt met meerdere criteria die elk een bepaald aspect van de gevolgen van een project meten. Misschien wel de bekendste MCA's zijn die van de Consumentenbond die leiden tot bijvoorbeeld de 'beste koop', maar MCA heeft vele toepassingsgebieden. Denken we opnieuw aan de Zuiderzeelijn dan zou de MCA bijvoorbeeld de landschappelijke impact als een criterium kunnen hanteren, het energiegebruik als een ander en het aantal reizigers als weer een ander; enzovoort. Om een eindrangorde te bepalen worden alle criteria op een gestandaardiseerde manier gemeten (zeg van 1-100) en worden de criteria veelal van gewichten voorzien die hun belangrijkheid aangeven. Op deze manier wordt een 'index'cijfer bepaald. Hoe hoger deze index hoe beter.

>>>Lees meer in 4.1 en uitgebreider in 4.2

1. Voor projectevaluatie in de duurzaamheidscontext is MCKBA een evenwichtige benadering, al reikt ze qua oordeelsvorming misschien niet zover als men zou willen. De KBA is te streng en te arm en de MCA te flexibel en te rijk; de combinatie van beide methoden is veel beter.

Om in een duurzaamheidscontext DE gevolgen voor DE sociale welvaart te willen beoordelen (zoals in een KBA) of ALLE zorgen en belangen van ALLE betrokkenen (zoals vaak in een MCA) is een onmogelijke opdracht.

De MCKBA erkent deze realiteit; niet het bepalen van de beste keuze is de taak van de evaluatie methode in de duurzaamheidscontext, maar het ondersteunen van beslissingen. Het gaat om zaken als: het scheppen van enige orde in een complexe besluitvormingssituatie; het beter structureren en begrijpen van de keuze situatie; het wegstrepen van slechte opties en het identificeren van opties die elkaar niet veel ontoppen; alsmede om het vinden van (nieuwe) alternatieven.

Het maximum dat een evaluator derhalve kan bereiken in de duurzaamheidscontext is onnodige onenigheid voorkomen en een verhelderend perspectief geven op een complex probleem. Een volledig oordeel of een complete rangorde van alternatieven is veelal niet aan de orde.

>>> Lees meer in 5.2; Verder in het KBA hoofdstuk in 3.3 en in het MCA hoofdstuk in 4.5

2. Een op ‘consensus gebaseerde aggregatie’ van impacts geeft een goed fundament voor evaluatie in de duurzaamheidscontext. Wat gemakkelijk in geld kan worden gewaardeerd komt in een beperkte KBA; voor de rest worden impacts geaggregeerd binnen een MCA stramien en qua criteria(structuur) aansluitend bij ‘jointly held objective conceptions’.

Het fundament van een evaluatie methode moet niet in een theorieboek worden gezocht. Het echte fundament van zowel de KBA als de MCA is de consensus onder betrokkenen dat de toepassing en meetmethoden van de evaluatietechnieken belangrijk en betekenisvol zijn.

Het echte fundament van de KBA is niet zo zeer de vaakgenoemde welvaartstheorie maar het is veeleer – de overigens niet vanzelfsprekende - consensus onder betrokkenen over het belang van in geld uitgedrukte rendementen, efficiency, budgetten en kosteninschattingen. Voor zover een evaluatie kan aansluiten bij de realiteit van markten, budgetten en kosten waar veel betrokkenen bekend mee zijn, dient men dit te doen.

Voor de MCA is het belang van consensus in feite hetzelfde. In de MCA literatuur wordt echter de basis van een consensus onder betrokkenen volop erkend. Zo staat in de MCA het zoeken naar ‘gezamenlijk gevonden objectieve concepten’ en het vinden van een door alle betrokken geaccepteerde ‘criteria-structuur’ centraal.

>>> Lees meer in 3.4.2 en 4.2.3

3. De ‘consensus’ in een MCKBA betreft de minimale relevantie van betrokken criteria en de begrijpelijkheid van meetmethoden. Ze betekent niet dat iedereen het over alles eens moet zijn.

Met name uit de case studies is gebleken dat de consensus in de evaluatie met MCKBA niet zozeer de compleetheid van de analyse zal betreffen. Een deel van de betrokkenen zal kunnen (blijven) vinden dat er meer aspecten in de analyse hadden moeten worden betrokken. De consensus dient met name te ontstaan over de relevantie van de wel gebruikte criteria en het relatieve gemak waarmee men de manier van meten van de impacts kan begrijpen.

De genoemde consensus over de relevantie van wel gebruikte criteria zou men kunnen uitleggen als consensus over het precieze gewicht van een criterium in relatie tot andere criteria. Een dergelijke interpretatie is logisch vanuit een ‘decision making’ (beslissing-nemend) perspectief op evalueren met MCA. Een dergelijke uitleg wordt hier echter niet bedoeld. Het gaat hier om relevantie die aangeeft dat het criterium een minimum niveau van belangrijkheid – voor iedereen – heeft. Een dergelijke interpretatie past in een ‘decision aiding’ perspectief: het doel van evalueren is het ondersteunen van beslissingen door structureren, inkaderen en meten.

>>>Lees meer in 7.3.6 en 7.3.7 (en 10.2)

4. Bij het betrekken van belanghebbenden bij een projectevaluatie dient de breedte van perspectieven – d.w.z. het bijeen brengen van mensen met sterk verschillende visies op de zaak - voorop te staan.

Het betrekken van belanghebbenden bij projectevaluaties in de duurzaamheidscontext is een tegenstrijdig iets. Enerzijds is het ronduit wenselijk: voor betrokkenheid, voor draagvlak, omdat twee meer weten dan één, en omdat stakeholders toegang kunnen geven tot cruciale empirische gegevens of hun eigen waarnemingen. Anderzijds is de lijst met wenselijke en mogelijke betrokkenen in een duurzaamheidscontext al snel aanzienlijk uitgebreid. Serieus mensen erbij betrekken is dan organisatorisch lastig of vrijwel onmogelijk en het heeft vaak negatieve gevolgen voor de kern van evaluatie: de criteria structuur.

Dit proefschrift bepleit een praktisch compromis. Het zoveel mogelijk betrekken van belanghebbenden, maar met name het verkrijgen van een breedte van inzichten binnen het ruimere evaluatie team: een dergelijke breedte is noodzakelijk voor het kunnen beproeven van de mate van bovengenoemde consensus elementen (zie consensus 1 en 2).

Het inhoudelijk zinvol en organisatorisch slim betrekken van een beperkte groep stakeholders is echter wel een ‘open eind’ in dit proefschrift en een onderwerp van vervolg onderzoek.

>>>Lees meer in 4.3 (inzichten over betrekken van stakeholders komen met name uit de MCA literatuur)

>>>Lees verder - vanuit de case studies - meer in 10.2 maar ook in 7.3.3 in 8.3.3 en in 9.3.3.

5. Hoe meer effecten van een project over ‘burger-waarden’ gaan hoe ongeschikter de KBA; hoe meer het gaat om ‘consument-waarden’ hoe geschikter de KBA.

De mogelijkheden van de KBA om veranderingen in *de sociale welvaart* te kunnen bepalen met *de meetlat van het geld* worden vaak overdreven ‘Geld maakt niet gelukkig’ zegt men in het Nederlands om de problematische relatie tussen geld en geluk aan te duiden en in het Engels zegt men ‘Best things in life are free’; vele waardevolle – fundamentele – zaken worden niet op markten gewaardeerd.

In de duurzaamheidscontext staat extreme armoede en mogelijk ernstige aantasting van het milieu centraal. Dit proefschrift bevat een nauwkeurige analyse van zowel de theoretische basis van de KBA als de waaier aan praktische monetaire instrumenten die beschikbaar zijn, waaronder instrumenten die expliciet vragen naar de Willingness To Pay; de bereidheid te betalen. Deze zogenaamde Contingent Valuation Methods worden door velen als belangrijk en vol potentie gezien om bijvoorbeeld de waardering voor milieuveranderingen te meten. De analyse in dit proefschrift laat de twee kernproblemen van deze instrumenten helder zien. Ten eerste veronderstelt CVM dat de voorkeuren van de respondenten gevormd zijn; terwijl het vaak gaat om zaken waarbij men ‘leert hoe men moet denken’; een realiteit die MCA technieken veel meer onderkennen. De nieuwste CVM technieken bewegen zich dan ook in de richting van het *ontdekken* van preferenties i.p.v. het simpelweg vragen van een waardering. Ten tweede laat de analyse in dit proefschrift zien dat extreme armoede en ernstige aantasting van het milieu geen zaken zijn die gemakkelijk in een KBA kunnen worden gewaardeerd; voor een deel omdat ze ‘te fundamenteel’ zijn. Dit kan men o.a. aflezen aan de onoplosbare moeilijkheden om voor milieu en armoede impacts zinvolle discontovoeten te bepalen.

Om in het algemeen de grens aan te geven waar de KBA ‘meetlat van het geld’ wel en niet zinnig te hanteren is maakt dit proefschrift gebruik van de zogenaamde ‘tweevoudige preferentie structuur’ van burger-consument. Gaat het om burger-preferenties dan is de KBA minder geschikt; hoe meer het gaat om consumentenvoorkeuren hoe geschikter de KBA.

>>>Lees meer in 3.5.5.2 en in 3.3.3.2b

6. In de duurzaamheidscontext is de kracht van de KBA veel meer gelegen in de strenge causaliteitsanalyse dan in het ‘op geld waarderen’ waar de KBA gewoonlijk mee wordt geassocieerd.

Het is gebruikelijk om de KBA te karakteriseren aan de hand van haar monetaire waardering. Voor het toepassen van een KBA is echter een goed begrip van de werking van markten en de causale relaties tussen markten en gedrag van individuen en bedrijven noodzakelijk: dit alles om ‘dubbeltelling’ van effecten te voorkomen. Nauwkeurige beschouwing van de KBA *in vergelijking met de MCA* in dit proefschrift leert dat juist die grote nadruk op een correcte analyse van de causaliteit en het daarna correct meten van impacts een belangrijker kenmerk van de KBA is dan haar monetaire gerichtheid.

>>>Lees meer in 3.4.2.2c-e

7. In de duurzaamheidscontext is de flexibiliteit van de MCA bruikbaar maar vaak ook te groot. Dit laatste kan in belangrijke mate verholpen worden door meer aandacht te schenken aan causale verbanden tussen de criteria en alleen de belangrijkste criteria te benadrukken.

De kracht van de MCA is haar grote flexibiliteit. De standaard elementen van de MCKBA (Triple E, globaal schaalniveau e.d.) rusten op deze eigenschap. Deze flexibiliteit resulteert in de duurzaamheidscontext echter gemakkelijk in een lange lijst criteria; omdat er veel onenigheid is over precieze definities, omdat er heel veel zorg- of aandachtspunten zijn die enigszins relevant zijn en omdat er sterk tegengestelde belangen zijn. Veel van deze criteria krijgen dan uiteraard een gering gewicht. Al met al leidt dit niet tot een verhelderende evaluatie waar men vervolgens dan ook weinig mee kan. En dat is weer niet in lijn met wat de MCA beoogt.

Door de combinatie met KBA en het daarmee veel scherper naar voren brengen van een analyse van causaliteit gekoppeld aan een versterkte focus op de allerbelangrijkste criteria kan de MCA haar kracht hervinden.

>>>Lees meer in 4.4.4b en in 4.4

8. De afwegingen en afruilen die MCA analisten en betrokkenen moeten maken tussen hoge orde criteria zijn in veel gevallen eenvoudigweg te moeilijk om te maken; daarom vermijdt de MCKBA het forceren van deze afwegingen. Interpretatie van de uitkomsten op die hoge orde criteria met behulp van ratio-analyse biedt enig soelaas – in de richting van verdere oordeelsvorming.

Om tot een uiteindelijke rangorde van alternatieven te komen geeft de MCA gewichten aan criteria. Het geven van gewichten aan hogere orde criteria – denk bijvoorbeeld aan milieu-impacts versus arbeidsmarktimpacts is echter vaak niet zo eenvoudig voor de betrokkenen. Gewichten toekennen aan lagere orde criteria – zeg het wegen van de schadelijkheid van verschillende stoffen – geeft veel minder problemen. In dit proefschrift wordt betoogd dat het gewichten geven aan die hogere orde criteria in de duurzaamheidscontext weinig realistische waarde heeft. Beter is het te werken met de tussenuitkomsten, de impacts per hoofdcriterium, en meer te zoeken naar objectievere interpretatie van deze uitkomsten m.b.v. ratio-analyses (bijv.: hoeveel natuurwinst per Euro). Ook ratio-analyses in vergelijking met andere projecten kunnen bijzonder krachtig zijn – wat dan overigens wel weer gestandaardiseerde meetmethoden vraagt om vergelijking mogelijk te maken.

>>>Lees meer in 5.4.8

>>>Lees voor concrete voorbeelden verder in 7.2.3.3 en 7.2.3.4

9. Voor een duurzame ontwikkeling is evaluatie op wereldschaal belangrijker dan de evaluatie over de lange termijn van de ‘toekomstige generaties’.

Ondanks de populariteit van de ‘toekomstige generaties’ bij definities van duurzaamheid is het wereldwijde of mondiale schaalniveau voor duurzaamheids evaluaties veel belangrijker. Immers, de gevolgen voor toekomstige generaties zijn altijd onzeker; gevolgen op wereldschaal kunnen met veel meer zekerheid worden blootgelegd.

Een deel van de ingewikkelde relaties en terugkoppelingen in de duurzaamheidscontext die vaak als belemmering worden gezien voor een heldere evaluatie, zijn niet echt ingewikkeld; ze liggen enkel op verschillende schaalniveaus. Daarom is het helder onderscheiden van verschillende schaalniveaus een belangrijk element van evaluatie in een duurzaamheidscontext.

De MCKBA die in dit proefschrift is ontwikkeld stelt bij evaluatie in een duurzaamheidscontext een mondiaal schaalniveau verplicht. Dit hoge schaalniveau garandeert dat geen belangrijke terugkoppelingen en weglekeffecten over het hoofd worden gezien. Om praktisch te blijven wordt dat gecombineerd met een analyse niveau dat dicht bij het beslissingsniveau ligt. Dit kan bijvoorbeeld een land zijn of een regio of een bedrijf.

>>>Lees meer in 2.3.2.4d en 5.4.2

10. Gevoeligheidsanalyses hebben een beperkte kracht. Om werkelijk het toegenomen begrip van het probleem naar betrokkenen te kunnen communiceren dient in de laatste fase van de evaluatie de vraag naar mogelijke nieuwe alternatieven veel centraler te worden gesteld.

Aan het eind van een evaluatie wordt vaak een gevoeligheidsanalyse verricht. In de MCKBA is de kerntaak van de analist aan het eind van de evaluatie het delen van zijn of haar toegenomen begrip. Een gevoeligheidsanalyse maakt hier een onderdeel van uit. Dit kan een nuttig instrument zijn inzake het ‘verantwoording afleggen’ naar verschillende betrokkenen. Gevoeligheidsanalyses hebben echter ingebouwde zwakte omdat het nu eenmaal lastig is om de hoofdboodschap van een evaluatie te ontcrachten. Een belangrijk middel om de uitkomsten van gevoeligheidsanalyses nadrukkelijker naar voren te halen is om – aan het eind van de evaluatie, op basis van het toegenomen begrip van de probleem situatie door de evaluatoren – expliciet de vraag naar nieuwe alternatieven aan de orde te stellen. Moeten nieuwe projectalternatieven worden overwogen? Moeten bestaande alternatieven worden gecombineerd?

Tenslotte weet de analist gewoonlijk als geen ander wat de beperkingen van de gehanteerde evaluatie techniek zijn. De analist heeft dan ook een belangrijke verantwoordelijkheid om expliciet te werken aan een correct begrip van zowel zeggingskracht als beperkingen van de methode. Naast de techniek is dit uiteindelijk ook in grote mate een kwestie van stijl.

>>>Lees meer in 5.4.9 en meer algemeen in 5.4.2

11. Het is gemakkelijk dingen ingewikkeld te maken.

Voor de analist die actief is in de duurzaamheidscontext liggen de kansen op complexe resultaten voor het oprapen. De kunst is de analyse eenvoudig te houden: weinig criteria, weinig schaalniveaus, houvast aan voor iedereen begrijpelijke concepten. Maar misschien is niet iedereen daar blij mee...

Tenslotte

Tenslotte, kunnen we terugkomen op het begin. Dit proefschrift gaat over projectevaluatie, duurzaamheid en verantwoording afleggen. Misschien enigszins paradoxaal, maar de complexiteit van die combinatie van drie zaken heeft geleid tot vereenvoudiging. De nieuwe methode MCKBA ziet in essentie af van zaken die evaluatie met ofwel KBA ofwel MCA ingewikkeld maken; en ze is gebaseerd op de – zeer Nederlandse en eenvoudige - consensus onder betrokkenen. Het veel explicieter testen dan tot dusverre is gebeurd in hoeverre consensus bestaat of kan worden bereikt binnen het soort evaluaties waarvoor de MCKBA is ontworpen is een buitengewoon interessante en belangrijke onderzoeksrichting voor de toekomst. De praktijk zal dan ook moeten leren in hoeverre de voorgestelde methode op grotere schaal nuttig kan zijn.

APPENDIX: ELABORATE TABLE OF CONTENTS

PREFACE
SUMMARY

PART ONE - INTRODUCTION

SUSTAINABILITY AND PROJECT EVALUATION

CHAPTER 1: CBA AND ACCOUNTABLE EVALUATION IN A SUSTAINABILITY CONTEXT

- 1.1 *Introduction*
- 1.2 *Central research question*
- 1.3 *The structure of this thesis*
- 1.4 *A note on the style of the thesis*

CHAPTER 2: THE CHALLENGES FOR PROJECT EVALUATION TOOLS IN AN ACCOUNTABLE SUSTAINABILITY CONTEXT

- 2.1 *Introduction*
- 2.2 *The judgement perspective in evaluation*
- 2.3 *Type of projects*
 - 2.3.1 General characteristics
 - 2.3.2 Evaluating on sustainability
 - 2.3.2.1 The concept of sustainability
 - 2.3.2.2 A long-term view (f1)
 - 2.3.2.3 A global (and local) perspective (f2)
 - 2.3.2.4 Economic development, environmental degradation and extreme poverty (f3)
 - 2.3.2.4a Three elements: Triple E
 - 2.3.2.4b Triple E relations: separate concerns or closely related concerns?
 - 2.3.2.4c Triple E – partly elements in their own right
 - 2.3.2.4d The three relations between the Triple E elements
 - 2.3.2.4e Triple E relations: facts and values
 - 2.3.2.4f Conclusions about Triple E relations
 - 2.3.2.5 Stakeholder participation (f4)
 - 2.3.2.5a Two reasons for stakeholder participation
 - 2.3.2.5b Various groups of stakeholders
 - 2.3.2.5c Minimum requirements for stakeholder participation
 - 2.3.2.5d Conclusion
- 2.4 *Synthesis of the essential characteristics*
 - 2.4.1 Four key items
 - 2.4.2 More on limitations, judgement and uncertainty
 - 2.4.2.1 Limitations and judgement
 - 2.4.2.2 Uncertainty
- 2.5 *Short summary of evaluation demands in a sustainability context*

PART TWO – THEORY

CBA, MCA AND THEIR INTEGRATION IN A SUSTAINABILITY CONTEXT

CHAPTER 3: CBA IN A SUSTAINABILITY CONTEXT

- 3.1 *Introduction*
 - 3.1.1 The discussion of CBA
 - 3.1.2 General background of CBA in economic science
 - 3.1.3 Two-steps
- 3.2 *The stages of CBA*
 - 3.2.1 Introduction
 - 3.2.2 Stage one: Define the project (or policy) and impact population
 - 3.2.2.1 General description
 - 3.2.2.1a Specification of project alternatives
 - 3.2.2.1b Determining the scale of the analysis: The impact population
 - 3.2.2.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.2.2a CBA literature silent on incorporation of newly emerging alternatives
 - 3.2.2.2b Measurement at a global scale is natural to CBA: Other scales can be included.
 - 3.2.3 Stage two: Identify project impacts

- 3.2.3.1 General description
 - 3.2.3.1a How to identify impacts in a CBA
 - 3.2.3.1b Identifying impacts: Opportunity costs
 - 3.2.3.1c Identifying impacts: Transfer payments
- 3.2.3.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.3.2a Problems with identifying costs and benefits
 - 3.2.3.2b Fundamental problems of identifying costs and benefits: ‘Social’ aggregation
 - 3.2.3.2c Fundamental problems of identifying costs and benefits: Relating ‘broad’ utility and narrow ‘economics’
- 3.2.4 Stage three: Quantify relevant impacts physically
 - 3.2.4.1 General description
 - 3.2.4.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.4.2a Empirical facts about (Triple E) impacts, their (un)certainly and relations become clear
- 3.2.5 Stage four: Value relevant effects monetarily
 - 3.2.5.1 General description
 - 3.2.5.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.5.2a Market values are the basis of CBA valuations of impacts
 - 3.2.5.2b Valuing environmental impacts often requires estimation of artificial market values
- 3.2.6 Stage five: Discount costs and benefits
 - 3.2.6.1 General description
 - 3.2.6.1a Discounting is natural for market related impacts
 - 3.2.6.1b Practical approach to choosing a discount rate: Government prescription
 - 3.2.6.1c Arguments for lower discount rates in the sustainability context
 - 3.2.6.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.6.2a The basis of discounting – alternative use – is not always meaningful
 - 3.2.6.2b The discounting debate is really about limitations to non-market valuation in CBA
- 3.2.7 Stage six: Apply the Net Present Value test
 - 3.2.7.1 General description
 - 3.2.7.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.7.2a Calculation of NPVs without difficulty, interpretation requires caution
 - 3.2.7.2b Distributional weights are sometimes considered
- 3.2.8 Stage seven: Perform sensitivity analysis
 - 3.2.8.1 General description
 - 3.2.8.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.2.8.2a Sensitivity analysis may be very rewarding in a conflict setting
 - 3.2.8.2b Uncertainty is not always best treated as footnotes to the main results: Uncertainty may turn out to be a central result
- 3.2.9 At this stage what is clear about CBA’s performance and what requires further examination?
- 3.3 *CBA and the incorporation of broad welfare notions*
 - 3.3.1 Introduction
 - 3.3.2 Broad welfare notions
 - 3.3.2.1 General description, subjective (and mainstream) welfare economists
 - 3.3.2.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.3.2.2a Welfare economics, despite its merit of realism, cannot be taken as a ‘basis’ for a practical- oriented CBA
 - 3.3.3 Mainstream: Willingness to Pay (WTP) as a shortcut to welfare measurement
 - 3.3.3.1 General description
 - 3.3.3.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.3.3.2a Monetary valuation of Triple E elements: Hesitation and context sensitivity
 - 3.3.3.2b The relevance of the dual-preference structure of consumer vs. citizen
- 3.4 *CBA and valuing gains for some, losses to others*
 - 3.4.1 Introduction
 - 3.4.2 The difficulty of distributional weights
 - 3.4.2.1 General description
 - 3.4.2.1a The rather limited success of distributional weights
 - 3.4.2.1b Ethical consensus
 - 3.4.2.1c Description of distributional impacts
 - 3.4.2.2 Assessment of pros and cons in the context of accounting for sustainability
 - 3.4.2.2a Judgement consensus among stakeholders is the final basis of CBA
 - 3.4.2.2b Careful description of distributional impacts the best option, with minimum demand of standardised attention to the extremely poor
 - 3.4.2.2c Everyday aggregate market values reflect parts of welfare changes connected with efficiency and individual use; they do not reflect fundamental values like distributional fairness

3.4.2.2d	Supply curves show info-richness of market values
3.4.2.2e	Aggregate market values are important and consistent everyday economic realities without broad welfare connotations
3.4	Conclusions on handling distributional impacts
3.5	<i>CBA and valuing environmental impacts</i>
3.5.1	Introduction
3.5.2	The concept of environmental value
3.5.2.1	General description
3.5.2.1a	The concept of Total Economic Value
3.5.2.2	Assessment of pros and cons in the context of accounting for sustainability
3.5.2.2a	Valuation is about anthropocentric value
3.5.2.2b	Non-direct use values are central, but they complicate comparability among Triple E aspects
3.5.3	Valuation techniques: Classification
3.5.4	Revealed preference techniques
3.5.4.1	Dose response and human capital
3.5.4.1a	General description
3.5.4.1b	Assessment of pros and cons in the context of accounting for sustainability
3.5.4.2	Effect on production or opportunity cost
3.5.4.2a	General description
3.5.4.2b	Assessment of pros and cons in the context of accounting for sustainability
3.5.4.3	Mitigatory and preventive expenditure and averting behaviour approaches
3.5.4.3a	General description
3.5.4.3b	Assessment of pros and cons in the context of accounting for sustainability
3.5.4.4	Replacement cost
3.5.4.4a	General description
3.5.4.4b	Assessment of pros and cons in the context of accounting for sustainability
3.5.4.5	Travel cost method and hedonic pricing
3.5.4.5a	General description
3.5.4.5b	Assessment of pros and cons in the context of accounting for sustainability
3.5.4.6	Conclusions on revealed preference
3.5.5	Expressed (or stated) preference: Contingent Valuation Method (CVM)
3.5.5.1	General description
3.5.5.2	Assessment of pros and cons in the context of accounting for sustainability
3.5.5.2a	Despite high hopes, CVM does not seem to provide reliable non-use values
3.5.6	Would MCA not provide a better starting point than CVM?
3.6	<i>Conclusions: CBA's performance in a sustainability context</i>
<i>Appendix to chapter 3 Illustration of pitfalls of environmental valuation - The Costanza attempt</i>	
	Introduction
	The method of the Costanza attempt
	The results of the Costanza attempt
	CBA basics: Comparability and aggregation
	Stakeholder understanding in the (policy) debate: Monetary information and other information
CHAPTER 4: MCA IN A SUSTAINABILITY CONTEXT	
4.1	<i>Introduction</i>
4.1.1	General
4.1.2	Different approaches to MCA
4.1.3	Presenting the essentials of MCA
4.1.4	Two-step approach
4.2	<i>The stages of an MCA</i>
4.2.1	Stage one: Identify objects and function(s) of the evaluation
4.2.1.1	General description
4.2.1.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.1.2a	MCA does not stress the need for a 'nul-alternative' or zero-option.
4.2.1.2b	MCA reflects on the functional context and tailors the evaluation to the project
4.2.2	Stage two: Identify stakeholders
4.2.2.1	General description
4.2.2.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.2.2a	Involvement of stakeholders makes the analyst's political and communication skills a critical factor for successful evaluation
4.2.2.2b	Involvement of all relevant stakeholders may be nearly impossible
4.2.3	Stage three: Elicit and organise criteria
4.2.3.1	General description
4.2.3.2	Assessment of pros and cons in the context of accounting for sustainability

4.2.3.2a	Triple E impacts can easily and equally be incorporated into MCA
4.2.3.2b	Values are structured in the process of MCA
4.2.3.2c	MCA depends on measurement options for clarifying and structuring judgement criteria
4.2.3.2d	In a complex decision-making context the issues of interdependence of preferences and double counting by unclear causality are important
4.2.4	Stage four: Assess the relative importance of criteria
4.2.4.1	General description
4.2.4.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.4.2a	Assigning weights may be difficult, especially with many stakeholders
4.2.5	Stage five: Estimate scores of alternatives on low level criteria
4.2.5.1	General description
4.2.5.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.5.2a	Achieving comparable measurement implies information loss and affects judgement
4.2.5.2b	Interdependency between knowing the relevant range of impacts and giving weights is not ideal for accountability
4.2.5.2c	The preferred personal interactivity of stakeholders with MCA analyst is a complicating factor
4.2.5.2d	The 'open' treatment of impacts spread over time within MCA contrasts sharply with CBA's discounting practice
4.2.6	Stage six: Aggregate scores with relative weights
4.2.6.1	General description
4.2.6.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.6.2a	Difficulty of establishing trade-off may lead to second-best MCA with only partial judgement
4.2.7	Stage 7: Perform sensitivity analyses
4.2.7.1	General description
4.2.7.2	Assessment of pros and cons in the context of accounting for sustainability
4.2.7.2a	Additional analysis may give some extra judgement possibility in case of trade-off difficulties
4.2.7.2b	A first MCA can be followed by another and yet another, including newly emergent alternatives
4.2.8	At this stage: what is clear about MCA's performance and what requires more attention?
	- Difficulty of establishing trade-offs may lead to a (second-best) MCA with only partial judgement and disaggregated results
4.3	<i>MCA and organising stakeholder involvement</i>
4.3.1	Introduction
4.3.2	Organising many distant stakeholders
4.3.3	Reconciliation of diametrically opposing interests
4.3.4	Conclusions on organising stakeholder involvement
4.3.4a	<i>Practical impossibilities limit representative stakeholder involvement</i>
4.3.4b	<i>Opposing interests of stakeholders discourages the search for a viable common perspective; broadness of perspectives can be achieved and can be useful</i>
4.4	<i>MCA and interdependence of criteria and double counting</i>
4.4.1	Large value trees
4.4.2	Interdependency of criteria and double counting
4.4.3	Trade-off difficulties
4.4.4	Conclusions on interdependence of criteria and double counting
4.4.4a	<i>Interdependency and double counting is hard to avoid in a complex decision-making setting</i>
4.4.4b	<i>Large value trees make MCA less effective – Focus on the most important criteria may help</i>
4.4.4c	<i>Partial judgement will often be the maximum result of an MCA</i>
4.5	<i>MCA and its judgement orientation</i>
4.5.1	Introduction
4.5.2	Descriptive and outranking approaches
4.5.3	Decision aid versus decision-making approaches in MCA
4.5.4	Assessment of the judgement orientation
4.5.4a	<i>MCDM and MCDA difference is not a matter of technique – Decision aid is the correct style of using MCA in the sustainability context</i>
4.6	<i>Conclusions: MCA's performance in the sustainability context</i>
CHAPTER 5:	MCCBA: INTEGRATING CBA AND MCA IN THE SUSTAINABILITY CONTEXT
5.1	<i>Introduction</i>
5.2	<i>Common ground for integrating MCA and CBA in MCCBA</i>
5.2.1	Full judgement is often impossible in a sustainability context
5.2.2	Decision aid is the realistic aim
5.2.3	The common ground
5.3	<i>Major building blocks for integrating MCA and CBA in MCCBA</i>
5.3.1	Introduction
5.3.2	Stakeholder involvement useful for broadness of views and for checking on consensus

- 5.3.3 Standardisation of global long-term Triple E criteria structure
- 5.3.4 Judgement and measurement are closely related
- 5.4 *Stages of the MCCBA approach*
 - 5.4.1 Introduction to the eight stages of MCCBA
 - 5.4.2 Stage one: Identify function, project alternatives and scale of the evaluation
 - 5.4.3 Stage two: Involve a broad group of stakeholders
 - 5.4.4 Stage three: Organise judgement criteria on Triple E impacts
 - 5.4.5 Stage four: Quantify impacts physically
 - 5.4.6 Stage five: Aggregate monetary scores consensus based
 - 5.4.7 Stage six: Aggregate non-monetary scores consensus based
 - 5.4.8 Stage seven: Interpret trade-offs
 - 5.4.9 Stage eight: Perform sensitivity analysis and reconsider project alternatives

PART THREE – USING MCCBA

THREE CASE STUDIES

CHAPTER 6: INTRODUCTION TO PART THREE

- Three chapters about applicability
- Essential aspects: Improving upon either CBA or MCA
- The character of the following chapters

CHAPTER 7: EVALUATION OF THE DUTCH ECOLOGISCHE HOOFDSTRUCTUUR (EHS): THE VALUE OF NON-MONETARY MEASUREMENT

7.1 *Introduction*

7.2 *Original results*

- 7.2.1 Brief summary of results
- 7.2.2 Results and method detailed
 - 7.2.2.1 Nature values
 - 7.2.2.2 Agricultural production
 - 7.2.2.3 Impact on recreation
 - 7.2.2.4 Costs of purchase, management and restoration
- 7.2.3 Overall evaluation of the results
 - 7.2.3.1 Introduction
 - 7.2.3.2 Government finance impacts as a threshold criterion
 - 7.2.3.3 The two major criteria in perspective
 - 7.2.3.4 Further interpretation per component of EHS
 - Efficiency with regard to costs to society
 - Efficiency with regard to government finance
 - Sensitivity analysis

7.3 *The Ecological Main Structure evaluation in light of the MCCBA approach*

- 7.3.1 Introduction
- 7.3.2 Stage one: Identify function, project alternatives and scale of the evaluation
 - 7.3.2.1 The function of the research
 - 7.3.2.2 The project alternatives
 - 7.3.2.3 The scale of the evaluation
- 7.3.3 Stage two: Involve a broad group of stakeholders
- 7.3.4 Stage three: Organise judgement criteria on Triple E impacts
- 7.3.5 Stage four: Quantify impacts physically
- 7.3.6 Stage five: Aggregate monetary scores consensus based
- 7.3.7 Stage six: Aggregate non-monetary scores (consensus based)
- 7.3.8 Stage seven: Interpret trade-offs
- 7.3.9 Stage eight: Perform sensitivity analysis and reconsider project alternatives
- 7.3.9 Epilogue
 - CBA results
 - Quantifying nature values and cost-effectiveness
 - Government budget

CHAPTER 8: EVALUATION OF THE GERMAN EMSSPERRWERK: THE VALUE OF MORE SCALE LEVELS

8.1 *Introduction*

8.2 *Original results*

- 8.2.1 Brief summary of results
- 8.2.2 Results and method in more detail
 - 8.2.2.1 Alternatives
 - 8.2.2.2 Nature impacts

8.2.2.3	Economic impacts
	Introduction
	Costs of dike improvement
	Costs of a flood barrier
	Costs of the Emssperrwerk
	Costs of renaturalisation
	Benefits of the different alternatives
	The benefit of the increased safety level
	The benefit of lower costs for dredging
	The benefit of secured growth possibilities for the Meyer shipyard
8.2.3	Overall evaluation of the results
8.2.3.1	Results at the German level
8.2.3.2	Results at the level of Niedersachsen
8.2.3.3	Results at the global level
8.2.3.4	Results at all levels
8.3	<i>The Emssperrwerk in the light of the MCCBA approach</i>
8.3.1	Introduction
8.3.2	Stage one: Identify function, project alternatives and scale of the evaluation
8.3.2.1	Function of the research
8.3.2.2	Project alternatives
8.3.2.3	Scale of the evaluation
8.3.3	Stage two: Involve a broad group of stakeholders
8.3.4	Stage three: Organise judgement criteria on Triple E impacts
8.3.5	Stage four: Quantify impacts physically
8.3.6	Stage five: Aggregate monetary scores consensus based
8.3.7	Stage six: Aggregate non-monetary scores consensus based
8.3.8	Stage seven: Interpret trade-offs
8.3.9	Stage eight: Perform sensitivity analysis and reconsider project alternatives
8.3.10	Epilogue on uncertainty
	The costs of the Emssperrwerk
	Meyer investment and the business cycle
	<i>Appendix 1 to chapter 8: Sensitivity analysis for four crucial parameters</i>
	Share Dutch employees
	An extra growth period of 20 years
	Lower relative growth of Meyer
	Higher inconvenience compensation
	<i>Appendix 2 to chapter 8: The results of the total nature score approximation for Niedersachsen and Germany</i>
CHAPTER 9:	DESIGN AND EVALUATION OF SUSTAINABLE CORPORATE PERFORMANCE (SCP) POLICY: THE VALUE OF CAUSAL ANALYSIS
9.1	Introduction
9.2	<i>Designing SCP policies following the MCCBA approach</i>
9.2.1	Introduction
9.2.2	SCP in general
9.2.2.1	Definitions
9.2.2.2	SCP business practices
9.2.2.2a	Practices related to all stakeholders
9.2.2.2b	Practices relating to employees and the community
9.2.2.2c	Practices relating to other stakeholder groups
9.2.2.2d	Conclusion
9.2.3	Existing SCP evaluation tools
9.2.3.1	Introduction
9.2.3.2	The Dow Jones Sustainability Index (DJSI)
9.2.3.3	Business Ethics 100
9.2.3.4	Limitations of existing evaluation tools
9.2.4	MCCBA and the design of SCP policies
9.2.4.1	Four MCCBA elements fruitful for SCP design
Box 9-1:	Individual or market/sector action?
9.2.4.2	More important elements of SCP policy design
	A 'sustainability audit' at the sectoral level
	Limitation of the number of issues
9.3	<i>Exploring a concrete example: Evaluating EurepGAP</i>
9.3.1	Introduction

- 9.3.2 Stage one: Identify function, project alternatives and scale of the evaluation
- 9.3.3 Stage two: Involve a broad group of stakeholders
- 9.3.4 Stage three: Organise judgement criteria on Triple E impacts
- 9.3.5 Stage four: Quantify impacts physically
- 9.3.6 Stage five: Aggregate monetary scores consensus based
- 9.3.7 Stage six: Aggregate non-monetary scores consensus based
- 9.3.8 Stage seven/Stage eight: Interpret trade-offs/perform sensitivity analysis and reconsider project alternatives
- 9.4 *Conclusions*
 - Ahold's challenges
 - Gains from the MCCBA approach
- Appendix to chapter 9: Some examples of sector-level SCP projects*

PART FOUR – CONCLUSIONS AND FINAL REFLECTIONS

THE MERITS OF THE MCCBA APPROACH

CHAPTER 10: CONCLUSIONS AND FINAL REFLECTIONS

- 10.1 Introduction*
- 10.2 Additional conclusions (additional to conclusions of chapter 5)*
- 10.3 Final reflections*

LITERATURE CITED

SAMENVATTING

APPENDIX: ELABORATE TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PUBLICATION SERIES OF THE STICHTING RUIMTELIJKE ECONOMIE GRONINGEN

LIST OF TABLES

Table ES-1	The eight stages of the MCCBA approach.....	viii
Table 3.1	Aspects of accountable-sustainability context explicitly addressed by different CBA authors of textbooks (although terminology may differ).....	31
Table 3.2	First conclusions about CBA's performance in the sustainability context.....	50
Table 3.3	Relevance of types of valuation and familiarity with giving WTP estimates for separate Triple E elements	58
Table 3.4	Total economic value of an environmental asset and the relevance of its several components in a sustainability context	72
Table 3.5	Total economic value of a car and the reflection of its several components in an aggregate demand curve	72
Table 3-6	Conclusions on CBA's performance in the sustainability context (per stage).....	83
Table 4.1	First conclusions about MCA's performance in the sustainability context.....	106
Table 4.2	The differences between decision-making and decision aid approaches	118
Table 4.3	Conclusions on MCA's performance in the sustainability context (per stage).....	121
Table 5.1	The eight stages of the MCCBA approach – compared to the stages of CBA and MCA	129
Table 6-1	Distinctions among the three chapters of part three	142
Table 7-1	Distribution of target species across species groups	145
Table 7-2	Average nature score per hectare for EHS and non-EHS sites, based on target species per se	146
Table 7-3	Nature score with optimality for EHS and non-EHS realised sites (based on target species with optimality).....	147
Table 7-4	Nature score with optimality and threat for EHS and non-EHS sites (based on target species with optimality under degree of threat.....	149
Table 7-5	Government finance impacts by EHS	155
Table 7-6	Results of Multi-Criteria Analysis at national level	156
Table 7-7	The overall impacts of EHS, on the three MCCBA criteria and the relative contribution of different EHS components.....	156
Table 7-8	Cost efficiency of EHS components.....	157
Table 7-9	Efficiency of EHS components with regard to government finance and nature values.....	158
Table 8-1	Calculation of nature value in the nul-alternative according to nature score method	172
Table 8-2	The nature scores of the different alternatives.....	173
Table 8-3	Costs of dike improvement in Euro mln	173
Table 8-4	Costs of a flood barrier in Euro mln.....	174
Table 8-5	Overview of costs of the Emssperrwerk to Germany in Euro mln (net present value).....	175
Table 8-6	Benefits of the Emssperrwerk alternatives (2abcd) and the renaturalisation alternatives (3abc) in mln Euro (German level; discounted totals).....	178
Table 8-7	Summary of results at the German level ('+' means net benefits, '-' means net costs).....	179
Table 8-8	Summary of results at the level of Niedersachsen ('+' means net benefits, '-' means net costs).....	180
Table 8-9	Summary of results at the global level	181
Table 8-10	Comparison of two alternatives (Emssperrwerk and renaturalisation) at	

three spatial scale levels	182
Table App.-8.1 Sensitivity analysis of the share of Dutch employees (level of Germany)	191
Table App.-8.2 Sensitivity analysis of 20 year extra growth period (level of Germany)	191
Table App.-8.3 Sensitivity analysis of lower relative growth of the Meyer ship yard (level of Germany)	192
Table App.-8.4 Sensitivity analysis of higher inconvenience compensation (level of Germany)	192
Table App.-8.5 Approximation of nature score total for Niedersachsen and Germany ...	193
Table 9-1 Division of SCP activities among stakeholder groups	197
Table 9-2 Examples of CSR evaluation methods according to user aim	202
Table 9-3 Overview of examples of measurement aspects of different sustainability indices	203
Table 9-4 Examples of questions of the DJSI	204
Table 9-5 Top 5 and lowest 5 of the Business Ethics 100 Best Corporate Citizens - 2002.	205
Table 9-6 Summary of views on the impact of EurepGAP	217
Table 9-7 Explorative results of the impacts of EurepGAP	219

LIST OF FIGURES

Figure 1.1: Graphical illustration of how the chapters in the thesis relate and how they are part of the “to-and-fro” between theory and practice.	7
Figure 2.1: Three perspectives in the evaluation process	10
Figure 2.2: Four basic options of connections between the Triple E elements	16
Figure 2.3: Illustration of how a higher spatial level and a longer term perspective may give a clearer view on the interrelatedness of Triple E elements.	20
Figure 2.4: Key items of accountable evaluation in a sustainability context	24
Figure 2.5: Illustration of how, from an initial state of unawareness, three different evaluations may result in various divisions of knowledge areas	27
Figure 3.1: The impact of several discount rates on the long-term effects	47
Figure 3.2: The relation of the different sections that follow	52
Figure 9-1: The preliminary criteria structure of the EurepGAP evaluatio	217

PUBLICATION SERIES OF THE STICHTING RUIMTELIJKE ECONOMIE GRONINGEN

1. J. van Dijk en R.J. Stoffelsma (red.), *Groningen, Friesland en Drenthe op eigen kracht naar Europa 2000?* 1990. ISBN 90-73709-01-6.
2. R.J. Stoffelsma en D. Strijker (red.), *De noordelijke akkerbouw: sector en regio op zoek naar een toekomst.* 1991. ISBN 90-73709-03-2.
3. D. Strijker (red.), *Intermodaal: op welke schaal?* 1994. ISBN 90-73709-05-9.
4. D. Strijker, J. van Dijk en M.J. Jongsma, *Vooruitzichten voor de arbeidsmarkt van de noordelijke land- en tuinbouw.* 1994. ISBN 90-73709-06-7.
5. F.J. Sijsma en D. Strijker, *Investerings met ruimtelijke werking in het landelijk gebied.* 1994. ISBN 90-73709-07-5, 2e druk 1995.
6. F.J. Sijsma, *Arbeidsmarktanalyse Provincie Groningen.* 1995. ISBN 90-73709-08-3.
7. G. Eding, T.M. Stelder, E.R. Vos en J. Oosterhaven, *Bi-regionale interactie: nieuwe input-output tabellen voor Groningen, Friesland, Drenthe en Overijssel voor 1990.* 1995. ISBN 90-73709-09-1.
8. F.J. Sijsma en D. Strijker, *Effect-analyse Ecologische Hoofdstructuur, Deel I: Hoofdrapport.* 1995. ISBN 90-73709-10-5.
9. F.J. Sijsma en D. Strijker, *Effect-analyse Ecologische Hoofdstructuur, Deel II: Natuurwaarde.* 1995. ISBN 90-73709-11-3.
10. G.J. Eding, T.M. Stelder en J. Oosterhaven, *Structuur en Interactie: Een interregionale structuurschets van Groningen, Friesland en Drenthe voor 1990.* 1995. ISBN 90-73709-12-1.
11. D. Strijker en F.J. Sijsma (red.), *Metten van natuur.* 1996. ISBN 90-73709-13-X.
12. F.J. Sijsma, T.M. Stelder, J.P. Elhorst, J. Oosterhaven en D. Strijker, *Ruimte te over, Ruimte tekort.* 1996. ISBN 90-73709-14-8.
13. J.P. Elhorst (red.), *Ruimte te over, Ruimte tekort: Woord en Weerwoord.* 1996, ISBN 90-73709-15-6.
14. J. Oosterhaven, F.J. Sijsma en J.B. Polak, *Schiphol-Flevo: Een verkenning van de werkgelegenheidseffecten.* 1998. ISBN 90-73709-16-4.
15. J.B. Polak (red.), *Differentiatie in de vervoerseconomie.* 1998. ISBN 90-73709-17-2.
16. J.P. Elhorst en J. Oosterhaven (red.), *Transport en Welvaart: Vervoer in Vogelvlucht.* 1998. ISBN 90-73709-18-0.
17. L. Broersma, M. Woolthuis en D. Strijker, *Effect-Analyse ISP 5 Agro-Business.* 1998. ISBN 90-73709-20-7.
18. J.P. Elhorst en D. Strijker (red.), *Het Economisch Belang van het Vervoer.* 1999. ISBN 90-73709-21-0.
19. G.J. Eding, J. Oosterhaven en T.M. Stelder, *Clusters en Linkages in Beeld. Een toepassing op de regio's Noord-Nederland, Groot-Amsterdam/NZKG en Groot-Rijnmond.* 1999. ISBN 90-73709-22-9.
20. Rijksuniversiteit Groningen en Centraal Bureau voor de Statistiek, *Regionale Samenhang in Nederland.* 1999. ISBN 90-73709-23-7.
21. D. Strijker, *Ruimtelijke Verschuivingen in de EU-Landbouw 1950-1992.* 1999. ISBN 90-73709-24-5.
22. J.P. Elhorst, T. Knaap, J. Oosterhaven, W.E. Romp, T.M. Stelder en E. Gerritsen, *Ruimtelijk Economische Effecten van Zes Zuiderzeelijn Varianten.* 2000. ISBN 90-73709-25-3.

23. J. Oosterhaven en D. Strijker (red.), *Effecten Magneetweefbaan Randstad-Noord-Nederland*. 2000. ISBN 90-73709-26-1.
24. D. Strijker, J.P. Elhorst, J. Oosterhaven en A.S. Zeilstra, *Kompas voor Kwaliteit*. 2001. ISBN 90-73709-27-X.
25. W.E. Romp, M.J.P.M. Thissen, J. Oosterhaven en J.P. Elhorst, *Indirecte Effecten Rondje Randstad: migratie en banen*. 2001. ISBN 90-73709-28-8.
26. D. Stelder, *Regions and Cities: Five Essays on Interregional and Spatial Agglomeration Modelling*. 2005. ISBN 90-73709-20-2.
27. F.J.Sijtsma, *Project evaluation, sustainability and accountability: Combining cost-benefit analysis and multi-criteria analysis*. 2006. ISBN 90-73709-24-5. ISBN (new) 978-90-73709-24-5

Overige publicaties Stichting Ruimtelijke Economie Groningen

J. Oosterhaven, *Het Noorden in Europees Perspectief*. 1991. (Oratie)

R.J. Stoffelsma, *Centrale Overheid en Regionale Welvaart*. 1992. (Dissertatie)

J.B. Polak (red.), *Differentiatie in de Vervoerseconomie*. 1998. (Liber Amicorum prof. Kuiler)

J.B. Polak, *Geef Vervoerseconomie de Ruimte*. 1998. (Afscheidsrede)

Publications of the Stichting REG can be ordered at:

Secretariaat Ruimtelijke Economie, RuG, Postbus 800, 9700 AV Groningen;

Phone +31 (0)50-3633740.

E-mail: H.C.W.E.van.der.Vleuten@rug.nl

Website: www.REGroningen.nl