Report 555000001/2005

SustainabilityA-Test inception report
progress to date and future tasks

W de Ridder*

* Corresponding author: Netherlands Environmental Assessment Agency (MNP associated with the RIVM), PO Box 303, 3720 AH Bilthoven, the Netherlands, e-mail: wouter.de.ridder@mnp.nl; telephone +31 30 274 4243; fax: +31 30 274 4433

This publication has been co-funded under the EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, contract GOCE-CT-2003-505328). Its content does not represent the official position of the European Commission and is entirely under the responsibility of the authors.

Netherlands Environmental Assessment Agency (MNP associated with the RIVM), PO Box 303, 3720 AH Bilthoven, the Netherlands, telephone +31 30 274 274 5, website: www.mnp.nl

Contributing authors: §3.2 Case study by Karlheinz Knickel (IfLS)

§3.4 Handbook by Marjan van Herwijnen (IVM)

§4.3 Physical assessment tools by Philipp Schepelmann (WI), Stefan Bringezu (WI), Matthias Nerger (WI), Stephan Moll (WI), Helmut Schütz (WI), Karl-Heinz Simon (CESR), Anne van der Veen (UT) and P.J. Stauvermann (UT)

§4.4 Monetary assessment tools by Anna Alberini (University of Maryland and FEEM), Onno Kuik (IVM , Vrije Universiteit Amsterdam), Harmen Verbruggen (IVM, Vrije Universiteit Amsterdam), Benjamin Goerlach (Ecologic), Anne van der Veen (University of Twente) and Julia Bartos (Czech Environment Center)

§4.5 Modelling tools by Hermann Lotze-Campen (PIK), Tom Kram (RIVM), Reyer Gerlach (IVM), Pim Martens (UM-ICIS), Pieter Valkering (UM-ICIS), Karl-Heinz Simon (UNIK-CESR), Alex Haxeltine (UEA-Tyndall), John Turnpenny (UEA-Tyndall), John Robinson (UBC-SDRI), Matthias Nerger (WI) and Gonçalo Lobo (EC-JRC)

§4.6 Scenario analysis tools by Karl-Heinz Simon (CESR), Alexa Matovelle (CESR), Joe Alcomo (CESR), Claudia Pahl-Wostl (ISF) and John Robinson (SDRI)

§4.7 Multi-criteria analysis tools by J. David Tàbara (IEST), Nadja Kasperczyk (IfLS), Karlheinz Knickel (IfLS), Gregor Meerganz (IEST), Daniela Russi (IEST) and Marjam van Herwijnen (IVM)

§4.8 Sustainability appraisal tools by Måns Nilsson (SEI), Kerstin Ehrhart (SEI), Dirk Günther (ISF), Anneke Klasing (Ecologic), Kata Wagner (Ecologic) and Gina Ziervogel (SEI-Oxford)

§4.9 Participatory tools by Matthijs Hisschemöller (IVM), Åsa Swartling (SEI), Marleen van de Kerkhof (IVM), Eefje Cuppen (IVM), Gonçalo Lobo (EC-JRC) and Ângela Guimarães Pereira (EC-JRC)

Please visit www.SustainabilityA-Test.net for all the available tool information.

Abstract

SustainabilityA-Test: progress to date and future tasks

Within the framework of the project, *SustainabilityA-Test,* an appraisal will be made of the methods and techniques (hereafter referred to collectively as *tools*) that may (can) be used in sustainability-related impact assessments in support of policy at various stages of development. The intrinsic qualities of the tools that are used in addressing aspects of sustainable development will be analysed, together with other relevant tool characteristics such as financial aspects and data and time requirements. The ultimate aim of the project *SustainabilityA-Test* is to improve the scientific basis of the application of tools in (sustainability) impact assessments.

This report provides an overview of the activities carried out during the first phase of the project, the most important of which were:
– The development of an overview and a preliminary evaluation of the tools that are to be evaluated at later stages during the course of the project;
– The development of an evaluation framework for the evaluation of tools; this evaluation framework also describes the aspects of sustainable development in more concrete terms;
– The design of a concrete working plan for the second phase of the project: the case study. In this phase of the project, the tools will be analysed on the basis of an actual case study which is set up around a European directive on biofuels and a regulation on energy crop premiums.

Keywords: impact assessment, sustainable development, sustainability assessment, policy analysis, tools, methods, techniques.

Rapport in het kort

SustainabilityA-Test: voortgang tot nu en toekomstige taken
In kader van het project *SustainabilityA-Test* wordt een beoordeling gemaakt van methoden en technieken (hierna: *tools*) die een rol (kunnen) spelen bij duurzaamheid gerelateerde effectrapportages ten behoeve van verschillende fases van beleid. Er wordt gekeken naar de kwaliteit van deze *tools* om verschillende kanten van duurzaamheid te onderzoeken. Daarnaast wordt ook een overzicht gemaakt van kosten, benodigde tijd en data en andere relevante eigenschappen van ieder *tool*. Het uiteindelijke doel van het project is om de wetenschappelijke onderbouwing te verbeteren van het gebruik van *tools* ten behoeve van effectrapportages van beleid op duurzame ontwikkeling.

Dit rapport geeft een beschrijving van de activiteiten van de eerste fase van het project. De belangrijkste zijn:

- Het creëren van een overzicht en voorlopige evaluatie van *tools* die verder zullen worden geëvalueerd tijdens het project;
- De ontwikkeling van een evaluatie kader ten behoeve van de evaluatie van de *tools,* waarin ook de aspecten van duurzaamheid verder uitgewerkt zijn;
- Het opstellen van een concreet werkplan voor de tweede fase van het project: de *case study*. In deze fase zullen de *tools* aan de hand van een concrete casus (een Europese richtlijn voor biobrandstoffen en een regeling voor premies op energiegewassen) worden toegepast.

Trefwoorden: effectrapportages, duurzame ontwikkeling, duurzaamheidanalyse, beleidsanalyses, tools, methoden, technieken

Project no. 505328

SustainabilityA-Test

Advanced Techniques for Evaluation of Sustainability Assessment Tools

Instrument: STREP
Thematic priority: [1.1.6.3] Global change and Ecosystems

D5: SustainabilityA-Test inception report
Progress to date and future tasks

Due date of deliverable: February 2005
Actual date of submission: 25 February 2005

Start of project: 1 March 2004 Duration: 30 months

Lead contractor for this deliverable:
RIVM-MNP
PO Box 303
NL-3720 AH Bilthoven
The Netherlands
Correspondance: Wouter de Ridder (wouter.de.ridder@mnp.nl)

Revision : version 6 (final)

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

Summary

Introduction

This report provides an overview of the first phase of the project *SustainabilityA-Test*. The ultimate aim of the project is to examine the theoretical and conceptual basis of the tools commonly used in (sustainability) assessments based on a literature review and a case study. The project will result in a synthesis in which these various tools and a number of assessment methods are compared within the framework of the requirements of sustainable development assessments. This synthesis will be made available in the form of a handbook.

There are three distinct phases to the *SustainabilityA-Test project*:

1. Inception
2. Case study
3. Integration and synthesis

Progress

By the end of January 2005, 9 months after the project had been initiated, the following progress had been made:

- creation of an evaluation framework for the evaluation of the tools;
- creation of a preliminary tool overview and evaluation papers, in which each tool group describes its tools as evaluated by means of the evaluation framework;
- setting-up of a concrete working plan for the project's case study;
- putting to paper the first ideas on the (electronic) handbook and a manner of having those interested access the information generated within *SustainabilityA-Test*.

Results

Evaluation framework

All tools within the framework of *SustainabilityA-Test* were evaluated for:

- their ability to support certain *policy processes*;
- their ability to address various *aspects of sustainable development*;
- their costs, time requirements, level of required expertise and further aspects (*operational aspects*).

The aspects of sustainable development are classified into topics that can be found within the three pillars of sustainable development (i.e. environmental protection, social development and economic development) and the so-called cross-cutting aspects of sustainable development (i.e. topics that cannot be attributed unambiguously to one of the three pillars, but which are important for sustainable development). An example of just such a topic of sustainable development is *global dimension*.

Knowledge of the tools' ability to address the aspects of sustainable development can also be used to analyse their ability to assess impacts of those policies that the European Commission has identified as having priority within its Sustainable Development Strategy.

With respect to the tool groups physical assessment tools and multi-criteria tools, the experiences with the evaluation framework and the relevant tools show that the evaluation framework works quite well. For the other tool groups, however, applying the evaluation framework seems less effective in providing the information required. The information referred to here is formed by the tool characteristics needed to actually build a toolbox capable of showing which tools can be part of which methods or broader instruments or approaches to measure and assess the three pillars of sustainable development.

Two main causes for the less effectiveness of the framework can be identified:

1. The evaluation framework does not contain the most relevant evaluation criteria: for example, tools of a more procedural character, such as s*ustainability impact assessment*, need additional criteria relative to those tools used in guiding assessment procedures, and/or
2. The evaluation framework is not being applied at the right level of the tools: for example, the evaluation of the tool cost/benefit analysis will not provide us with any information on the applicability of the various methods to monetise benefits.

Further work is therefore needed to overcome the difficulties outlined above:

1. critically analysing each tool group to determine if we are evaluating on the most efficient level;
2. critically analysing – and if needed, adjusting – the evaluation criteria for their ability to address the most relevant tool characteristics;
3. continuing with the evaluation of tools within the evaluation framework.

Tool overview

The tool overview paper has so far resulted in the identification of 42 tools. The figure below provides an overview of these tools.

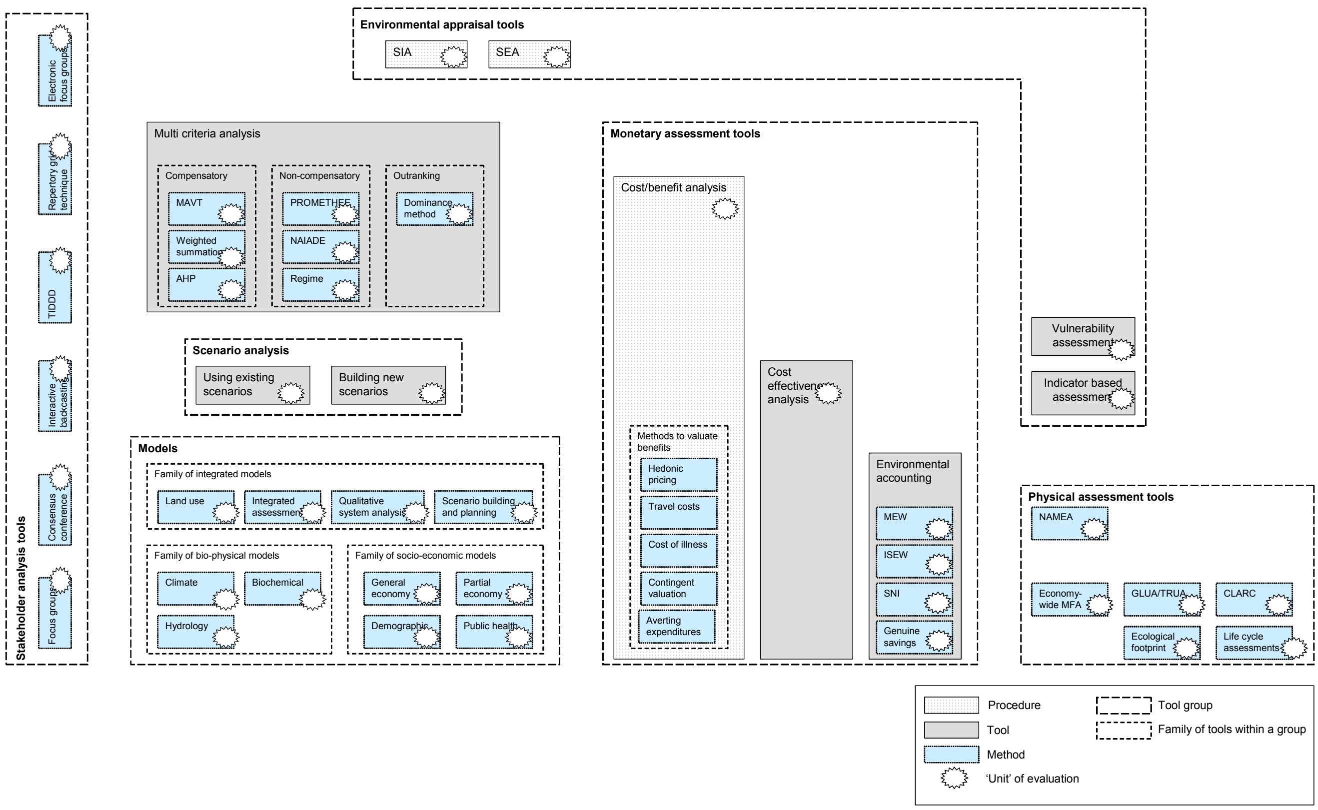


Figure 1: Overview of tools considered in SustainabilityA-Test

By means of an additional review of all tool papers (both within the project team and from external sources), we should do a further analysis to see if all tools commonly used in assessments have indeed been included.

Case study

The case study was launched during a meeting in Berlin with all of the project partners on March 2 and 3, 2005. During the first part of the case study, which will run until June 2005, the phase-1 preliminary tool evaluation will be extended by analysing how tools actually have been used at the European and national levels in the development of the Biofuel Directive and the Energy Crop Premium Regulation. In addition, the tool experts will describe how their tools could have been used (differently) within the context of these directives. The key question will be what extra areas could have been covered (or which areas that were already covered could have been covered better) by a tool. These areas refer to the twelve priority areas set by the European Commission's Sustainable Development Strategy.

Work package 1 of the project has clarified the necessity to distinguish between tools that can be used stand-alone and tools that are fully dependent on other tools. In addition, the need to make combinations in order to cover a broad spectrum of sustainability aspects has been confirmed. Both aspects will therefore be researched further during the case study, although at a later stage (part 2 of the case study). Policy makers could also be consulted with the aim of obtaining a better integration of the policy-making perspective into the final outcome of *SustainabilityA-Test*.

Handbook

The development of an electronic and/or web-based handbook, with some search capacity to find the various tools, is under consideration. An electronic handbook will facilitate the navigation through the vast amount of information generated during the project. Options for an electronic handbook are currently under investigation and were discussed with all project partners at the Berlin meeting.

1 Introduction

> 'The EU and Member States need to pursue the development of impact assessment tools in order to help them make well-informed decisions. These should take into consideration all costs and benefits, including short and long term, as well as global competitiveness' (High Level Group, 2004).

Assessment tools play an important role in decision-making processes. The collection of tools that can be used to carry out assessments is huge. Each tool has its own specific quality and contributes in a particular way to decision-making processes. Each tool can be used to address different issues, like costs and benefits, short and long-term effects, global competitiveness and many more key aspects in relation to sustainable development.

SustainabilityA-Test evaluates tools that can be used for sustainability assessments, amongst others to improve the scientific underpinning of these assessments. This report is the inception report of the project. It provides an overview of the first phase of the project *SustainabilityA-Test*.

1.1 SustainabilityA-Test

SustainabilityA-Test is a so-called 'specific targeted research or innovation project' (STREP) under the 6th framework research programme of the European Commission (priority 1.1.6.3 – Global change and ecosystems)[1].

In short, *SustainabilityA-Test* will take stock of and evaluate tools that can be used for carrying out (parts of) assessments. This task will provide us with an overview of tools and the issues that can be addressed with them. In addition, *SustainabilityA-Test* will analyse what assessment questions could be asked when one wishes to address the various aspects that are relevant when assessing in light of sustainable development. Lastly, *SustainabilityA-Test* will develop a handbook that brings together the outcome of both tasks. This handbook should help those 'on the verge of carrying out an assessment' with formulating assessment questions and finding the most suitable tools to answer these.

> To illustrate this in a simplified manner: when a proposal for an extension of road infrastructure needs to be assessed in the light of sustainable development, first, we need to know what kind of sustainability aspects need to be taken into account. One of these aspects obviously is the long-term effect on traffic flows. Other aspects could be economic effects to the region affected and environmental effects near the planned roads. The next step would be to find the most appropriate tools to assess these aspects. Possibly models exist that can be used. These models probably require exogenous input in the form of economic scenarios, current and expected traffic flows and information on job and business markets for the regions involved. Such information should ideally become available when using the handbook.

The European Commission's main interest in the outcome of the tool inventory and evaluation is an overview of tools that can be used under the umbrella of the Commission-style impact assessments[2] procedure. This procedure has been designed by the Secretariat

[1] See http://europa.eu.int/comm/research/fp6/index_en.html for further information about this 6th framework programme.
[2] Personal communication with Mr. Deybe, European Commission Directorate General for Research, Directorate I (Environment), Unite 1 (Strategy and policy for sustainable development)

General of the European Commission. The *SustainabilityA-Test* takes stock of and evaluates all kinds of tools used for all kinds of assessments, and thus not only those tools that are used under the umbrella of the impact assessment procedure. Considering the importance of the impact assessment procedure at the European level, the results of the tool inventory and evaluation will also be discussed in light of the Commission's impact assessment procedure.

The word 'tools' should be interpreted in the broadest sense, ranging from small aids to complex methodologies for carrying out an assessment. 'Assessments' in turn refers to all kinds of assessments to determine if and to which degree an observed development, or certain policy proposal, contributes to sustainable development. Both ex-ante and ex-post assessments are taken into consideration within *SustainabilityA-Test*, thus addressing the various stages of the policy cycle.

1.1.1 Goal, purpose and main output

The overall goal of the project *SustainabilityA-Test* is to:

1. support the definition and implementation of the EU Sustainable Development Strategy by describing, assessing and comparing tools that can be used to measure or assess sustainable development; and thus
2. improve the scientific underpinning of sustainable development impact assessment.

The purpose of *SustainabilityA-Test* is to:

1. examine the theoretical and conceptual basis of the tools and their uses based on a literature review of tool applications and a case study;
2. develop a synthesis in which the various tools and a number of assessment methods are compared with the requirements of sustainable development assessments and to formalise this synthesis in a tool framework.

SustainabilityA-Test will do so by generating the following outputs:

1. provide a consistent and peer-reviewed appraisal of the potential of common and emerging tools (i.e. methodologies, tools, approaches and appraisals) for sustainable development-related assessments in support of the various stages of policy;
2. make the appraisal of the tools vis-à-vis key aspects of sustainable development, as provided in the project proposal and to sharpen these key aspects on the basis of this project;
3. provide and apply a framework (matrix) for evaluation of the tools;
4. increase insights into how the various scientific tools relate to the requirements of participation and consultation;
5. disseminate the results widely among assessment practitioners as well as users;
6. identify important and promising issues for targeting subsequent research and development efforts;
7. build on the considerable knowledge with regard to integrated environment assessment that is available among the members of the European Forum for Integrated Environment Assessment (EFIEA) and in international organisations.

The main deliverable of the project will be the handbook, *Advanced Tools For Sustainability Assessments*, which aims to support policy makers on the verge of making a sustainability assessment in finding the most suitable tools for doing so.

1.1.2 Overview of the project as a whole

SustainabilityA-Test comprises three phases: the inception phase, the case study phase and the phase for integration and synthesis of the results.

The first phase of the project aims to create a provisional evaluation framework and a bundle of preliminary tool overview and evaluation papers. In addition, a detailed working plan for the case study will be set up, and a first outline for the final outcome of the project, the handbook *Advanced tool for Sustainability Assessments*, will be drafted.

The second phase of the project, the case study, is designed to further deepen and broaden the preliminary tool evaluation performed in phase 1. This work complements the tool-by-tool evaluation of phase 1 by analysing at national and European level how tools have been used, or could have been used, on the basis of two concrete policy cases (the Directive on Biofuels[3] and a regulation on Energy Crop Premium[4]). During the case study specific attention will be given to combinations of tools (including ‘recipes’), for the most part during the second half of the case study. In addition, the case study aims to further analyse (the European Commission’s) assessment practice and the role tools have in assessments.

During the third phase of the project, the integration and synthesis phase, the results of phases 1 and 2 will be integrated and synthesised. Due attention will be given to the link between different tools, the link between tools and assessment practice in general and the Impact Assessment procedure in particular. This phase will also draw conclusions with respect to the role of different tools in various stages of an assessment (e.g. scoping, assessing and interpreting the results), the role of scientific knowledge and the suitability of tools and, lastly, the role of participation and stakeholder consultation in assessments.

1.1.3 Project team

The project team of SustainabilityA-Test consists of 18 project partners from different institutes and disciplines (see Annex 1 for the list of project partners). For managerial reasons, several groups have been created: the tool teams, the I&S team and the peer group (see Annex 2 for an overview of the project teams and their members).

Tool teams

Most project partners are members of at least one tool team. Each tool team is responsible for a cluster of tools. The following tool teams exist (see Annex 2 for the members of each team):

- Tool team 1: Physical assessment tools
- Tool team 2: Monetary assessment tools
- Tool team 3: Modelling tools
- Tool team 4: Scenario analysis tools
- Tool team 5: Multi-criteria analysis tools
- Tool team 6: Sustainability appraisal tools

[3] Directive 2003/30/EC of the European Parliament and of the Council of 8 May 2003, (OJ L 123, 17.5.2003, p. 42-46); http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_123/l_12320030517en00420046.pdf

[4] Commission Regulation (EC) No 2237/2003 of 23 December 2003 (OJ L 339, 24/12/2003, p.52–69); http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_339/l_33920031224en00520069.pdf

- Tool team 7: Stakeholder analysis tools
- Tool team 8: Transition management tools

After analysing the preliminary tool overview and evaluation papers, in which the tool teams describe what tools will be evaluated, the project team came to the conclusion that for one tool group, ‘Transition management (TM) tools’, it is difficult to define actual assessment tools that belong under that heading, mainly because transition management is still a rather new area of work. The team responsible for the evaluation of transition management tools will therefore provide a background to TM based on the role of TM in sustainability assessments and describe how TM can be used as an instrument in assessments (following the common structure designed for the tool reports). Based on the latter, the team will also describe how (combinations of) tools can support the transition process at various stages. As the Matisse project could also benefit from these tasks and vice versa, tool team 8 will also act as an intermediary between these two projects, as far as it concerns tools and methods that can be used in the context of transition management.

I&S team

The leader of work package 1 (design) and 3 (integration and synthesis) is the same (RIVM); this is because of the nature of the work. The RIVM is supported in this task by the integration and synthesis team (I&S team). The I&S team’s main responsibility is to support the development of the evaluation framework, the verification of the objectivity of the tool evaluations and the integration of the results.

Originally, the I&S team consisted of representatives from EC-JRC, IVM-VUA, Tyndall-UEA, USF-UOS and RIVM. During the first phase of the *SustainabilityA-Test* it became evident that the leader of work package 2 (the case study) should become a member of the I&S team too. Also the tool expert on the Commission-style Impact Assessment (IA) procedure became a member of the I&S team, since this IA-procedure is such an important tool in the Commission. Integration of knowledge with respect to this assessment procedure and the evaluated tools is crucial for the success of the project as a whole (see Annex 2 for a list of all I&S team members)

During the final phase of the project each of the seven the tool team leaders will become a member of the I&S team in order to guarantee an effective integration of the project’s outcomes.

Peer group

A peer group strategically guides the project (see Annex 2 for its members). This group includes the overall manager of the project, the 3 other leaders of work packages, a representative from the EEA and the European Commission desk officer for *SustainabilityA-Test*. The peer group meets at least 6 times (around crucial milestones and the major workshops) during the whole project.

1.2 Progress so far

In the next sections an update is given of the progress made within each work package.

1.2.1 Work package 1: inventory of tools and tool-by-tool evaluation

Work package 1 comprises the development of an evaluation framework, an inventory of tools and a preliminary evaluation of the tools. These tasks support the stock-taking of tools and their preliminary evaluation. The design phase has also been used to analyse assessment practice at the Commission and – more general – to analyse what sustainability assessment is about (or should be about). Both tasks, the bottom-up tool inventory and evaluation and the top-down assessment analysis, provide a starting point for, and will continue during, work package 3 (integration and synthesis).

Work package 1 was originally planned to last 9 months. However, the importance and complexity of this inception phase forced the project team to take more time for it. Phase 1 lasted until February 2005 instead of November 2004. These three months of delay will be compensated during the remainder of the project.

Evaluation framework

The evaluation framework has been designed in several steps. First, the evaluation criteria were developed, described in the 'Analytical Framework' and reviewed by the project partners in June 2004. Second, an improved set of evaluation criteria, together with an evaluation instruction and reporting format, have been described in the 'draft methodology report' (D3), disseminated prior to the inception workshop in September 2004. On the basis of comments received from the tool teams on the evaluation criteria, on the instruction and on the reporting format, minor adjustments were made to it. From September 2004 onwards, the tool teams started using the evaluation framework. Results of this are presented in the bundle of preliminary tool overview paper (see www.SustainabilityA-Test.net) of which a brief summary is given in chapter 4. Experiences during this first phase of the project with the applicability of the evaluation framework are discussed in chapter 5.

Preliminary tool overview and evaluation papers

The preliminary tool overview and evaluation papers contain an overview of tools and the results of the preliminary evaluation by means of the evaluation framework. In addition, existing experiences with the tools are described in these papers, as well as research questions and challenges associated with the tools that will be addressed during *SustainabilityA-Test*. Each tool team is responsible for drafting such paper. The first versions of the papers were drafted by August 2004 and discussed at the inception workshop in September 2004. The final draft papers were ready by January 2005 and bundled and disseminated in February 2005 (see also www.SustainabilityA-Test.net).

1.2.2 Work package 2: case study

A concrete proposal for the content and design of the case study was distributed to all project partners in December 2004 for review. This proposal was the result of intensive discussion between the members of the I&S team and between the I&S team and all other project partners. Details with respect to the case study are included in §3.3.

1.2.3 Work package 3: integration and synthesis

The work done so far with respect to the integration and synthesis of results is for the most part captured in work package 1 (design). The results of work package 1 will feed into work package 3.

1.2.4 Work package 4 and 5: management and dissemination

Dissemination activities: handbook
The handbook *Advanced Tools for Sustainability Assessments* is the main output of the project. The handbook will assist those on the verge of carrying out an assessment to find the most suitable tools to do so, which could also include steering the user of the handbook towards formulating suitable assessment questions. An outline for the handbook was developed and discussed in July 2004 together with *SustainabilityA-Test*'s desk-officer. Ideas for the handbook were further developed afterwards and presented and discussed at I&S team meetings and at the inception workshop. A more detailed proposal for the content and structure of the handbook is presented in §3.3.1 and will be further discussed with all project partners (at the case-study kick-off meeting and thereafter).

Dissemination activities: other
The public project website (www.SustainabilityA-Test.net) gives a summary of the project, an overview of the project partners and the calendar of events. Also the deliverables of the project can be downloaded from here. In November 2004, a meeting took place with the project coordinator of the SEAMLESS project. In this meeting the possible contributions of the SustainabilityA-Test project in the SEAMLESS project and vice versa was identified. Furthermore, a contribution has been made to a project proposal initiated by Pietro Caratti of FEEM (the MAPSIA proposal submitted under the FP6-2004-Global-3 call). Within this project we want to bring together the coordinators and key partners of the main IA projects (SustainabilityA-Test, Insure, IQ Tools) to exploit synergies with regard to dissemination and networking. Finally, the project managers of the following projects were invited to attend the case-study kick-off meeting in March 2005 in Berlin: SENSOR, SEAMLESS, IQ tools and Insure.

Management
To exchange information and documents within the project easier and more consistently, an Interest Group at the CIRCA website has been utilised. To ensure that the project is run according to plan, on time and within the budget, regular contact has been made with the project participants through email and by telephone. In consultation with the peer group the deliverables D5 (inception report) and D6 (the bundle of preliminary tool overview paper) were postponed by a few months and the I&S group supplemented by two persons (see Annex 2).

To date, the following meetings have been organised:
- Kick-off meeting on April 15-16, 2004, IVM, Amsterdam
- First Peer Group meeting on April 14, 2004, IVM, Amsterdam
- I&S meeting on August 23-24, 2004, RIVM, Bilthoven
- Inception meeting on September 9-10, 2004, IVM, Amsterdam
- Second Peer Group meeting on September 8, 2004, IVM, Amsterdam
- I&S Telephone conference on October 13, 2004
- I&S Telephone conference on November 22, 2004
- I&S meeting on January 17-18, 2005, IVM, Amsterdam
- Case study kick-off meeting in March 2005, Berlin

1.3 Aim of this report

This inception report presents the outcome of phase 1 (work package 1) of *SustainabilityA-Test*. It presents terminology that will be used during the project (chapter 2), the methodology for tool inventory and evaluation, which includes an evaluation framework (chapter 3), the results of the preliminary evaluation (chapter 3) and a discussion of these (chapter 4), on the basis of which further research questions for the remainder of the project are formulated (chapter 5). This should ensure that the whole project team will start from a common understanding of the different concepts relevant to the context of sustainable development and the methodology used. As such, the inception report becomes a 'book of reference' for the project. By means of this inception report, work package 1 has been concluded.

2 The concept of sustainable development in *SustainabilityA-Test*

Sustainable development is a term that can be interpreted in different ways. Scientific discourses have, for decades, been trying to define it. Still, most of such debates have not resulted in a generally acknowledged definition, and most likely, they never will.

The biggest challenge lies in agreeing on the exact meaning of sustainable development with respect to analysing behaviour and/or actions we take (e.g. policy interventions): will certain behaviour or actions lead to becoming sustainable? Is certain behaviour or are certain actions in line with sustainable development?

In this chapter we will argue the point that the one exact definition does not exist. Depending on person's world view, many different ideas exist about what sustainable development is about. What one person considers to be perfectly sustainable could be considered unsustainable by others.

Within a project like *SustainabilityA-Test* we need a common understanding with respect to sustainable development. This chapter sets out a way to understand sustainable development without defining exactly what it means for our current behaviour and/or policy interventions. By doing so, we conceptualise sustainable development to make it concrete enough to work with, while avoiding the risk of getting stuck in, or (re)opening the debates about, its exact definition.

SustainabilityA-Test also uses a characteristic terminology with respect to tools. This terminology is explained in chapter 4, in which all tools considered in the project are presented as well.

Below, we will first discuss the basics of sustainable development for as far as these are commonly agreed upon. The following section will discuss the importance of world views, followed by a section that explains how sustainable development can be conceptualised so that it can actually be used in assessments. The chapter ends with discussion on different types of assessments (mono-disciplinary, integrated, sustainability assessments) and the role of science in sustainability assessments.

2.1 Sustainable development: the parts we agree upon

Sustainable development is derived from sustainability. It is useful to briefly discuss the meaning of sustainability before embarking on discussing sustainable development.

2.1.1 Sustainability from the dictionary

Sustainability is the 'property of being sustainable'[5] and sustainable means 'capable of being sustained or maintained'[6], 'to keep in existence; to maintain'[7]. In an anthropocentric view, to sustain refers to *humankind* sustaining itself on earth.

Earth, the natural system in which humankind lives, continuously changes, making it possible or impossible, easier or more difficult, for humankind to exist. Thus, to sustain humankind on earth refers to sustaining humankind living and interacting with and within a natural system that continuously changes. Therefore, in the strictest sense of the word, sustainability means:

> The property of being capable to sustain (i.e. to maintain) earth or to sustain humankind on earth living in and interacting with a natural system that continuously changes.

Although many more variations with respect to what sustainability means exist[8], this variation covers most relevant parts of it. Sustainability is, however, often mentioned in the context of sustainable development and many times considered equal to it.

2.1.2 From sustainability to sustainable *development*: a political evolution

When the concept of sustainability landed in the political arena, it evolved gradually into the term sustainable development. It is worthwhile examining the history of this process.

The quality of the natural system we live in determines the ease to sustain mankind on earth. For that reason, *sustainability* was initially mainly connected with the ecological qualities of our world. That these qualities are influenced by humankind and that they could perhaps even be finite, was brought under the attention by the report *Limits to growth* in 1972 (Meadows et al., 1972). This report concluded that present growth trends are unsustainable and that bringing them to a halt is a necessity for preventing further deterioration of the world's ecological qualities:

> The present growth trends in world population, industrialisation, pollution, food production, and resource depletion continue unchanged, the limits to growth on this planet will be reached sometime within the next one hundred years. […] It is possible to alter these growth trends and to establish a condition of ecological and economic stability that is sustainable far into the future.

In the same year, at the United Nations Conference on the Human Environment (Stockholm, 5 to 16 June 1972) the developing countries brought forward that not until the gap between the poor and the rich countries was substantially narrowed, any progress could be made in improving the human environment[9]. These concerns were reflected in the adopted declaration by linking (ecological) sustainability explicitly to socio-economic development[10]:

> '[…] the developing countries must direct their efforts to development, bearing in mind their priorities and the need to safeguard and improve the environment. For the same purpose, the industrialised countries should make efforts to reduce the gap between themselves and the developing countries.'

---

[5] www.dictionary.com: WordNet (r 2.0), August 2003, Princeton University.
[6] Webster's Revised Unabridged Dictionary, MICRA, 1998.
[7] The American Heritage Dictionary of the English Language, Fourth Edition, Houghton Mifflin Company, 2000.
[8] A brief summary can be found at page 93, box 1.2 of the IPCC report 'Climate Change 2001 – Mitigation', available at http://www.grida.no/climate/ipcc_tar/wg3/060.htm
[9] See the brief summary of the general debate at http://www.unep.org/Documents/Default.asp?DocumentID=97.
[10] See: http://www.unep.org/Documents/Default.asp?DocumentID=97&ArticleID=1503

The next Summit on Environment and Development, held in Rio in 1992, contributed further to the shift from the ecologically inspired notion of sustainability towards sustainable *development*, which is also inspired by socio-economic concerns (MNP, 2004). It also emphasised that sustainable development had to take place without transferring damage from local and current behaviour to other areas or future generations[11], in line with the definition of sustainable development that the World Commission on Environment and Development, better known as the 'Brundlant Commission', gave to it five years earlier (WCSD):

> To meet the needs of the present without compromising the ability of future generations to meet their own needs.

At the most recent World Summit on Sustainable Development, held in Johannesburg in 2003, the so-called pillars of sustainable development were specifically mentioned in the adopted declaration. According to it, people have to 'assume a collective responsibility to advance and strengthen the interdependent and mutually reinforcing pillars of sustainable development – economic development, social development and environmental protection – at the local, national, regional and global levels'[12]. Note that *sustainable development* moved further away from ecological sustainability by referring to economic development and social development in the framework of environmental protection. In a way, the focus on the preservation of our ecological qualities has been gradually replaced by the aim to facilitate economic and social development within certain ecological boundary conditions. *Development*, originally connected to ecological sustainability by developing countries, is now considered a necessity not only for the developing countries, but also for the developed world.

2.1.3 Three dimensions of sustainable development: here and now, elsewhere and later

As explained above, sustainability has gradually evolved into sustainable development. At the same time the so-called pillars of sustainable development came in, emphasising the existence of and need for striking a balance between economic, social and environmental qualities. Worldwide debates also added that the currently developed countries had their responsibility with respect to the less developed countries and that our current generation had its responsibility for future generations.

It is therefore that sustainable development has evolved into the idea of striking a balance between certain economic, social and environmental qualities along the following three dimensions:

- here and now
- elsewhere; and
- later

These three dimensions are an important step forward in making concrete what we should take into account when assessing for sustainable development. However, it is still not concrete enough: more guidance is needed to determine what actually the 'right balance' between these dimensions is. Which qualities need be guaranteed for economic, social and

[11] Principles 2 and 3 of the Rio Declaration on Environment and Development, 1992.
[12] See number 5 of the Johannesburg Declaration (http://www.un.org/esa/sustdev/documents/WSSD_POI_PD/English/POI_PD.htm).

environmental concerns, what quality is desired elsewhere, and how long? In other words: we have to conceptualise sustainable development further.

2.2 Sustainable development: a diversity of views (world views)

One way of conceptualising further the notion of sustainable development is by reaching consensus about where lies the balance between economic and social development and ecological qualities, between local economic, social and environmental qualities and those elsewhere, and between present qualities and future ones. In this section, however, we argue that it is impossible to precisely define this equilibrium, as the relative importance of economic, societal and environmental concerns, the local and global scale, and present and future generations, depends on a person's 'world view'.

What one person believes to be sustainable development could be interpreted by another as something that is unsustainable development. In general, what is considered to be *sustainable* depends on a person's personal values (i.e. his/her world view) (MNP, 2004). To illustrate this, the characteristics of two different world views on the role of technology in the future are sketched out below (translated from (MNP, 2004) – consistent with the IPCC-SRES scenarios (IPCC, 2000)):

> World view 'A1 / global market': rapid technological developments. Confidence in bio-technology, new materials, health-technology, GMO, nuclear power. Research financed through internationally operating industries. Emphasis in industrialised countries on the replacement of semi- and unskilled workers by knowledge, capital goods, energy and materials.
>
> World view 'B2 / caring regions': local inventions in new forms of governance and cohabitation ('social technology') are more likely than 'hard' technological breakthroughs. If necessary, energy and primary materials are replaced by labour and recycled materials.

Recognising that different world views exist is important for understanding that consequently different views exist on what is considered to contribute to *sustainable development*. People seek different objectives, have different ideas about the availability of collective means that can be used to reach these objectives and they consider themselves to a greater or smaller degree co-responsible for how these collective means should be distributed. In addition, people have different views on the value of scientific knowledge and on what risk levels (associated with not knowing everything) are acceptable. Therefore people identify different sustainable development problems and solutions (MNP, 2004; MNP/TNS-NIPO, forthcoming).

2.3 Aspects of sustainable development

From the above it can be concluded that it is impossible to define the exact balance between economic, social and environmental concerns, here and now, elsewhere and later, that contributes to sustainable development. We can, however, find aspects that *could* be considered important in the context of sustainable development. These aspects can be used to describe this balance, irrespective of where exactly it lies.

In *SustainabilityA-Test*, an aspect of sustainable development is some topic, issue, problem, challenge, quality or dimension that is considered relevant (by some or by many) in the context of sustainable development. Some of these aspects have already been mentioned: economy, society, environment, here, elsewhere, now and in the future. These are too aggregated to be useful and need be further specified.

There are two main lines of approach for creating a detailed list of aspects of sustainable development: bottom-up and top-down.

- Bottom-up: this approach tries to list *all possible aspects* one could think of in the context of sustainable development. Such a list is inexhaustible. Aspects are added to the list by advancing (scientific) knowledge. When, for example, the presence of a substance appears to be harmful for humans, the concentration of that substance in the air becomes an aspect relevant for sustainable development.
- Top-down: this approach tries to list all aspects that are specifically linked to *principles* – also referred to as *main challenges* or *objectives*, *priorities*, *goals*, *targets* et cetera – of sustainable development. These principles are usually agreed upon by a certain institutional body, or by (groups of) countries. Examples of principles are the ones found in the Rio Declaration on Environment and Development (UN, 1992), the Millennium Development Goals (UN, 2000) and the European headline objectives for sustainable development (CEC, 2001).

There is a significant difference between the lists of aspects created by the two approaches. In the bottom-up list, an assessment becomes ticking off whether an aspect is affected and if so, determining the magnitude. In reality, the person responsible for carrying out the assessment will most likely predetermine which aspects need further analysis and which can be left aside. Scoping, i.e. the process of determining what to include in the assessment and what not to include, therefore becomes a crucial step in the assessment. The interpretation of the final outcome of the assessment is critical too when using a bottom-up list of aspects, because the aspects are not linked to objectives per se. The lack of objectives means that separating the acceptable effects from unacceptable ones is left to the policy maker. When using the top-down list, scoping and interpretation become less crucial, assuming that the assessment will analyse only those effects that are set as priority objectives in a prevailing strategy.

In *SustainabilityA-Test* we will use both the bottom-up and top-down approach for selecting aspects that are relevant in the context of sustainable development.

2.3.1 Creating a list of aspects I: bottom-up (all aspects)

A common way for finding all possible aspects that could be relevant when assessing in the light of sustainable development is by dividing our complex world into a number of less complex subsystems and identifying topics that characterise each subsystem. The three-pillar approach, dividing the world into economy, environment and society, is perhaps the best known example of this approach. It is derived from the notion of sustainable development that emphasises finding the right balance between economic development, social development and environmental protection[13].

[13] See, for example, Johannesburg Summit.

Different pillars and different numbers of pillars can be found in literature. The preferred number and exact description of pillars depends on the emphasis that one wishes to put on a certain subsystem. The most common pillar approach is ‘people, planet, profit’, or ‘society, environment and economy’.

Three-pillar aspects
Within *SustainabilityA-Test* we shall use this three pillar approach. The main reason for that is its long history and, as a result, its function as a basis for many other pillar-based varieties[14]. The pillar approach is not without criticism, though. Most concerns relate to the risk of overlooking interdependencies and interconnectivities between the pillars, making it difficult to identify trade-offs and win-win situations.

Crosscutting aspects
The three-pillar aspects alone are incapable of capturing all relevant aspects in the context of sustainable development. There are aspects relevant for sustainable development that cannot be captured by the pillars. Two of these have been mentioned before: the transfer of our local problems to other areas (‘elsewhere’) and the transfer of our current problems to future generations (‘later’). To provide room for those aspects, the category of crosscutting aspects is created in *SustainabilityA-Test*.

The lists of three-pillar and cross-cutting aspects are presented in §3.1.2.

Box 2.1: Why ‘cross-cutting’?

‘Cross-cutting aspects’ is used to refer to those aspects that in fact can be applied to each aspect that can be found in one of the three pillars (economy, society and environment). Applying a cross-cutting aspect to one of these aspects will broaden it. An example to illustrate this: when assessing a European proposal concerning the appropriateness of nuclear power, we can identify relevant aspects in each pillar (e.g. waste in the environmental, costs in the economic and regional employment rates in the social pillar). Adding a cross-cutting aspect like ‘intergenerational effect’ (i.e. looking at the effect of present action on future generations) will broaden the three mentioned aspects to waste, costs and regional employment effects in about 20–30 years from now.

2.3.2 Creating a list of aspects II: top-down (aspects relevant for principles of sustainable development)

Principles of sustainable development, also known as sustainability criteria, describe the main challenges we stand for (or better: we recognise) to become more sustainable. There have been many initiatives to develop them. An impressive collection has been composed by the International Institute for Sustainable Development[15].

It is impossible to identify one list of sustainability criteria that is better than all the others. *SustainabilityA-Test* will evaluate tools used in EU policy making. It is therefore logical to look for a list of principles for sustainable development that is endorsed by the European Council and that sets the framework for the European Commission. This is the EU Sustainable Development Strategy. The headline objectives set out in this strategy will function as aspects of sustainable development from a top-down perspective. What aspects exactly can be extracted from the strategy will be described in §3.1.2.

---

[14] The three pillars are formulated in the Rio Declaration by the principles 3 to 5 (see http://www.unep.org/Documents/Default.asp?DocumentID=78&ArticleID=1163)
[15] http://www.iisd.org/sd/principle.asp

2.4 Sustainability assessment

Sustainability assessment, *sustainable development assessment* and *assessments in light of sustainable development* will all mean the same in *SustainabilityA-Test*. *Sustainability assessment* (SA) will be used from now on solely for the reason of it being shorter than the other options. In this section we clarify what is actually meant with *sustainability assessment* in the context of *SustainabilityA-Test*.

Within the project we primarily consider sustainability assessments carried out within the government domain. These assessments can be done for different reasons, ranging from ex-ante impact assessments (to appraise if a proposed government intervention will contribute to sustainable development) to ex-post assessments and evaluations (to appraise if an intervention has contributed to sustainable development). In principle, sustainability assessments can be made to support any kind of policy process found in a policy-making cycle.

2.4.1 Minimum standards for sustainability assessments

Not all assessments are *sustainability* assessments in the sense of how sustainability and sustainable development is described in §2.1. An assessment of the short-term economic, social and environmental impacts, for instance, within Europe of a certain European proposal cannot be considered a *sustainability* assessment. Two important dimensions that are embedded in the notion of sustainable development lack in such assessment: the effects of that proposal elsewhere (in this case outside Europe) and the effects of that proposal in the future (long-term effects).

The completeness of an assessment can be described in terms of the level of integrating different disciplines and the level of taking into account the external dimension and the longer-term time horizon. Within *SustainabilityA-Test* we draw an arbitrary line between *sustainability* assessments, *integrated* assessments and *mono-disciplinary* assessments to express the fact that a sustainability assessment should contain certain minimum standards.

Sustainability assessments
Sustainability assessments refer to assessments that bring together as many relevant aspects in the context of sustainable development as possible. This implies that the following three dimensions should be included:

1. here and now – the balance between the three pillars of sustainable development;
2. elsewhere – the balance between the external and internal dimension;
3. later – the balance between current needs and long-term needs.

In order to tackle everything, sustainability assessments could be made by carrying out various integrated and/or mono-disciplinary assessments in parallel.

Integrated assessment
The term *integrated* assessment has in the past been used for assessments that integrate *environmental* concerns in different types of sector specific assessments. In *SustainabilityA-Test* we use this term to refer to multi-disciplinary assessments that cannot really be regarded as sustainability assessment because of a serious deficit in the coverage of such assessment

(e.g. by not addressing the external dimension or long-term effects). A sustainability assessment is an integrated assessment per se, but not the other way around.

Mono-disciplinary assessments
Mono-disciplinary assessments refer to assessments that focus on one (scientific) discipline. For example, an assessment calculating the CO_2 emission reduction potential would be considered a mono-disciplinary assessment in *SustainabilityA-Test*. Mono-disciplinary assessments are often the building blocks for integrated and sustainability assessments.

2.4.2 The role of knowledge in (sustainability) assessments[16]

Not everything can be assessed and we do not always know what exactly to assess, how to assess it and how to use the outcome. In other words: it is not always clear what assessment questions to ask (scoping), what tools best to use to answers the questions and how to interpret the answers in the decision-making process.

There are two key factors determining if and how assessment questions can be answered and what the value of these answers could be in the decision-making process (see also Box 2.2, based on Hisschemöller and Gupta, 1999):

1. the level of consensus on scientific knowledge about the issue at stake;
2. the level of consensus on values[17] about the issue at stake.

These key factors determine the role of science and thereby the role of assessment tools in assessments and decision making processes:

- The outcome and usage of an assessment is least contested when it concerns assessments of issues with a high level of consensus on the applied scientific knowledge and on the values on the issues at stake. In these cases, tools are often used to assess something and the methodology and outcome are not much contested;
- When little consensus exists on how to use the outcome of an assessment, despite a high level of consensus on scientific knowledge, the decision maker has the obligation to clarify that such an outcome has different meanings in different world -views. People will respond differently to the question how to weigh GDP loss with healthier trees; science, and thus tools, can be used to *mediate* between the different world views.
- When there is no consensus on scientific knowledge, but high consensus on values of the issues at stake, the policy maker should realise that the outcome of an assessment is 'just' one outcome: using different scientific approaches could lead to completely different outcomes of the assessment. Science and tools can be used to *advocate* different scientific opinions.
- Finally, when no consensus exists on scientific knowledge and no consensus on values, a policy maker should realise that an assessment cannot be much more than a tentative exploration of unknown domains. Science and tools can be used as problem recogniser, after which gradually a common recognition of the problem might rise (i.e. increasing consensus on values).

Obviously, the role of scientific knowledge in assessments and in the interpretation of the assessment outcome is important for the selection of 'the most suitable tools' for carrying out an assessment. In determining how best to answer certain assessment questions and how best

[16] This paragraph is largely based on (Boersema Jan J. and Reijnders, Lucas (eds), to be published) and (Bert de Vries, 2004)
[17] A person's set of values determines his/her world view, and, vice versa, world views influence the formation of different values.

to interpret assessment outcomes the thinking framework as discussed here could be useful. It offers a structured approach to linking the formulation of assessments questions to tool selection, and to linking the outcomes of assessments to the decision making process.

Box 2.2: A framework of thinking for the role of science in assessments

As explained in the main text, science has different roles in assessment situations with high or low consensus on values. These values represent value-laden interpretations of situations or of problems at stake. To illustrate this: some are in favour of addressing climate change by abating emissions, whereas others are in favour of addressing climate change by adapting to it.

Figure 2.1 below provides a framework of thinking for the interaction between scientific consensus and consensus on values.

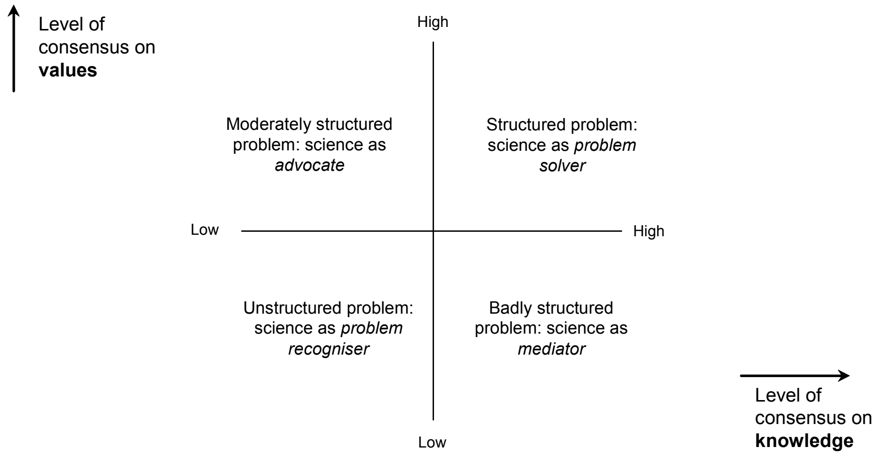


Figure 2.1: A framework of thinking for the level consensus on scientific knowledge and values (after Hisschemöller and Gupta, 1999)

This framework is shortly explained by describing each quadrant of Figure 2.1 and by giving an example:

- Top-right quadrant: there is strong consensus on knowledge and values. Assessments can be carried out by using approaches that are not contested and that can be used for providing answers to the questions posed. Both the outcome of an assessment and the meaning of this outcome for the decision maker are likely to be broadly accepted and supported. Example: air dispersion models.
- Top-left quadrant: there is little consensus on knowledge and a high level of consensus on values. The lack of knowledge consensus provides room for alternative scientific approaches and thus for e.g. alternative tools used in assessments. The outcome of an assessment will be ambiguous and risks being contested by advocates of different scientific approaches. Progressing scientific knowledge is expected to gradually decrease these controversies. Example: the potential of renewable energy sources.
- Bottom-right quadrant: there is strong consensus on knowledge, but little consensus on values. The lack of consensus on values provides room for different interpretations of the scientific results, and thus for alternative ways to incorporate assessment results in the decision making process. The outcome of the assessment itself is unambiguous, but the meaning of that outcome is. Science can be used to mediate between different opinions. Example: a cost-benefit analysis (assumed that the methods used are broadly accepted) and the usage of that outcome in decision making.
- Bottom-left quadrant: there is no consensus on knowledge and no consensus on values. One could speak of chaos. The role of science is to tentatively explore the unknown zone for problem recognition. In terms of assessments, this situation corresponds to not knowing if certain aspects are important and how these aspects could be assessed. Example: consequences of changes in biodiversity.

3 *SustainabilityA-Test*'s methodology

This chapter describes the different building blocks of *SustainabilityA-Test* in more detail. It successively discusses:

- the evaluation criteria;
- the evaluation methodology;
- the case study; and
- the handbook.

A more detailed description of each tool that will be evaluated and the results of the evaluation itself can be found in the next chapter.

3.1 Evaluation criteria

The evaluation framework has been designed in several steps. First, the evaluation criteria were developed, described in the 'Analytical Framework' and reviewed by the project partners. Second, an improved set of evaluation criteria, together with an evaluation instruction and reporting format, have been described in the 'draft methodology report' (D3), disseminated prior to the inception workshop in September 2004. On the basis of comments received from the tool teams on the evaluation criteria, on the instruction and on the reporting format, minor changes were made to it. From September onwards, the tool teams have started using the evaluation framework. Results of that are presented in the bundle of preliminary tool overview paper (see www.SustainabilityA-Test.net) of which is brief summary is given in chapter 4. Experiences during this first phase of the project with the applicability of the evaluation framework is discussed in detail in chapter 5.

In this section the tool evaluation criteria are presented (how these criteria are used is subject of §3.2). All tools evaluated within *SustainabilityA-Test* will be evaluated for:

- their ability to support certain *policy processes*;
- their ability to address various *aspects of sustainable development*;
- their costs, time needs, level of required expertise et cetera (*operational aspects*).

Each category of evaluation criteria is discussed in further detail in the next paragraphs.

3.1.1 Policy processes

There is a variety of types of policy processes thinkable in which sustainability assessments could play a role. Table 3.1 lists these processes. It is an indicative list, which is based on a theoretical framework for policy analysis (Brewer, DeLeon, 1983). Using this theoretical framework ensures having a structured approach for the identification of policy processes.

With this list we aim to have a list of policy processes that could *possibly* be supported by (sustainability) assessments. It is by no means said that we need *all* processes listed here. The experiences with evaluating the tools by means of these policy processes will clarify the need for making adjustments to them.

Table 3.1: Policy processes

Policy processes	Explanation
− Recognition of a problem	The process of analysing if (or discovering that) an observed development/trend leads to a problem (including drawing conclusions from monitoring and strategic outlooks, and the process of getting the problem on the political agenda)
− Investigating the nature of a problem and identifying conflicting assumptions with respect to the problem situation	The process of analysing the problem in further detail, in order to understand the nature of the problem (i.e. conceptualising and outlining the problem, identifying the driving forces underlying the problem, causal relations between the driving forces and the observed problems, and the identification of conflicting assumption) by means of strategic outlooks, models et cetera
− Identification of possible solutions to alleviate, mitigate or resolve the problem	The process of generating ideas to address the problem, including the collection of information and data needed to lay out a range of possible responses and the further specification of potential policy choices within that range, and including the process of tentative burden sharing analysis
− Analysis of policy proposals (ex-ante)	The process of analysing the likelihood that any of the policy proposals will prove to be a success or failure by means of predetermining the risks, costs and benefits associated with each proposal, using empirical, scientific and/or projective knowledge (including the process of screening for possible impacts)
− Selection of a policy option	The process of the authoritative policy maker or body debating and exploring and comparing in detail all policy proposals in order to reduce the level of uncertainty and to reach the best decision, incorporating into the decision making process all the work that has been done prior to this stage
− Implementation of the selected policy option	The process of developing Directives, regulations, guidelines et cetera to execute the selected policy option, including the definition of policy goals (targets) and the burden sharing analysis supporting that
− Evaluation of the selected and implemented policy option (ex-post)	The process of assessing the efficiency and results (and possibly other aspects) of the selected and implemented policy option, including the process of determining what aspects/impacts should be accounted for during this evaluation and the process of actually carrying out the evaluation
− Discontinuation of poorly performing or unnecessary policy options	The process of the authoritative policy maker or body debating and exploring in detail if policy options are poorly performing or unnecessary, in order to reach a decision on the termination of such policy and to specify new problems stemming from termination, if any

3.1.2 Aspects of sustainable development

As mentioned in §2.3 two approaches exist for creating a list of aspects of sustainable development: bottom-up, creating a list of all conceivable aspects in light of sustainable development, and top-down, creating a list of aspects that are relevant in light of principles of sustainable development endorsed at EU level.

Bottom-up: three-pillar aspects and cross-cutting aspects

The number of aspects that could belong to each of the three pillars and the category of crosscutting aspects is in principle inexhaustible. The European Commission’s Handbook ‘How do an Impact Assessment’ (CEC, 2003) supplies a long list of aspects that could be taken into consideration during an impact assessment forms a good starting point for such list. Other lists exist and have been analysed in order to complement the list from the European Commission’s. These are:

1. **SDI**: Sustainable Development Indicators, developed by the SDI-taskforce[18] and designed to monitor progress with the European Commission’s Sustainable Development Strategy[19];
2. **SI**: Structural Indicators, developed by Eurostat, and designed to report progress made with the Lisbon-agenda in annual spring reports from the European Commission to the European Council[20];

[18] Members of this task force are Eurostat (lead) and representatives from EU Member states, EFTA countries, other parts of the European Commission, the EEA, OECD and UNCSD.
[19] See http://forum.europa.eu.int/Public/irc/dsis/susdevind/library

3. **IA^{STAR}**: factors for sustainability appraisals, developed by the European Commission's Joint Research Centre[21].
4. **EEA's key aspects of SD:** the European Environment Agency (EEA) developed a list of eight aspects that form a good basis for a list of the crosscutting sustainability aspects (unpublished).

Taking these four lists together, an indicative list of aspects that can be placed in each pillar (see Table 3.2) and in the category of crosscutting aspects (see Table 3.3) can be constructed. The aspects listed in Table 3.2 are main categories only; details of what exactly falls within each category can be found in Annex 3. Just as for the policy processes both lists are indicative and aim to include the most relevant and commonly seen aspects within each pillar of sustainable development and with respect to the cross-cutting aspects. Note that the list of crosscutting aspects contains alternative ways of looking at the dimensions 'elsewhere' and 'later' (see §2.1). Both lists are a functional starting-point for the evaluation and most likely need to be adjusted in the course of the project.

*Table 3.2: List of main categories of environmental, social and economic aspects**

Environmental	Social	Economic
– Air, water, soil or climate – Renewable or non-renewable resources – Bio-diversity, flora, fauna – Land use – Natural and Cultural heritage – Waste production/generation or recycling – Human safety or health – The likelihood or scale of environmental risks – Mobility (transport modes), or the use of energy	– Social Cohesion – Employment Quality – Public health – Health systems and security – Social Protection and Social Services – Consumer interests – Education – Social Capital – Liveable communities – Equality of opportunity and entitlement – Culture – International co-operation – Governance and participation – Fundamental human rights – Security, crime or terrorism – Ageing of society and pensions	– Economic growth – Price levels and stability – Effects on public authority budgets – Human capital formation and employment – Economic cohesion – Innovation – International performance – Market structure – Microeconomic effects on enterprises, non-profit organisations etc. – Effects on households – Global partnership

*) A more detailed list of aspects that fall under each main category listed in this table is given in Annex 3.

[20] See http://europa.eu.int/comm/eurostat/newcronos/reference/display.do?screen=welcomeref&open=/&product=EU_strind&depth=2&language=en
[21] See: http://www.jrc.es/projects/iastar/

Table 3.3: List of crosscutting aspects

Short name	Explanation
Elsewhere	
Distributional effects	The distribution of (dis-)advantages over different regions, societal/income groups, sectors etc.
Global dimension	Worldwide effects, or effects outside Europe
Spatial scale	Magnitude of impacts in terms of spatial scale (local, regional, national, European and/or global)
Later	
Inter-generational effects	Long-term (at least one generation, 25-odd years) effects in the social, economic and environmental potential – 'potential' can be conceptualised by looking at stocks, or capacity, found in the economic, social and environmental domains
(Ir-)reversibility	The arising of (ir-)reversible and long term effects to economies, societies/humans and/or ecosystems
Adaptability	Changes in the capability of the economic, environmental and social system (or the system as a whole) to adapt to external influences
(De-)coupling	Effects to the economy's resource use compared to the economic development, or effects to welfare growth compared to consumption of natural resource

Top-down: aspects linked to principles

The second approach towards creating a list of aspects of sustainable development is by analysing what sustainability criteria, or principles, have been formulated at a certain policy level. As *SustainabilityA-Test* serves the European Commission, the policy level to look at is the European policy level. This section will explain what aspects linked to principles of sustainable development, formulated at the European level, can be distinguished.

There is one prevailing (internal) European strategy for sustainable development, adopted in 2001: 'A Sustainable Europe for a Better World: A European Strategy for Sustainable Development' (CEC, 2001). The external dimension, added to this strategy in 2002, is formulated in the Commissions communication 'Towards a global partnership for sustainable development' (CEC, 2002a). Both documents together form the prevailing strategy for sustainable development in Europe.

Twelve issues that pose the biggest challenge to sustainable development in Europe are identified in the EU SDS (including its external dimension):

1. limit climate change and increase the use of clean energy
2. address threats to public health
3. manage natural resources more responsibly
4. improve the transport system and land-use management
5. combating poverty and social exclusion (Lisbon)*
6. dealing with the economic and social implications of an ageing society (Lisbon)*
7. harnessing globalisation: trade for sustainable development (ensure that globalisation contributes to sustainable development)
8. fighting poverty and promoting social development
9. sustainable management of natural and environmental resources
10. improving the coherence of European Union policies
11. better governance at all levels
12. financing sustainable development

The priorities marked with an (*) were already identified at the Lisbon European Council in 2000 (Council of the European Union, 2000). These so-called 'Lisbon-objectives' form an integral part of the EU SDS. Priorities 7–12 come from the 'external dimension' added to the 'internal' Gothenburg strategy.

These twelve priorities are still textual and need to be transformed into concrete 'topics' that can be used for determining what topics can be covered by a tool. In other words, an indicator (or a set of indicators) needs to be attached to each of them. This can be done by taking the indicators that are actually being used to monitor progress with these priorities.

For monitoring the EU SDS the 'Sustainable Development Indicators' are being developed (Eurostat, 2004). This is a process that should lead to a final report by the end of 2004 in which a set of indicators will be proposed together with a critical assessment of the main difficulties and of the main data and conceptual gaps that remain in the European Statistical System. Preliminary indicator lists are already available[22]. The Lisbon strategy has its own indicators: the 'structural indicators'[23]. These cover only part of what is covered by the Sustainable Development Indicators (mainly economic and social reform in Europe), but cover these parts in more detail. The structural indicators are used in progress reports (in the so-called Spring-reports[24]).

Thus, the indicators to use to transform the textual priorities into 'topics' can be taken from both the SDI- and the SI set, despite the former not being finalised yet. These indicator sets together provide an excellent basis for attaching an (set of) indicator(s) to each priority. Table 3.4 lists the main categories of both indicator sets (details can be found in Annex 4).

*Table 3.4: Main categories of SDI and SI indicators **

SDI – 1 Economic development
SI – 1 Employment
SI – 2 Innovation and Research
SI – 3 Economic Reform
SDI – 2 Poverty and social exclusion
SI – 4 Social Cohesion
SDI – 3 Ageing society
SDI – 4 Public health
SDI – 5 Climate change and energy
SDI – 6 Production and consumption patterns
SDI – 7 Management of natural resources
SDI – 8 Transport
SDI – 9 Good governance
SDI – 10 Global partnership

*) SDI refers to categories taken from the Sustainable Development Indicators, SI to those taken from the list of Structural Indicators. More details with respect to this list is given in Annex 4; SI – 5, 'environment', is not mentioned here specifically. This SI–5 category refers to climate change, energy consumption and transport and therefore provides no added value to the categories that belong to the SDI-list.

If we know how well a certain tool can cover the indicators listed in Annex 4 (of which the main categories are given in Table 3.4), we can derive from that the tool's capability to address each of the twelve priorities of sustainable development.

[22] Various versions of this list can be downloaded from http://forum.europa.eu.int/Public/irc/dsis/susdevind/library?l=/datastablessandscharts&vm=detailed&sb=Title
[23] See http://europa.eu.int/comm/eurostat/structuralindicators and http://forum.europa.eu.int/irc/dsis/structind/info/data/index.htm
[24] See http://europa.eu.int/comm/lisbon_strategy/reports/index_en.html

3.1.3 Operational aspects

Each tool has its own costs, data needs et cetera. These so-called operational aspects are important when determining which tool best to use for a specific assessment. Table 3.5 lists the operational aspects that shall be used for the evaluation of tools.

Some operational aspects (e.g. specificity of results, geographical coverage) are relevant only for arithmetic tools like models. At this point it remains questionable if these aspects need to be reported during the evaluation. After all, each and every model has its own geographical coverage, resolution etc. while it is not within the scope of *SustainabilityA-Test* to evaluate on a model-by-model basis[25]. Therefore, some operational aspects listed below might be too specific or detailed, and therefore unneeded.

Table 3.5: Operational aspects

Operational aspect	**Explanation**
INPUT / OUTPUT	
Costs (in monetary terms) for applying the tool (or a range of costs)	Estimated average costs in euros – costs include fees for using tools, hiring experts, labour costs
Manpower needs for applying the tool (or a range of manpower needs)	Estimated average number of man-months needed when using/applying the tool
Time needs for making the assessment (or a range of time needs)	Estimated average number of days involved in applying the tool (duration of its execution)
Data needs	Required data input, based on an expert judgement (provide an explanation of the judgement)
Data availability	Availability of the required data sets, based on an expert judgement (provide an explanation of the judgement)
Data type input	Description of the kind of input data that is required by the tool
Data type output	Brief explanation of what kind of output data is generated by the tool
Technical equipment required	Computation power, computer systems etc needed for using the tool, based on an expert judgement (provide an explanation of the judgement)
COMPLEXITY / TRANSPARENCY	
Complexity (of the tool itself)	Is the essence of what the tool does difficult to understand? Tools can be difficult to apply, while at the same time it is easy to understand what they do. Are the results easily to interpret? High, medium or low complexity, based on an expert judgement (provide an explanation of the judgement)
Transparency	Is the tool transparent? Are users able to understand what has been done to get to the outcome of the tool? High, medium or low user friendliness, based on an expert judgement (provide an explanation of the judgement)
User friendliness	Is the tool easy to work with? Can users apply the tool themselves? High, medium or low user friendliness, based on an expert judgement (provide an explanation of the judgement)
Reliability	Will the tool generate the same outcome when used more than once? High, medium or low reliability of the outcome, based on expert judgement (provide an explanation of the judgement)
Uncertainty	Is the tool capable of taking into account the uncertainty related to the outcome of the tool? Yes or no (provide an explanation)
TOOL CHARACTERISTICS	
Marginality	Is the tool capable of specifying the extent to which observed effects can be attributed to a policy proposal or an observed problem, i.e. can it help to assess whether the magnitude of effects are marginal, or significant, compared to the 'no (policy) change' or 'business as usual baseline'? Yes or no (provide an explanation)
Intensity	Can the tool specify what the extent is of an impact or an effect compared to the potential overall scope for change? Yes or no (provide an explanation)
Experience (with applying the tool)	Is there a lot of experience with applying the tool, or is the tool still rather new (or even in its development-phase)? High, medium or low level of experience, based on an expert judgement (provide an explanation of the judgement)
Mandatory usage	Legal obligation to use the tool (yes/no, and if yes, specify the obligation)

[25] *SustainabilityA-Test* is not designed to evaluate in detail all available models supporting sustainability assessments. IQ Tools (another 6th Framework Research Programme) will make such evaluation for models used by the European Commission.

Operational aspect	Explanation
Time before results become outdate	Estimated average number of years before tool results are out-of-date
Time scale / time horizon	Estimated number of years a tool can 'look back' (retrospective), and forward (prospective)
Geographical coverage	Number of (and specification which groups of) countries can be covered by the tool
Geographical resolution	Level of geographical detail (world, EU, country, region, NUTS)
Specificity of results	Level of aggregation or specification of results (level of specificity in terms of economic sectors, or socio-economic groups, etc)

3.2 Evaluation methodology

The evaluation of tools consists of two parts: 1) tool evaluation by means of the evaluation framework on the basis of existing knowledge and experiences, and 2) an elaboration of these desk-study results by means of drawing lessons from actually applying the tools to a real policy case during the case study (discussed in §3.3).

The (preliminary) tool evaluation is done by means of the evaluation framework and a common reporting format on the basis of existing information and experiences with respect to the tools. As indicated before, the tool-by-tool evaluation comprises an evaluation with respect to the ability of tools to support various policy processes and to address (crosscutting) aspects of sustainable development. In addition, a number of operational aspects need to be specified and other relevant issues to be reported.

3.2.1 How to evaluate which policy processes are supported?

The policy processes that are distinguished in *SustainabilityA-Test* are given in Table 3.1on page 32. The scoring used for how well a tool can support certain policy processes is given in Table 3.6.

Table 3.6: Overview of the scoring categories that can be applied in SustainabilityA-Test

Score	Value
0	Undetermined
1	Particularly well
2	Fairly well
3	Unsuitable
4	Unknown
5	Unknown – subject of further work beyond *SustainabilityA-Test*
6	Unknown – subject of further work within *SustainabilityA-Test*

Some tools can be regarded as processes themselves (i.e. not a one-off event, but a process that runs parallel to various stages in policy making, like e.g. *Strategic Impact Assessment*). In such cases, the score 'particularly well' is reserved for the policy process in which the tool normally 'starts'. The policy processes that run parallel with the tool application should all be given the score 'fairly well'.

When tools can be used in different policy processes, fulfilling different functions, each tool function must be separately evaluated: one tool evaluation sheet will be made for each function of the tool. There should be a different name assigned to each function of the tool, to prevent confusion.

3.2.2 How to evaluate the ability to address (crosscutting) aspects of sustainable development?

The (crosscutting) aspects of sustainable development that can be addressed by a tool are specified in Table 3.2 and Table 3.3 on page 34. How well aspects can be addressed by a tool is specified conform Table 3.6.

The tool teams have mainly scored the coverage of *main* categories of these aspects. It appeared so far to be not necessary to use the detailed lists of (crosscutting) aspects for the tool evaluation.

Only the coverage of tools of the (crosscutting) aspects of sustainable development has been evaluated by the tool teams. The coverage of aspects related to principles of sustainable development can be extracted from that evaluation.

3.2.3 How to evaluate the operational aspects?

The operational aspects, as specified in Table 3.5 on page 36, are reported in various quantities (e.g. costs in euros, manpower needs in man months) or scoring categories (e.g. user friendliness in high/medium/low categories). The following scoring system should be used.

Table 3.7: Scoring to be applied to the operational aspects

Operational aspect	**Score 1**	**Score 2**	**Score 3**
INPUT / OUTPUT			
Costs (in monetary terms) for applying the tool	Costs in euros*		
Manpower needs for applying the tool	Manpower needs in man-months*		
Time needs (for making the assessment)	Number of days*		
Data needs	Low	Medium	High
Data availability	High	Medium	Low
Technical equipment required	Low	Medium	High
COMPLEXITY / TRANSPARENCY			
Complexity (of the tool itself)	Low	Medium	High
Transparency	High	Medium	Low
User friendliness	High	Medium	Low
Reliability	High	Medium	Low
Uncertainty	Yes / No		
TOOL CHARACTERISTICS			
Experience (with applying the tool)	High	Medium	Low
Mandatory usage	Yes / No (and specify what obligation)		
Time before results become outdate	Years		
Time scale / time horizon	Retrospective: year (e.g. 1990) Prospective: year (e.g. 2010)		
Geographical coverage	Number of countries and group of countries (e.g. EU25, oldEU, newEU, World)		
Geographical resolution	NUTS1, NUTS2, NUTS3, EU-regions, countries, EU, World etc.		
Specificity of results	Sectors (agri, transport, energy, etc), socio-economic groups etc		

*) In first instance, costs, manpower needs and time needs will be specified in euros, man months and days respectively. This data can in a later stadium of the project be recalculated to high, medium and low categories, which allows for easy comparison across tools.

Determining whether a tool is for example *user-friendly* is not straightforward and depends, amongst others, on the user. It is the task of the tool teams to make an expert judgement, and to explain why a certain judgement has been made. The I&S team verifies objectivity in coordination with the tool teams. If necessary, the scores and arguments are discussed with all tool team leaders present.

3.2.4 Other relevant tool characteristics

The evaluation is not limited to tool characteristics enquired with the evaluation framework. After all, there might be relevant tool characteristics that do not necessarily come to the surface when only looking at the evaluation criteria as described above. Each tool team therefore also reports those tool characteristics not (yet) tackled with the evaluation framework.

3.2.5 Reporting format

The tool inventory and the results of the evaluation are reported by means of the tool overview and evaluation papers. In addition, a *tool information sheet* has been created for each tool, summarising the main findings and providing an overview of the evaluation scores given to each tool for each evaluation category. These sheets will be used as reference sheets in the handbook.

3.3 Case study

Brief summary of the case-study proposal developed by Karlheinz Knickel (IfLS), available at www.SustainabilityA-Test.net.

3.3.1 Purpose

The case study of *Sustainability Advanced Test* should help:

- to better understand the theoretical and conceptual basis of the various tools,
- to analyse the role played by tools in decision making processes,
- to compare the different tools in their ability to address key aspects of SD,
- to give insights on the interrelation between and best combinations of tools.

In other words, the case study should add to and deepen the preliminary tool evaluation of WP1 that was primarily based on a critical document review and review of experiences with the tools.

The case study's objectives can best be met when it allows for the tools[26] that have been evaluated during WP1 to be applied to the *same* policy case. It will then be possible to *directly compare* and identify:

- differences in the theoretical and conceptual basis,
- the practical implications of these differences in terms of assessment outcomes,
- the ability of the tools to address the different aspects of SD,
- the most suitable (effective, necessary) combinations of different tools, and, also very important,
- their operational characteristics[27].

[26] Or 'tools' that are representative for a particular tool type; depending on the most suitable level for the actual application and evaluation. The 'most suitable level' will presumably correspond with the one that has been used for the tool information sheets (TIS).

[27] In the TT evaluation papers as they are until now it can be seen that particularly these last two areas (combinations, operational aspects) remain rather vague.

All this is indicative of the particular strengths and weaknesses of the different tool(s) (types), and it improves (and illustrates) the information that will in the end be provided in the handbook.

3.3.2 Content

The topic for the case study

The policy case proposed is the *increasing governmental support* that is given to the expansion of energy crop production. It is expressed in terms of two specific policy decisions that are the starting point for the case study:

1. Biofuels Directive: Adoption of Directive 2003/30/EC of the European Parliament and of the Council of 8 May 2003 on the promotion of the use of biofuels or other renewable fuels for transport (OJ L 123, 17.5.2003, p. 42-46);
2. Energy Crop Premium: Introduction of an energy crop premium in the Common Agricultural Policy (CAP) (Reg. (EC) No 2237/2003 of 23 December 2003 (OJ L 339/52-69)[28].

These two policy decisions are related to each other because each of them will contribute to an expansion of energy crop production.

Why are both policy decisions important with respect to sustainable development and why this topic?

Due to the broad diversity of possible impacts, both cases form an interesting subject for the case study. The expansion of energy crop production that is being aimed at has environmental, social as well as economic impacts, and it has for example distributional and intergenerational effects. Many of them are expressed in the EU Sustainable Development Strategy (CEC, 2001):

- climate change goals: Increase the use of clean energy; reduce CO_2 emissions; energy crops are described as CO_2 neutral;
- decreasing the use of non-renewable resources: Energy crops substitute fossil fuels;
- manage natural resources more responsibly: Supporting a more sustainable land use; ensure that chemicals are used in such way that they pose no significant threat to the environment. A substantial increase in energy crop production may increase diversity of land use but may also lead to an increase in overall land use intensity (e.g. fertiliser and pesticides use);
- protect and restore habitats and natural systems and halt the loss of biodiversity by 2010: A substantial increase in energy crop production may increase the cultivation of and pressure on marginal high nature value land;
- economic and social goals: Reduce disparities in economic activities and maintain the viability of rural (…) communities. An increasing production of energy crops (and non-food crops in general) will provide new income sources to the agricultural sector and rural areas.

[28] Laying down detailed rules for the application of certain support schemes provided for in Title IV of Council Regulation (EC) No 1782/2003 establishing common rules for direct support schemes under the common agricultural policy and establishing certain support schemes for farmers.

At the same time the biofuel Directive is an outstanding example of a policy proposal where the chosen scope of assessments has a great effect on the outcome of assessments. Assessing the biofuel directive for its cost effectiveness as a GHG emission abatement measure, or assessing the directive for its impact on the security of energy supply in Europe, or assessing the directive for its implications on European land use, renders completely different outcomes.

Moreover, the case is not only significant in the EU as a whole, but also at national level. This will increase the likelihood that we will find sufficient policy processes within the case in which sustainability assessments (could) play a role. And finally, the case is favourable because the required state-of-the-art knowledge with respect to both agriculture and land-use is present within *SustainabilityA-Test*'s project consortium.

3.3.3 Design

The case study will be divided into two parts:

- In Part 1 the tool-by-tool evaluation as performed in phase 1 will be further deepened and elaborated on the basis of the case study. This part lasts until June 2005. The September 2005 workshop has been shifted forward to June, so that more time can be used to research (promising) tool combinations – the heart of *SustainabilityA-Test*. On the basis of part 1, the workshop in June provides an excellent opportunity to analyse efficient and promising tool combinations, thereby preparing for the second part of the case study.
- In Part 2 the emphasis will be on the analysis of (promising) combinations of tools. Most likely, members of different tool teams will have to join up and create across-the-tool team groups for efficiently researching combinations of tools (recipes) and the linkage of such recipes to assessment practices. This part lasts until the end of the case study (early 2006).

Part 1: deepening the tool-by-tool evaluation (February – June 2005)
Step 0 (preparatory work): Description of existing EU level assessments by IfLS (February 2005).
A description of the EU level assessment(s) that have been made in the course of the preparation and adoption of the Biofuels Directive has already been prepared by RIVM and will be provided to all tool teams as an input (no comparable assessment is available yet on the Energy Crop Premium). In this description information is given on the assessment questions that have been asked, the scope of the assessments that have been made at EU level (assessment questions; SD aspects [cross-cutting issues] and SD impacts addressed) and the tools have been used in the assessments. The description of the EU level assessment(s) will be given to the TTs for a critical examination (see Step 1).

Step 1: Critical review of EU level assessments by tool teams (March 2005)
On the basis of the description of the *EU level* assessment(s), the tool teams should ask how the assessment that has been carried out by or on behalf of the European Commission could have been improved. The tool teams should in particular critically examine the assessment from the point of view of what their tools have to offer.

Step 2: Review of national level assessments (March 2005)
In addition to the critical review of the EU level assessments, we will try to obtain an overview of the *national level* assessments that have been made in the context of the

preparation of the Biofuels Directive and the Energy Crop Premium. This will facilitate a more complete overview of actual assessment practice.

Step 3: to conclude the work of part 1 of the case study, each tool expert will describe how his/her tool could have been used for the biofuel directive or energy premium crop regulation or both, and what that tool would have contributed to the assessments that have been carried out already – specifically: what parameter aspect relevant for sustainable development could have been included. By means of an illustrative application of the tool, the tool expert shall than describe how the tool could have been applied (i.e. what input data would have been required, how much time would have been needed, an estimation of the expected costs and a description of the expected output). It is not the intention to actually carry out the assessment, as for most tools this would be too time and resource consuming.

Part 2: Combinations of tools and linking to assessment practice (June – December 2005)
The second part of the case study consists of the actual planning of a full-fledged assessment that satisfies the needs of a more comprehensive sustainability assessment. The underlying assumption is that the needs of a sustainability assessment could only partially be covered by individual tool teams in Step 3. Particular emphasis will be given on an effective coverage of relevant aspects, the creation of efficient linkages between tools and on the development of a convincing assessment procedure.

The development of suitable combinations of tools is the predominant goal. Tool teams should develop *links with other groups* wherever useful.

Presentation of the results of Step 4 (October - December 2005)
The plan for a more comprehensive sustainability assessment will be documented in a report that should follow common guidelines (approx. 30 p.).

3.3.4 Particular roles of tool teams during the case study
Some tool teams (TT5, TT6, TT7 and TT8) cannot apply their tools to the case study directly without using substantial inputs from other tools. They are more 'up-stream' tool groups. These TTs therefore play a different role in the case study. The different roles will be developed further during the Berlin workshop.

The 'new' roles during the implementation of Step 3 could be:
- TT5 Multi-Criteria Analysis: Examining coverage of relevant criteria by individual tool teams 1 – 4. Elaborating the question of a more integrated / summative evaluation with particular emphasis on MCA. TT5 will present the results of the observation in a short paper (4-5 pages; delivery: 17 June 2005).
- TT6 Environmental Appraisal: Playing the role of an observer and consultant. Thereby examining the consistency of the different approaches used by tool teams 1 to 4 with the declared aims and contents of the EU SDS and the related assessment procedures. TT6 is also asked to identify promising / interesting links between tool teams. TT6 will present the results of the observation in a short paper (4-5 pages; delivery: 17 June 2005).
- TT7 Stakeholder Analysis: Playing the role of an observer and consultant. Thereby examining the way the position of stakeholders is taken into account and the way participatory approaches are used by tool teams 1 – 4. TT7 would also provide advise

on possibilities for involving stakeholders (without actually implementing it in this phase). TT7 will present the results of the observation in a short paper (4-5 pages; delivery: 17 June 2005).

- TT8 Transition Management: Playing the role of an observer and consultant. Thereby examining the consistency of the different approaches used by tool teams 1 – 4 with the main ideas of the Transition Management concept. TT8 is also asked to identify promising / interesting lines of assessment thinking in respect of a transition to a more sustainable situation. TT8 will present the results of the observation in a short paper (4-5 pages; delivery: 17 June 2005).

TT5 and TT7 are expected to become actively involved in the assessment teams in the second part of the case study. TT6 (Environmental Appraisal) and TT8 (Transition Management) will continue playing the role of observers and consultants. TT6 examines the consistency of the approaches developed by the assessment teams with the declared aims and contents of the EU SDS and the related assessment procedures. TT8 is asked to identify promising / interesting lines of assessment thinking in respect of a transition to a more sustainable situation. TT6 and TT8 will present the results of their observation in two short papers (4 to 5 pages each).

3.4 Hand book

By Marjan van Herwijnen (based on the proposal for the handbook which can be found at www.SustainabilityA-Test.net).

The next sections will briefly discuss the options for one of the project's main outcomes: the Handbook *Advanced Tools for Sustainability Assessments*.

3.4.1 Handbook

The handbook is the main outcome of the project. It should enable those interested in finding all information generated during *SustainabilityA-Test*, and specifically those on the verge of carrying out an assessment in finding the most suitable tools to do so.

The handbook will be based on the following pieces of information and products developed in the course of the project:

- the tool overview papers – these form the core of the handbook and can be used as a book of reference in which all details regarding the tools can be found;
- the tool information sheet – these sheets summarise the main characteristics of each tool and can therefore be a useful introduction into each tool;
- the evaluation framework – the evaluation criteria will form the criteria for looking up the tools in the handbook.

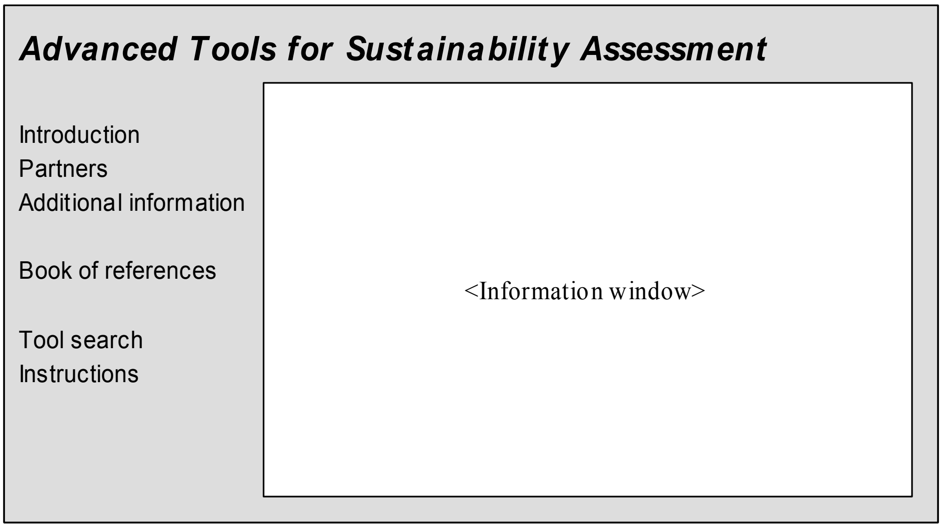
Case study output is not mentioned separately, as the information generated during the case study will feed into each of the products mentioned above.

The bundle of all tool overview papers already contains over 400 pages of text, excluding the tool information sheets and output of the case study. In order for the handbook to become manageable and useful, an electronic version of it will be made. This allows easier distribution and better access to the information that is included. When such electronic version is stored on a website (we refer to that as a web book), updating the information becomes a possibility, thereby extending the lifetime of *SustainabilityA-Test*'s output.

3.4.2 Web book

The first ideas with respect to the web book have been put on paper (see for details www.SustainabilityA-Test.net). The most important items are briefly discussed below.

The home page of the web book will have the following structure:



Introduction

If one clicks on *Introduction*, the following options will appear below this link on the left:

Aim	Aim of the website
Content	Introduction to the website's content
Context	Introduction to the project *SustainabilityA-Test*
Terminology	Introduction into tools, procedural tools and recipes

If one clicks on one of these options, information will appear in the information window at the right.

Partners

If one clicks on *Partners*, the list of partners and a link to their websites will appear in the information window.

Additional information

If one clicks on *Additional information*, the following options will appear below this link on the left:

Evaluation criteria	Description of the evaluation criteria
Tool information sheets	Explaining the structure and content of the tool information sheets
Tool evaluation papers	Explaining the structure and content of the tool evaluation papers
Tool search program	Explanation on how the search algorithm works
Case study	Information on the case study used in the evaluation
Other deliverables	List of other relevant deliverables from *SustainabilityA-Test* and links to there location

If one clicks on one of these options, information will appear in the information window.

Book of references
The book of references allows users to navigate to the tool information and evaluation results in two ways:

1. by using the evaluation criteria (i.e. policy processes, SD aspects, coverage of impacts or operational aspects) as entry points. The user will be presented with e.g. table showing all policy processes, all tools and how well each tool can support each policy process;
2. by using the tool groups as entry points. The user can navigate to the tools through these groups.

Instructions
If one clicks on *Instructions*, the following options will appear below this link on the left:

Introduction	General introduction to the search option
How to …	Detailed instruction on how to execute a search

If one clicks on one of these options, information will appear in the information window.

3.4.3 Tool search program

In addition to navigating to the tool information on the basis of the evaluation criteria and the tool groups, the aim is to also include some search program that helps users in funding suitable tools for carrying out assessments.

The search criteria for a tool search program could be largely based on the evaluation criteria. A shell should be designed around that in order to ensure a user friendly program that actually helps in finding tools and promising tool combinations, and that proposes realistic solutions for carrying out assessments.

4 Tool overview and preliminary evaluation

This chapter presents the tools that will be evaluated within *SustainabilityA-Test*. First, the tool groups and tool types will be explained, followed by an overview of all tools considered within the project.

A detailed tool-by-tool description and discussion of the evaluation results can be found from section 4.2 onwards.

4.1 Tool groups and tool types

Eight groups of tools have been distinguished, although it is difficult to apply a strict classification to the selected tools. Some tools overlap, and the applied classification is also partly shaped by the concentration of expertise with certain tool types within the project's consortium, solely for the purpose of efficiency. The classification presented here forms a useful starting-point, but should remain open for discussion until the end of the project.

The following tool groups are distinguished within the project:

1. physical assessment tools – tools that assess some physical parameter;
2. monetary assessment tools – tools that assess some financial/economical parameters;
3. models – tools that used (computer) model;
4. scenario analysis – tools with a prospective character;
5. multi-criteria analysis – tools that help with the consideration of various criteria;
6. sustainability appraisal tools – tools prescribing how sustainability appraisals could/should be done;
7. stakeholder analysis tools – tools that aim to involve stakeholders;
8. transition management – tools that can support transition management.

In *SustainabilityA-Test* the word 'tool' is used to refer to all types of tools, methods, methodologies and procedures that can be used to carry out (part of) an assessment. It is a collective term used only as such to prevent us from having to write 'tools, methods, methodologies and procedures' each time we want to refer to all kinds of tools. Obviously, different types of tools exist. Some have a more procedural character whereas others have a more instrumental character. Different types of tools have different functions in assessments, and in order to be able to distinguish between them, we distinguish the following tool types within *SustainabilityA-Test*: methods, tools, procedures and recipes[29].

Note that the word 'tool' is used both as a collective term, referring to methods, procedures and tools, and as a specific 'type of tool'. This might lead to confusion. However, we decided to keep this twofold usage of the term tool, as both are already well rooted in the project, and confusion will be avoided by the context in which both terms are used.

Each tool type (method, tool, procedure and recipe) is discussed in more detail in the next sections.

[29] Terminology proposed in the draft methodology report, elaborated by John Robinson (SDRI, Canada) and discussed and agreed at the inception workshop 9 and 10 September 2004.

4.1.1 Method

A method is a tool type on the most detailed level within *SustainabilityA-Test*. It is a specific analytical procedure within a tool; a specific 'way of doing something'. Most tools contain multiple methods. Methods themselves can have many levels, and methods can be embedded in methods. E.g. a method is *non-market valuation* (methodology to estimate the benefits of a policy intervention) within which *contingent valuation* is a method too (non-market valuation methodology to estimate the benefits of a policy intervention).

4.1.2 Tool

A tool is a tool type which makes up a recognisable methodological approach (e.g. CBA, scenario analysis, et cetera). Tools can use various methods. An example of a tool is the cost-benefit analysis. Different methods exist to estimate the costs and the benefits.

The word 'tool' is also used as a collective term, referring to all tools that will be evaluated within *SustainabilityA-Test*. This collective term is used to prevent us from having to write 'methods, tools, procedures and/or recipes' each time we want to refer to all 'tools'.

4.1.3 Procedure

A procedure is a tool type, which describes how tools can be used to accomplish a type of assessment. Procedures can use various tools and/or methods, but they do not consist of certain tools and/or methods per se.

Examples of procedures are the Commission's Impact Assessment procedures (CEC, 2002b) and the Strategic Environmental Assessment procedure, describing what elements should be included in the assessment (which tools or methods to use, is left open).

4.1.4 Recipe

A recipe is a combination of procedures, tools and methods to undertake (parts of) sustainability assessments. A recipe could be one procedure with pre-described tools and methods, but it could also be a combination of different procedures, tools and methods.

Tools are seldom used alone and for the most part used in combination with other tools, with or without procedures describing which tools to use and how to use them. The heart of *SustainabilityA-Test* is to analyse common assessment practice and to identify leads for making combinations of tools that could strengthen assessments. Combinations that will be evaluated within *SustainabilityA-Test* shall be referred to as 'recipes'. This word has been chosen to capture the fact that recipes refer to more than just making combinations of tools. It is a combination of tools *and* an instruction of how to create and use it.

The word 'recipe' will be used within the project team to refer to these combinations. It remains to be seen whether this word can also be used outside the project team – e.g. in the final handbook – as it is currently unknown if such word provides added value or leads to confusion.

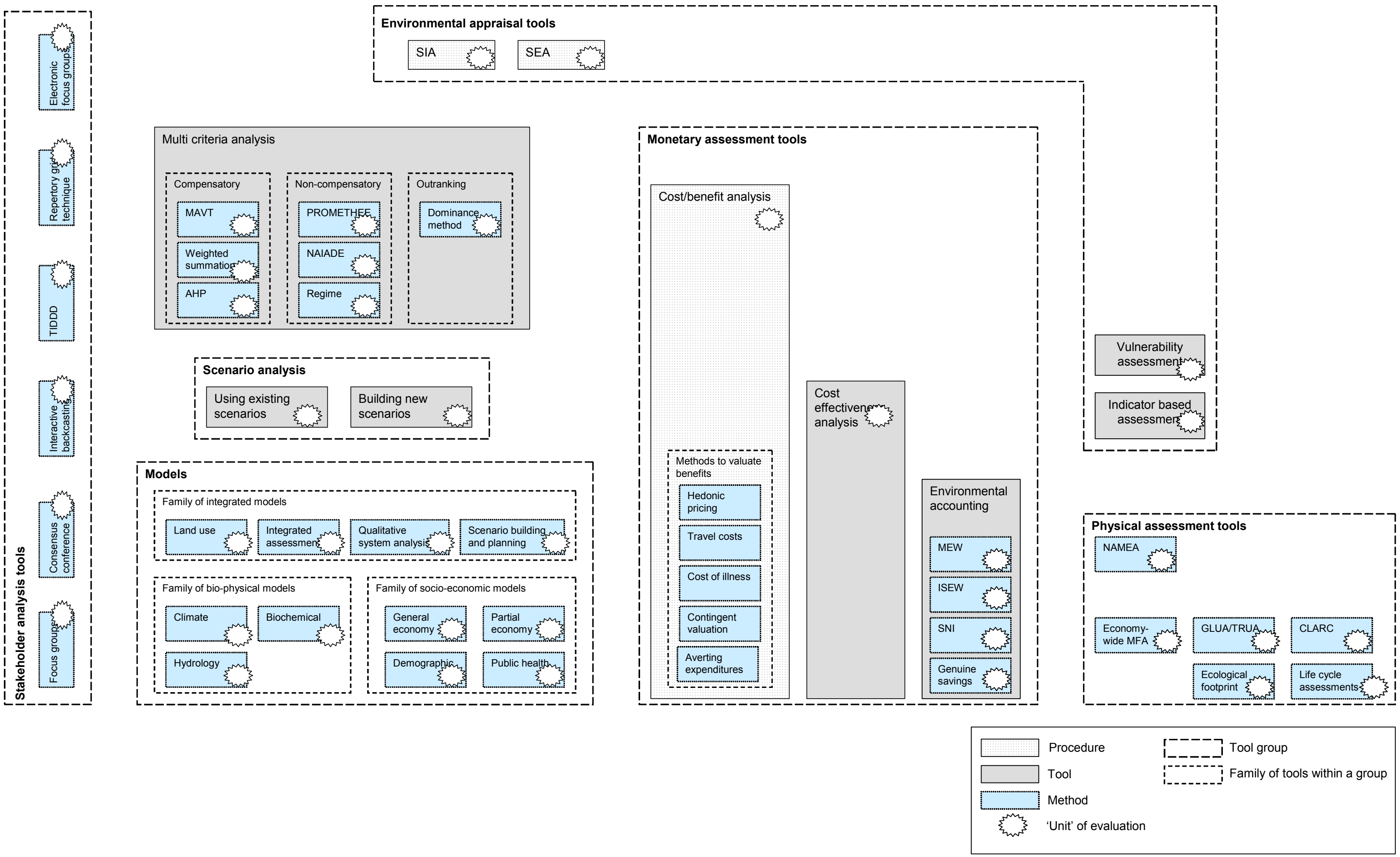


Figure 4.1: Overview of tools considered in SustainabilityA-Test

Note: the (vertical) 'hierarchy' suggested by this figure is roughly based on what tools could be considered a building block for other tools. Obviously, there exists no one true hierarchy. The figure provides an overview and the hierarchy should therefore not be attached too much value.

4.2 Overview of tools

All tools that are commonly used in Europe for carrying out assessments initiated by policymakers will be considered in *SustainabilityA-Test*. Not all of these will be evaluated, though. *SustainabilityA-Test* shall focus on the most common and/or promising tools. ***Figure 4.1*** (page 49) shows which tools have – so far – been included in the project. The little stars in the figure denote the 'units' that will be evaluated. A description of each tool category and the tool/methods belonging to each category can be found in chapter 4.

Sections 4.2 to 4.9 present detailed descriptions of the tools included in each tool group. These sections also contain the results from the preliminary evaluation. The information is taken from the tool overview and evaluation papers (deliverable D5), which are available at www.SustainabilityA-Test.net.

The focus of the preliminary evaluation lies on the strengths and weaknesses of each tool, as this provides ground for further exploring the tools during the case study. Particular attention is given to weaknesses of each tool, thereby possibly giving a too negative impression. The reason for choosing this approach is that the weaknesses of each tool provide leads for further improvements. A more in depth and comprehensive presentation of the evaluation results, including a more complete overview of each tool's strengths, can be found in the tool overview and evaluation papers.

4.3 Physical assessment tools

By: Philipp Schepelmann (WI), Stefan Bringezu (WI), Matthias Nerger (WI), Stephan Moll (WI), Helmut Schütz (WI), Karl-Heinz Simon (CESR), Anne van der Veen (UT) and P.J. Stauvermann (UT)

All tools within the group of physical assessments tools relate – in one way or another – human activities to environmental pressure. For the assessment of impacts it is decisive how the pressure on the environment is put into a cause-effect (pressure-impact) relation.

Bringezu et al. (2003) distinguish between two types of pressure:

1. specific pressure: the assessment of impacts is based on a sufficient knowledge of physiochemical and toxicological properties of specific substances for which quantitative measures are available (e.g. ozone-depleting or greenhouse potential).
2. generic pressure: the assessment of impacts is based on a system-specific generic environmental impact potential associated with the amount of human consumption of natural resources.

Measuring specific pressures have the advantage of being more accurate. The drawback is that it is selective, because there are only few impacts on the environment for which sufficient quantitative knowledge is established (and generally accepted).
Measuring generic pressure has the advantage of giving a systemic overview and providing information on the metabolic performance of a region or economy. The drawback is that impacts are only estimated to indicate the order of magnitude of impact potential (unspecific).

4.3.1 Economy-wide MFA

Economy-wide MFA is used to analyse the physical relation between the economy and the environment in order to understand how it works (Bringezu et al. 2003). The results are the basis for the evaluation and possible structural change. Economy-wide MFA provides an overview of annual material inputs and outputs of an economy including inputs from the national and foreign environment and outputs to the environment and the physical amounts of imports and exports (Eurostat, 2001). The accounts are in physical units, usually tonnes per year. Economy-wide MFA can be used as a physical complement to the monetary System of National Accounts.

Economy-wide MFA accounts for material exchange between national or regional economies and:

a) the environment (via resource extraction on the input side and waste deposition, and releases to air and water (on the output side), and
b) other economies (via trade) through measuring flows in physical units.

Economy-wide MFA constitute the basis from which a variety of material flow based indicators can be derived to measure the metabolic performance of economies and regions with regard to material input and output, used and unused extraction of resources, domestic and foreign resource requirements, renewable and non-renewable resource input, balanced vs. unbalanced physical trade (with and without hidden flows).

The principles of statistical approaches towards material flow accounts and material balances have been formulated in the 1970s (see e.g. United Nations, 1976). It is also used by the European Commission (Moll et al., 2003), the EEA (EEA, 2000) and Eurostat (Eurostat, 2001). Material flow accounts also are part of official statistics in several EU Member States and EFTA countries (see Eurostat, 1997). The idea of economy-wide aggregated material flow accounts and balances (as opposed to single-material or substance accounts) has been applied already in the late 60s (Ayres and Kneese, 1969), and was re-vitalised in the early 90s and put into statistical practice in e.g. Austria (Steurer, 1992, Fischer-Kowalski and Haberl, 1993), Germany (Schütz and Bringezu, 1993), Japan (Japanese Environmental Agency, 1992) and the USA (Rogich et al., 1992, Wernik et al., 1996). The application of MFA is also recommended by the OECD Council[30].

Evaluation: This tool provides an overview of the annual material inputs and outputs of an economy. It is therefore a measure of the metabolic performance of an economy, linking economic activities with environmental impacts through generic pressure indicators. The tool is not capable of measuring substance-specific impacts, although the relation of generic pressures (e.g. total outputs to the environment) specific pressures (e.g. carbon dioxide emissions and waste deposition) can be determined. The economy-wide MFA indicates pressures at an aggregated level, i.e. for a whole nation, or by sectoral attribution for branches and product groups (linked to NAMEA), on a micro level (households, firms or products and services) some of the MFA-based indicators can also be applied in a consistent manger (link to LCA-type of analysis (Ritthoff et al., 2002).

[30] Recommendation of the Council on Material Flows and resource productivity. Endorsed by Environment Ministers on 20 April 2004 Adopted by the OECD Council on 21 April 2004

4.3.2 LCA

Life cycle assessment (LCA) is a technique generally used for assessing specific environmental impacts associated with a product. As a systematic tool, LCA analyses and assesses environmental impacts over the entire life cycle of a product. The main areas of the application of LCA within public environmental politics have been waste treatment options, means of transport, energy sources, and product's choice (Frankl and Rubik, 2000). Authorities may use LCA, for instance, in work on product related aspects of environmental action plans and in environmental labelling of products (Wenzel et al., 1997). Companies may use them in product development, environmental management and marketing. Consumer organisations may use them in counselling consumers.

Full LCA methodologies are codified in the ISO standard series 14040 (ISO, 1996). Carrying out a LCA contains four main phases (ISO 1996, SETAC, 1993, Wenzel et al., 1997, Frankl and Rubik, 2000):

- Phase 1: defining the goal and scope
- Phase 2: inventory analysis to define the system under investigation and its boundaries by means of describing the input and output flows of materials, energy, water and pollutants, necessary data, calculation procedures, allocation rules etc.
- Phase 3: carrying out the impact assessment (classification and characterisation of environmental impacts based on the inventory analysis, regarding goal and scope)
- Phase 4: interpreting the results

Evaluation: LCA can systematically address established impacts of a product from raw material acquisition to final disposal. Decisions based on LCA therefore stimulate minimisation of the use of materials and energy for existing processes, which in turn minimises or avoids effluents, air emissions and (hazardous) wastes of production systems. On the downside, LCA requires vast amounts of data, of which the availability and quality is often a big problem. In addition, any LCA necessarily involves assumptions and subjective valuation procedures, which often lack transparency, and lead to results which are usually very specific for the comparison of certain products and may not be transferred to other systems. LCA can furthermore cover only a limited subset of environmental impacts. It can be considered a rather complex tool, which is difficult to communicate and which requires high expertise and much time to use.

4.3.3 EF

The main question answered by the Ecological Footprint (EF) is how much biologically productive land would be required on a continuous basis to provide for the necessary energy and material resources consumed by a population and to absorb the wastes discharged by that population. It does not compute actual land use; the footprint is a hypothetical figure. But, it enables to compare regions and countries. Wackernagel et al. (1999) calculate that per global citizen 2 ha is available. This figure thus becomes the ecological benchmark for comparing people's ecological footprints.

EF expresses results in spatial units. It applies optimistic yield figures, not including all impacts. Several land categories are distinguished: consumed/degraded land, gardens, crop land, pasture land and grass lands, productive forests, and energy land. Energy takes a special place in footprint analyses: More than 50% of EF estimates relate to land needed to catch CO_2 emissions from burning fossil fuels. in biomass. It is thus a combination of a systemic overview on the physical turnover of an economy with an association of specific impacts on the environment (land).

The EF is used as an indicator for the (un)sustainability of a person, city, region or nation, as it is an appealing and easy to communicate concept (see e.g. www.myfootprint.org). With regard to operationalisation at various levels, however, the applicability and use seems limited, especially when priority fields of action are to be determined. Although we are rather sceptical on the use of Ecological Footprint in policy design, policy analysis and policy evaluation, we are of the opinion that EF can be applied in a didactically context in order to improve understanding of the scale and impact of environmental problems.

Evaluation: The ecological footprint provides an easily communicable indication of the pressure a certain region poses on the world with a focus on green house gas emissions. This is why it is a rather popular indicator. In the calculation of the EF a variety of assumptions have to be made, some of which seem contestable. As an assessment tool, it seems questionable whether the EF provides more or better information than the basic parameters (e.g. GHG emissions) used for recalculation into area units.

4.3.4 GLUA/TRUA

The first steps towards GLUA were taken in the Eurostat guide (Eurostat 2001), proposing the assembling of trade balances for land use (in hectares) as an additional resource use indicator in international trade. The land based trade balance of a country quantifies the land area used for the production of its imports abroad as well as the domestic and foreign land area needed to produce the goods and services exported to the rest of the world. In order to provide a more comprehensive picture it seems to be useful to quantify the total mass flow-related land use. This approach of a ‘global land use accounting’ (GLUA) considers total resource flows and land use associated with the domestic activities of a national economy, region or product chain. For practical reasons it may focus on the global land use associated with the domestic consumption of agricultural products. The method measures resource use on a life-cycle-wide basis, i.e. also trans-regional resource requirements. Thus, it allows to detect the shift of environmental burden between regions, e.g. in the course of globalisation, and as a consequence of technological change, e.g. with a shift towards biofuels and biomaterials.

GLUA accounts for the land use related with agricultural commodities (Schütz et al., 2003). Combining the yields of agricultural primary commodities with the physical imports and exports data provides information that accounts for land use related with a particular agricultural commodity (e.g. coffee, fruits, vegetables, fibres). The results show to which extent the examined economy requires more or less land for its consumption of agricultural products than is available within the region. Further, the absolute level of land use per capita of the examined economy demonstrates how much this requirement is above or below the global average of available agricultural land.

Evaluation: GLUA can be used to calculate how much land a certain region uses for the production and consumption of agricultural commodities, inside and outside country or region under investigation. It can therefore be used to calculate the shift of environmental burden from one region to another. The assessment of such shifted impacts is derived from quantified generic pressure. When GLUA is combined with (economy-wide) MFA, total resources use accounting emerges (TRUA), which could be used to calculate e.g. land use associated with biofuel production and consumption, both inside and outside that nation. TRUA therefore offers the opportunity to reveal possible impacts of a transition towards biofuels including inequalities in distributions of land use (also per capita). Hence, the results contribute to evaluation in terms of sustainability.

4.3.5 NAMEA

The NAMEA (National Accounting Matrix including Environmental Accounts) is a statistical information system to combine national accounts and environmental accounts in a single matrix. It is a hybrid accounting system. The NAMEA system contains no economic assumptions; it is only descriptive. It maintains a strict borderline between the economic and the environmental aspects. It is represented in monetary units on the one hand and in physical units on the other hand. The interrelationship between the economy and the environment has two perspectives, an economic one and an environmental one. The economic perspective contains the physical requirements in the economic processes, like energy and material and spatial requirements. The environmental perspective puts forward the consequences of these requirements with respect to the availability of the natural environment. Consequently, the optimal allocation of natural resources requires the consideration of both perspectives.

The fundamental idea of the NAMEA is to extend the conventional national accounting matrix with two additional accounts. One additional account is the account for environmental problems like the greenhouse effect or the ozone layer depletion. The selected environmental themes are partly global environmental problems and partly national and local environmental problems. The second additional account is for environmental substances, like carbon dioxide or sulphur dioxide, where these substances are expressed in physical quantities, like kilograms, tons et cetera. NAMEA generates consistent summary indicators for those environmental problems, which are considered most pressing at the political level.

The NAMEA distinguishes between households and industries including public services. Further, the NAMEA consists of two types of physical accounts: the substances accounts and the environmental themes accounts. Besides the conventional economic aggregates, the NAMEA contains a summary of environmental indicators. As a result it can be investigated how much a specific economic activity contributes to the GDP, employment, exports et cetera and how much it contributes to the major environmental problems, like the greenhouse effect, ozone layer depletion et cetera. Given the NAMEA it is possible to decompose the changes in emissions by industry into several effects:

1. demand composition shift effects;
2. output growth effects;
3. eco-efficiency change effects.

<u>Evaluation</u>: NAMEA is a tool to relate environmental themes with economic structure and performance. It is possible to integrate social accounts into the NAMEA system. This is done in the so-called System of Economic and Social Accounting Matrices Extensions (SESAME)[31]. Consequently, it is possible to get insight into the problem of who should pay for environmental damages[32]. Depending on political preferences sustainability can be translated into environmental indicators combined with the economic indicators, capturing

[31] See e.g. Keuning (1997), Keuning (1998), Van de Ven, Kazemier & Keuning (1999), Keuning & de Haan (1996).
[32] See Steenge (1997, 1999).

part of sustainability only. NAMEA can be used in all stages of the policy processes by informing the public and policy makers about the status quo of the environmental assets and environmental pollution. Especially, NAMEA provides policy makers with a data-framework, which can be used to sketch the trade-offs between costs of prevention of environmental damages and macro-economic policy objectives.

4.3.6 CLARC

Societal dynamics which are of relevance for energy and material fluxes can be found at different levels and in many different fields: land use, industrial transformation, mega-cities development, etc. The tool 'Characterising lifestyles and their resource consumption' uses the household level and the consumer's decisions as the focus of analysis. It is possible to characterise different types of households in different states with divergent development conditions (industrial states, least developed countries, et cetera). A look at the resource consumption patterns will give clues about the society analysed: the consumption patterns of today give much information about the past and the present state of society, and the actual and ongoing dynamics in societies will determine the consumption patterns of the future. However, in most cases it is not of much help to include the typical lifestyle of a certain society only – one has to deal with different fractions (societal groups, classes) in a society because the differences in social aspects usually determine also differences in resource utilisation and environmental impacts.

The tool is applicable with different grades of resolution. Hence, data requirements are depending on the chosen approach. Approaches can be imagined that work without any concrete empirical data. The use of cultural types, for example, establishes a conceptual framework only and within that framework scenarios are worked out (using general hypotheses about demand structures and technological options) that emphasise different patterns of demands of energy or goods and products. Other approaches are using costly data surveys or use statistical data, e.g. in correlating energy consumption and income or social status.

Evaluation: CLARC completes the analysis of the energy and material fluxes and their evaluation by defining a starting point (demands, household consumption patterns based on lifestyles). It has the potential to be an eye-opener for future problem situations. When qualitative approaches are chosen, the data requirement is costly.

4.4 Monetary assessment tools

By: Anna Alberini (University of Maryland and FEEM), Onno Kuik (IVM , Vrije Universitaat Amsterdam), Harmen Verbruggen (IVM, Vrije Universitaat Amsterdam), Benjamin Goerlach (Ecologic), Anne van der Veen (University of Twente) and Julia Bartos (Czech Environment Center).

The tool category monetary assessment tools focuses on the theory and practice of the following tools:

- cost-benefit analysis,
- cost-effectiveness analysis, and
- environmental accounting.

Evaluation: Each monetary assessment tool builds on the monetisation of impacts. By expressing impacts in monetary terms multiple impacts are made comparable with one another, in an easy to understand and appealing fashion. Monetisation of impacts is therefore generally used throughout the world.

A general weakness of monetisation is obvious: not everything can be straightforwardly monetised, and even if something can be monetised, assumptions have to be made which are sometimes severely challenged. A risk associated with the usage of all monetary assessment tools is therefore that the selection process of impacts that should be accounted for becomes dictated by the possibility to monetise them, in which case the outcome of the assessment does not take into account all relevant aspects. In addition, the distribution of costs and benefits, and issues like equity and fairness cannot be addressed, as the monetary assessment tools usually work at a higher level of aggregation (e.g. x million euro costs to society, without specifying the 'winners' and the 'losers'.

4.4.1 Tool: Cost-Benefit Analysis

Cost-Benefit Analyse (CBA) monetises the positive (benefits) and negative (costs) impacts of a proposed policy or project, and compares these sets of figures with one another. CBA can be used to estimate costs and benefits both ex-ante and ex-post, i.e. before and after a proposed policy is being implemented.

The first step in a CBA is determining which costs and benefits to include, along with the parties who incur such costs or experience such benefits. Next, the analyst needs to place a monetary value on the various categories of costs and benefits, and aggregate the various categories of monetised costs and benefits into cost and benefit totals. This task is complicated by the fact that many policies entails costs and benefits that are incurred in different time periods – which requires discounting them to compute their present values – and that many categories of benefits are non-market goods, and as such are not bought and sold in regular marketplaces.

In the next sections categories of costs and benefits, and methods to estimate categories of benefits are further discussed, followed by a discussion of some important aspects related to the tool CBA and relevant for cost/benefit calculations.

Categories of costs

When economists estimate the costs of a policy or project for the purpose of conducting a CBA, they refer to the social costs of the policy, which may well be different from the private costs (i.e. the costs to a private entity) of the policy. The costs of a policy can be placed into five broad categories:

1. real-resource costs, including compliance costs: all of the resources that are used up to implement the programme or policy. Real-resource *compliance* costs are all of the resources that must be expended for complying with the regulatory aspects of the programme;
2. government regulatory costs: the monitoring, administrative, and enforcement costs associated with the policy, especially when the latter has a regulatory aspect;
3. social welfare losses: the consumer and producer losses associated with possible rises in prices or decreases in output that occur as a result of the policy;
4. transitional costs: the value of all the resources that are displaced by the policy, and the private costs of reallocating these resources;

5. indirect costs: costs such as adverse effects that the policy may have on product quality, factor productivity, innovation, discouraged investment, and changes in markets indirectly affected by the policy.

Categories of benefits

Which specific non-market valuation technique is appropriate for one's benefit estimation exercise depends on the context, on the category of benefits being considered, and on whether the market or individuals' behaviours are consistent with the assumptions of the method being used. The estimation of the benefits of a policy is generally a difficult task, much more so than the estimation of the costs of a policy. The benefits of projects or policies will often take the form of improvements in individuals' welfare that are not traded in markets, and for which there is no market price.

The first order of business for the analyst who wishes to estimate the benefits of this policy is to identify the beneficiaries of the policy. The next step involves quantifying the physical effects of the policy. Once the physical effects are established, a monetary value must be attached to them. ***Table 4.1*** lists examples of categories of benefits and methods top estimate their monetary value.

Table 4.1: Examples of categories of benefits of environmental policies

Benefit category	**Example of service flows affected by the policy**	**Possible monetary valuation methods**
Human health benefits: morbidity and mortality risks	Reduced risk of cancer, reduced risk of respiratory symptoms	– averting behaviour – contingent valuation – hedonic pricing methods – cost of illness
Amenities	Visibility affected by air quality	– averting behaviour – contingent valuation – hedonic pricing methods
Ecological benefits: market products	Provision of food, fuel, timber, fibre, fur	– Market approaches
Ecological benefits: recreation and aesthetics	Viewing, fishing, boating, swimming, hiking, etc.	– production function – contingent valuation – hedonic pricing methods – travel cost method
Ecological benefits: ecosystem services	Flood moderation, climate moderation, water filtration, sediment trapping, groundwater recharge, soil fertilisation, pest control	– production function – averting behaviours – hedonic pricing methods
Ecological benefits: existence and bequest values	No associated services (passive use values)	– Contingent valuation
Materials damage	–	– Averting behaviour – market approaches

In theory, the benefits of a policy are correctly captured by the beneficiaries' willingness to pay (WTP) for the policy. The total benefits of the policy are the sum of the individual beneficiaries' WTP. There are several ways to estimate the WTP:

- Using information available in regular markets (market approach): in some cases, it is possible to determine the WTP for the proposed policy by using information available in regular markets. For example, if an environmental or agricultural policy results in increased agricultural output, the benefits of the policy should be equal to the change in crop, multiplied by the market price of the crop. When valuing ecological benefits of a proposed policy, it might be possible to identify service flows that can be bought and sold competitively as factors of production or final consumption goods;

- Contingent valuation (non-market, stated preference approach): a survey-based approach that asks individuals to report their WTP for a public programme that would deliver a public good or a specified improvement in environmental quality;
- Travel costs method (non-market, revealed preference approach): model-approach for estimating the recreational benefits of a certain resource;
- Hedonic property value (non-market, revealed preference approach): asserts that individuals perceive housing units as bundles of attributes and derive different levels of utility from different combinations of these attributes;
- Compensating wage study (non-market, revealed preference approach): using the hedonic property value methodology on labour markets;
- Benefit transfer (non-market, revealed preference approach): using existing information from previous studies and applying that to a new context.

Methods to estimate benefits

The choice between the various methods available depend on a number of factors, including the type of benefit being valued, data availability, whether housing or labour markets satisfy the assumptions needed for hedonics, the need to focus on certain categories of beneficiaries and benefits, and the amount of funding available to perform the analysis. These considerations are made explicit in ***Table 4.2***.

Table 4.2: Non-market valuation methods

Method	Suitable for…	Type of values	Conditions
Stated-preference approaches			
Contingent valuation	Virtually any public policy or programme; extremely flexible	Use values, Non-use values	Design and administration of the questionnaire are difficult, a number of biases are possible that can limited through careful construction and pretesting of the survey instrument.
Revealed-preference approaches			
Travel cost methods	Only for amenities, natural resources (e.g., beaches, bodies of water, national parks or wildlife reserves) or cultural sites (monuments) that people actively visit.	Use values	Travel cost can be subject to measurement error, especially if the researcher wishes to include the opportunity cost of time. It may be difficult to identify substitute sites. Questions about trips taken under hypothetical conditions may be necessary to trace out the demand function at post-policy conditions.
Hedonic pricing methods	Only for changes in environmental or urban quality that can be captured into housing markets; only for job risks that are captured into compensating wage differentials.	In theory, both use and non-use	Individuals are assumed to be perfectly aware of the environmental, urban quality, job risks. Market must clear. Sufficient transactions must be observed to estimate the hedonic regression, and sufficient variability in environmental or urban quality or job risks must exist to identify their effect. Difficult to separate the effect of these variables from other factors that can influence housing prices or wages.
Averting expenditures	Human health effects or other effects (e.g., materials damage) from which people can protect themselves	n/a	Possible when individuals can document actions and expenditures incurred to reduce risks. In some cases, it is possible to engage in actions that reduce risks (e.g., staying indoors in days with high air pollution) but it is not easy to place a monetary value on these actions. Fails to capture the value of the discomfort of being sick.
Cost of illness	Human health effects	n/a	Relatively easy to perform, but fails to capture the value of the discomfort of being sick.

Conducting a valuation study is generally expensive and time-consuming. In some cases, it is possible to answer policy questions without necessarily conducting an original valuation study. This is called benefit transfer. The benefit transfer approach relies on existing information from previous study, and seeks to apply it to a new context or locale.

4.4.2 Tool: Cost-Effectiveness Analysis

A Cost-Effectiveness Analyses (CEA) focus primarily on the costs of attaining certain policy goals, usually measured in physical units (e.g. acres of wetlands restored, number of cases of illness avoided). A CEA seeks to find the best alternative activity, process, or intervention that minimises resource use to achieve a desired result. Analysts and agencies perform CEAs when the objectives of the public policy have been identified and the only remaining question is to find the least-cost option of arriving at these objectives. CEA, therefore, does not ask, nor attempts to answer, the question whether the policy is justified, in the sense that its social benefits exceed its costs: cost benefit analysis can be used for that.

Like cost-benefit analysis, cost-effectiveness analysis can be used either to assess the expected impacts of alternative policy measures before they are implemented (ex-ante), or to assess the effectiveness of a policy measure that is already in place (ex-post). While the approach and the methods used are the same for ex-ante and ex-post CEAs, the purpose of the instrument is different. Ex-ante CEA is used up front, at an early stage in the policymaking process, to identify the path of action that promises to be most cost-effective. An ex-ante assessment will need to rely on assumptions and projections, as well as on cost and effectiveness data from different contexts, in order to anticipate the impacts of future measures. By contrast, the ex-post CEA aims to assess whether a problem has been tackled effectively through the policy measure or project investigated. In other words, it provides a measure for the efficiency of policy implementation.

A CEA normally consists of four steps. The first is to determine the objectives of a policy and programme, and to quantify these in physical terms. In the second step, the total costs of the policy or programme are assessed. The main focus is on direct monetary costs, using market prices. In addition, the assessment may also include wider economic impacts or opportunity costs. Future costs are discounted to give the net present value. The third step is to quantify the impact that the policy or programme on the predefined target. In the case of an ex-ante analysis, this may involve scenario building or modelling. For an ex-post analysis, the level of target achievement is easily observed, but the impact of a particular policy measure may be difficult to disentangle. The fourth step is then to assess the cost per unit of output for different strategies through a simple division.

<u>Evaluation</u>: Both CBA and CEA use discount rates to recalculate costs and benefits in the future to present value. This discount rate strongly influences on the outcome of assessments and is often disagreed upon. With a CEA it is not possible to determine whether society is better of by implementation of a certain policy (CBA should be used for that), or by reaching a certain target. It should be used only to calculate which policy option is most effective in implementing a certain policy or reaching a certain target. A weakness related to this characteristic of CEA is that multiple objectives are difficult to account for.

4.4.3 Environmental accounting

Modern national accounting is based on the standard national accounting identity that equals total income to total expenditures. To the extent possible, all measurements are based on observable transactions in the market[33]. The national accounts and their main aggregates, such as gross domestic product (GDP), gross national product (GNP) or national income (NI), record the level and nature of economic activity in a country in the accounting period. These accounts do, however, not account for negative externalities generated by economic activity – like e.g. degradation of environmental quality – or for the reduction of renewable and non-renewable natural resources.

If one wants to 'green' national accounts, one has first to decide which elements of sustainability should be included in it. In general, there are five main reasons for the fact that national income may be a poor approximation to economic welfare. The main reasons may be grouped as follows:

1. the treatment of non-marketed good and services and leisure time. Non-paid household labour is an often quoted example. Non-paid housekeeping provides services to the members of the household. These services are not recorded in the present national accounts. The number of working hours in a week have gradually fallen over the last century. We may assume that the additional leisure time is valued positively by most workers. This amenity is not recorded in national in-come accounts, however.
2. the treatment of consumer durables. Consumer durables provide services to their owners over their lifetimes. Current accounting practice, however, treats them as final consumption in the year of their purchase only. This may lead to foolish paradoxes as the implication that deliberate efforts to make goods more perish-able raise national output.
3. equity. The aggregate measure of national income or national income per capita is silent about the distribution of this income. If the distribution of income is an argument in the social welfare function, a change in national income may be a poor approximation of its effect on welfare.
4. instrumental, regrettable or defensive expenditures. Some of the expenditures that are currently classified as final output could be regarded as intermediate expenditures. The reason is that certain activities do not yield direct utility but are regrettably necessary inputs to activities that yield utility. Pollution control expenditures by governments and households could be classified as intermediate rather than as final output. Some authors go much further than this and would classify a whole range of expenditures that are related to the 'necessary overhead costs of a complex industrial nation-state' as instrumental or defensive.
5. environmental damage costs. Environmental externalities from economic activities produce economic 'bads' such as environmental pollution and natural re-source degradation that are not recorded in national accounts.

Different methods of green national accounting put a different emphasis on these key issues. Four of these are:

- MEW: Measures of Economic Welfare
- ISEW: Index of Sustainable Welfare
- Genuine Savings
- Sustainable National Income

[33] Notable exceptions are food and fuel produced and consumed by farm families and the rental value of owner-occupied dwellings.

Each of these methods is discussed in more detail in the next sections.

4.4.4 Method: Measures of Economic Welfare (MEW)

MEW (Nordhaus and Tobin, 1972) is calculated by adjusting conventional national income accounts by the factors mentioned below. It is mainly based on a reclassification of expenditures and it therefore requires little additional research.

The MEW includes corrections of conventional Net National Product (NNP) for the following factors:

- non-market activities and leisure time. The authors valued these activities and leisure time at their presumed opportunity cost, the money wage rate. The imputation of non-paid activities and leisure time more than doubles NNP.
- a reclassification of government final expenditures into intermediates, consumption and net investment, and a reclassification of some household expenditure. Education, medicine and public health expenditures are considered gross investments that raise productivity or yield household services.
- consumer durables. The treatment of consumer durables as capital goods turns out to have little quantitative effect.
- instrumental or defensive expenditures. Among these expenditures are classified: costs of commuting to work, and government services such as police, sanitation, road maintenance and national defence.
- disamenities of urbanisation. This category which includes the environmental damage costs of environmental pollution is valued by a 'disamenity premium' that is estimated as the income differential between people living in densely populated locations and people living in rural locations. The 'disamenity premium' is estimated to be about five percent of GDP.

4.4.5 Method: Index of Sustainable Welfare (ISEW)

ISEW is partly based on a reclassification of expenditures and partly on the valuation of income distribution and environmental damages. In the original study, changes in environmental pollution and resource stocks were valued with 'off-the-shelf' values from the literature. As such, the method would not prevent a more sophisticated approach towards the valuation of environmental externalities.

In contrast to Nordhaus and Tobin's MEW, the ISEW does take environmental dam-age and natural resource depletion explicitly into account. The ISEW distinguishes between water pollution, air pollution, noise pollution, loss of wetlands, loss of farm-land, and long-term environmental damage. Assessments of environmental damages for specific years are taken from literature (especially Freeman, 1982) and extrapo-lated to other years. The depletion of mineral and fuel resources is valued at their to-tal production value. Long-term environmental damage includes climate change and depletion of the ozone layer.

4.4.6 Method: Genuine Savings

The Genuine Savings indicator measures aggregate net savings in a country that takes account of the depletion of natural resources and the accumulation of pollutants. The problems of measurement and valuation of natural resource depletion and accumulation of pollutants are basically the same as in other approaches to adjust national income.

Most proposals for adjustments to national accounting focus on income or product accounts, such as GDP, NDP or national income. Pearce et al. (1998) propose an indicator of 'weak' sustainability that focuses on capital formation: they call it a measure of 'genuine savings'. In the weak sustainability 'paradigm' different types of capital are distinguished such as man-made capital, human capital, natural capital and even social capital. The operational definition of sustainable development in the weak sustainability paradigm is that the total stock of capital should be maintained as a necessary (though possibly not sufficient) condition for the sustenance of future well-being (Pearce et al., 1998). The sufficiency of the weak sustainability paradigm hangs on the extent to which different types of capital can be substituted for each other. Pearce et al. (1998) give a number of arguments (irreversibility, uncertainty, and thresholds/discontinuities) why the substitution possibilities between man-made and natural capital may be less than perfect. However, Pearce et al. argue, while an indicator based on weak sustainability may not necessarily tell us what development is sustainable, it certainly tells us what development is not sustainable. Persistent negative genuine savings rates must lead to non-sustainability in the sense that the welfare of the country will eventually decline.

4.4.7 Method: Sustainable National Income

The Sustainable National Income (SNI) approach to Green Accounting computes the hypothetical level of national income in a year that is consistent with sustainable patterns of production and consumption. The sustainable patterns of production and consumption are determined by imposing so-called 'absolute sustainability standards' on environmental themes, so that the stock of natural capital will not diminish over time. SNI measures the level of net national income (NNI) that can be consumed without future impoverishment: it is therefore a guide to prudent conduct in economic matters. The change of SNI over time measures the change towards, or away from, sustainable development.

The SNI is computed by an applied general equilibrium model in which the 'sustainability standards' act as restrictions on the feasible set of solutions. The model contains detailed abatement cost curves for environmental themes. The difference or 'gap' between SNI and conventional NNI measures the dependence of the economy on that part of its natural resource use that exceeds the sustainable exploitation levels. The gap between SNI and NNI is therefore an indicator of the extent of unsustainability of an economy.

4.5 Modelling tools

By: Hermann Lotze-Campen (PIK), Tom Kram (RIVM), Reyer Gerlach (IVM), Pim Martens, Pieter Valkering (UM-ICIS), Karl-Heinz Simon (UNIK-CESR), Alex Haxeltine, John Turnpenny (UEA-Tyndall), John Robinson (UBC-SDRI), Matthias Nerger (WI), Gnçalo Lobo (JRC).

Many scientific disciplines rely to an increasing extent on the use of computer modelling tools to represent, describe, analyse and simulate major processes related to research questions in their realm. Modelling tools help to structure scientific thinking, to focus on the most relevant processes, analyse important trade-offs between conflicting goals, define scenarios, and, to a certain extent, make predictions of likely future developments. Modelling applications to be discussed should have some relevance to the actual policy making process with regard to sustainability questions. Hence, pure theoretical models and rather conceptual models, dealing with basic mechanisms without direct link to empirical information, are not included in *SustainabilityA-Test*.

The tools have been grouped into three categories. While many other forms of categorisation would also be plausible, the current approach was chosen, because:

(1) it reflects the development and application of models in their traditional disciplinary setting and hence will be most familiar to model developers and potential users, like policy makers and the wider public,

(2) it shows the relative position of available models, along the lines of increasing thematic integration, and

(3) it illustrates the challenges for more integrated modelling, which reflects the more general challenge of truly integrated research on sustainability impacts.

Each of the three categories covers a number of modelling approaches which are treated as ‘tools’ in this report. The different tools are discussed on the basis of specific model examples, which are called ‘methods’ in this report.

The following tools will be evaluated:

1. Biophysical models, primarily covering natural-scientific phenomena:
   - Climate
   - Hydrology
   - Biochemistry
2. Socio-economic models, focussing on human dimensions of sustainability:
   - General economy
   - Partial economic sectors
   - Demography
   - Public health
3. Integrated models, covering attempts to bridge the gap between natural and social sciences and combine aspects from both domains into one modelling framework:
   - Land use:
   - Integrated assessment
   - Qualitative system analysis
   - Scenario building and planning

4.5.1 Climate

Climate models simulate long-term changes in atmospheric conditions, like temperature, precipitation and atmospheric contents of various gases. They are used both to reconstruct climate conditions in the past and predict future trends. Most climate models have a global focus and cover very long time scales, from several decades to millions of years.

4.5.2 Hydrology

Hydrology models contain mathematical descriptions of the major elements of the water system, i.e. rivers, lakes, oceans, soil, atmosphere. They describe the impact of natural (e.g. climate change) and/or anthropogenic (e.g. water withdrawals) disturbances on water flows and water cycles. They can be applied on different scales, ranging from local to global. Some models cover water quality aspects.

4.5.3 Biochemistry

Biogeochemistry models explain vegetation growth and related natural exchange processes, based on climate conditions, soil quality, nutrient and water supply. Some models focus on natural vegetation, while others deal with agricultural crops or forestry only. They can be used to simulate external effects, e.g. climate change, on vegetation growth and related material fluxes, e.g. change in soil carbon, water balances. They can also be used to simulate potential natural vegetation, e.g. for reconstructing past vegetation cover or for excluding current anthropogenic disturbance.

4.5.4 General economy

General economy models are aggregated representations of an economic system, usually a nation state (or a group of nations). They are 'closed' in a sense that they are based on a consistent accounting framework that covers the whole economy. GEM can be highly stylised, consisting of only one or very few sectors, but focussing on the complex dynamic processes of investment, innovation and economic growth, especially in the longer run. Others, like general equilibrium models or multi-sector econometric models can cover a large number of economic sectors, but have to be more restricted with respect to dynamics, structural flexibility, and time horizon.

4.5.5 Partial economic sectors

Economic sectors are important descriptive units of an economic system. Partial economic sector models have a focus on a certain sector of the economy, for which they provide much more structural detail than multi-sectoral general economy models can do. Sector models work on the simplifying assumption that major feedbacks between the specific sector and the economy as a whole, e.g. effects on employment and growth, can be neglected. Taking macroeconomic conditions and certain prices as given, the allocation and distribution effects within the sector can therefore be looked at more realistically. Moreover, specific environmental conditions and constraints can be taken into account.

4.5.6 Demography

Demography models provide long-term projections of future population changes, based on external scenarios on natural and anthropogenic influences. Major driving forces are changes in environmental conditions, fertility, epidemiology, mortality, and more general socio-economic conditions.

4.5.7 Public health

Public health models are discussed and evaluated here with special regard to the Modelling framework for the health Impact ASsessment of Man induced Atmospheric changes (MIASMA). MIASMA consists of five models: the Thermal Stress Model, the Skin Cancer Model, the Malaria Model, the Dengue model, and the Schistosomiasis Model. The models are driven by scenarios of population figures and atmospheric changes, superimposed on baseline data regarding disease incidence, climatic conditions, and ozone-layer thickness.

4.5.8 Land use

Land use change models (LUC) are integrated models that link important economic activities like agriculture, forestry, transport or energy production, with environmental processes. Land use activities represent some of the most intensive and closest links between society and nature. Some LUC focus more on biophysical determinants of human land use activities, while others are more closely linked to economic decision models that treat biophysical conditions as decision constraints. LUC have been applied on very different spatial scales, ranging from single farms to global coverage.

4.5.9 Integrated assessment

Integrated assessment models try to link, within a single modelling framework, main features of society and economy with the biosphere and climate systems. Starting with a focus on the connection between anthropogenic greenhouse gas emissions and climate change, the agenda of IAM now includes aspects of land use, biogeochemistry, hydrology, demography and health. The goal is to make more and more parts of the 'Earth system' endogenous to the modelling framework. This is an ongoing process, in which major methodological barriers between scientific disciplines have to be overcome. E.g. the optimisation mode of most economic models has to be linked with process-based time-step models in climate or biogeochemistry research.

4.5.10 Qualitative system analysis

QSA approaches try to structure and analyse socio-economic processes and their environmental implications only based on qualitative information. Qualitative information (only directions of change, no numerical measures) is less demanding for data providers and can be used under circumstances where quantitative assessments are not available, or where quantitative information is not strictly comparable. Nevertheless, QSA is based on mathematical functions and hence retains a rigorous approach. This assures that consistent chains of argumentation can be derived even from relatively vague initial statements.

4.5.11 Scenario building and planning

SBP models are highly integrative tools which are capable of representing a wide variety of social and natural compartments of the Earth system. They can be used to develop and structure complex scenarios, and also for analysing these scenarios in interactive stake-holder participation settings. They can be easily applied to analyse various policy measures and other human activities, e.g. environmental management.

4.5.12 Evaluation of all model categories

Whether or not computer models are useful for the analysis of complex real-world phenomena, especially in an interdisciplinary setting, is subject to debate. Modellers would argue that the application of rigorous mathematical methods provides more structure and transparency to the analysis of complex problems. It also makes it in some cases easier to communicate problem formulation, as compared to pure verbal descriptions of a research approach. Critics would probably argue that models have to make use of simplifications and exogenous assumptions to such an extent that in many cases it renders them useless for saying anything substantial about the problems to be analysed. The choice of the model structure may in many cases to a certain degree predetermine the outcomes. Moreover, while being simplifications of reality, many scientific models remain so complex that they are rather seen as black boxes instead of transparent research machines. Hence, some of the great strengths of modelling tools are felt as serious weaknesses by non-modellers.

Complexity vs. simplicity: In order to model complex real phenomena, it is necessary to make simplifying assumptions. This reduces the real problem to the core aspects relevant for the analysis. Models have the strength that they make simplifications in a structured and transparent way. Only this approach facilitates the simulation of complex processes in the first place. However, oversimplifications and unrealistic assumptions can be a weakness. The acceptance of model results for non-modellers very often depends on the acceptance of basic assumptions.

Quantitative vs. qualitative aspects: The strengths of mathematical representations and simulations of natural or social processes can only be fully exploited if the quantification of these processes is acceptable to the modeller and the model-user. For many phenomena, especially in the social sciences, only insufficient quantitative information is available to describe and analyse the underlying processes. In these cases a quantitative analysis may be misleading. However, most mathematical models are ill-equipped for dealing with qualitative information and usually quantifiable proxy variables are used.

Endogenous vs. exogenous processes: Especially in the field of SIA it becomes apparent that ‘almost everything depends on everything else’, i.e. in a dynamic world there are hardly any exogenous variables. However, most models require a distinction of what to include in the model and what to leave out, in order to keep the involved processes tractable. This important decision during the model development may have strong implications for the model results and their interpretation. A structured discussion about exogenous and endogenous elements of a certain model can also play an important role in structuring and tailoring the problem to be analysed. Hence, this aspect should be made clear and transparent to the model user.

Specialisation vs. integration: There is a clear trade-off between model specialisation and integration. The more focussed a model is on a specific real-world process, the more appropriate will be the model representation with regard to this process, but on the other hand more related external processes have to be ignored or treated as exogenous. Integrated models try to include many linkages between different domains in an explicit way, but in order to keep the overall complexity under control they will often have to rely on simplified representations of the single elements involved. Increasing computer power partly helps to make integrated models also more sophisticated, but the basic trade-off remains and has to be acknowledged by model developers and model users. There will be no ‘one-size model that fits all purposes’, as this would probably be not very useful.

4.6 Scenario analysis tools

By: Karl-Heinz Simon (CESR), Alexa Matovelle (CESR), Joe Alcomo (CESR), Claudia Pahl-Wostl (ISF), John Robinson (SDRI)

A manifold of definitions exists in practice and in the literature on what ‘scenario analysis’ is. However, most scenario developers would agree that scenarios are made up of a set of explicit ‘if-then’ propositions that explore the consequences of a range of driving force assumptions. A scenario can take many forms including an image, a graphic, a table, or text.

In particular, scenarios are constructed especially to assist in the understanding of possible future developments of complex systems.

From a theoretical perspective, scenarios can be classified in five different dimensions:
1. Qualitative / Quantitative scenarios: meaningful for the approach (e.g. combination with modelling tools) or problem area chosen;
2. Anticipatory / Exploratory scenarios: relevant for the direction of analysis (looking forward or back in time);
3. Participatory / Non-participatory scenarios: whether stakeholder participation is, or is not, included in the process of scenario analysis;
4. Baseline / Policy scenarios: whether we deal with policy or baseline scenarios
5. Scale (from global to theme-specific);

These dimensions will be taken into account during the evaluation of the tool scenario-analysis. However, for categorising scenarios from a perspective of their role in assessments, these dimensions cannot be used, as within assessments often scenarios are being used that comprise a combination of the different dimensions. Therefore, a differentiation is made on the basis of the use of scenarios in assessments, namely whether existing scenarios are to be used or new ones have to be developed. In this tool group, therefore, the following 'tools' are being evaluated:

- using existing scenarios: selecting a relevant scenario as the basis for various assessment tools that require data or qualitative information with a time-dimension as their input;
- developing new scenarios: the process of developing a new (set of) scenarios on the basis of which various assessment tools calculate various aspects (impacts, costs, benefits et cetera).

4.6.1 Using existing scenarios

Several reasons justify such a decision to rely on existing scenario analyses. Often the exercise is nearly finished and not much additional time is left to start new studies. Also the budget restrictions often will prevent the involvement of additional personnel for additional scenario studies. Other arguments for using existing scenarios are that in some problem areas indeed well elaborated and reliable studies exist, as is the case with population projections for example.

When existing scenarios are the material used in the assessment a careful analysis is necessary to see whether there is sufficient correspondence in problem recognition and framework parameters (time characteristics, fundamental conceptual basis, etc.).

In Table 4.3 there is a preliminary listing of EU-related scenario studies. As can be seen many issues relevant in sustainability assessments have still been treated in these studies, like energy, land use, transport, employment etc.

Table 4.3: EU and Global Change Related Scenario Projects

Issue	Year	Author / Institution	Title	Reference
Green house gases	2002	EEA	The ShAir scenario – Towards air and climate change outlooks, integrated assessment methodologies and tools applied to air pollution and greenhouse gases	http://reports.eea.eu.int/topic_report_2001_12/en
Environment	2003	EEA	Europe's Environment: The Third Assessment	http://reports.eea.eu.int/environmental_assessment_report_2003_10/en
Environment	2003	EEA	Europe's Environment: The Second Assessment	http://reports.eea.eu.int/92-828-3351-8/en
Environment	1999	EEA	Scenarios and Outlooks in Environment in the European Union at the turn of the century. CD-ROM / Background Information	http://eea.eionet.eu.int:8980/irc/eionet-circle/eu98ap/info/data/main.htm
Transport	2002	European Commission DG Transport & Energy	SCENES European Transport Scenarios	http://www.iww.uni-karlsruhe.de/SCENES/#funding
Sustainability	2001	European Commission DG R&D	VISIONS	http://www.icis.unimaas.nl/visions/
Land use	1999	Van Latesteijn (WRR)	Land Use in Europe : A methodology for policy-oriented future studies. WRR-voorstudie nr.106	http://www.wrr.nl/ne/voorhoofdstukken.php?rapportnr=106
Population	1999	European Commission (ECFIN)	The Economic Consequences of ageing population (A comparison of the EU, Us and Japan).	http://www.edwardhugh.net/files/population_and_ageing.pdf
Sustainability	1999	European Commission, Forward Studies Unit	Scenarios Europe 2010	http://europa.eu.int/comm/cdp/scenario/index_en.htm
Transport	1998	European Commission DGVII	POSSUM – policy scenarios for sustainable mobility	www.cordis.lu/transport/src/possum.htm
Transport	1998	Nijkamp P., Rienstra S., Vleugel J.	Transportation Planning and the Future (book)	http://www.bookworkz.com/construction/planning/0471974080.html
Transport	1997	Hey, Hijkamp, Rienstra & Rothenberger	Assessing Scenario on European Transport Policies by means of Multicriteria analysis	http://www.tinbergen.nl/discussionpapers/97086.pdf
Green house gases	1997	ECN	Scenarios for Western Europe on long term abatement of CO2 emissions	http://www.ecn.nl/docs/library/report/1997/c97051.pdf
	1997	Smith, D.	Eurofutures – Five scenarios for the next Millenium (book)	http://www.ecn.nl/library/reports/1997/c97051.html
Sustainability	2004	European Commission	Resource use scenarios for Europe in 2020	http://www.seri.at/Data/seri/publications/documents/SERI%20Studies%201.pdf
Sustainability	2003-2006	European Commission	MOSUS	http://www.mosus.net/index.html
Sustainability	1996	European Commission	Progress report from the Commission on the Implementation of the European Community programme of policy and action in relation to the environment and sustainable development	
Energy	1996	European Commission	European energy to 2020: a scenario approach	http://europa.eu.int/comm/energy/en/etf_2_en.html
Sustainability	1996	Wuppertal Institute	Towards a sustainable Europe	

Rural environment	1995	Schoute	Scenario studies for the rural environment: selected and edited proceedings of the symposium Scenario studies for the rural environment	http://www.neutrino.co.jp/abi_enpo/0-7923-3748-4.PDF
Energy	1995	ECN	Energy scenarios for a changing Europe: Integration versus fragmentation	http://www.ecn.nl/library/reports/1995e/c95075.html
Employment	1994	European Foundation for the Improvement of Living and Working Conditions	The potential of employment opportunities from pursuing sustainable development	
Rural environment	1992	WRR	Ground for choices, four perspectives for the rural areas in the European Community	
Environment	1992	RIVM	The environment in Europe: a global perspective	
Urbanisation	1992	European Commission	Urbanisation and the functions of cities in the European Community	
Environment	1989	IIASA	Future environments of Europe	

Evaluation: In general, including scenarios into assessments has the advantage of bringing into discussion various options and different strands of alternative modes of development or problem solutions. And it provides a methodological framework to foster communication amongst scientists, politicians and laymen.

When existing scenarios are used that tool has relatively modest requirements of fiscal and time resources. On the other side, flexibility is restricted because of the obligation to use the concepts and models development in external study groups and, possibly, under different problem recognitions. Policy options chosen in the existing scenario studies can not be extended or inclusion of additional options is extremely difficult. Therefore, also the transparency is sometimes low because of the necessity to make ad-hoc decisions during the procedure and because the documentation of scenario finding and evaluation processes might be incomplete or preliminary. An additional problem might arise from the manifold of existing scenario studies which makes it difficult to decide which ones to trust and which ones to reject.

One source for these kinds of problems is the lack of a standardised, well accepted method for scenario definition, evaluation and documentation.

4.6.2 Developing new scenarios

Developing new scenarios is a procedure for evaluating future developments and/or assessing response strategies and, as has been described above, is helpful to explore alternative futures. In assessments scenario analysis is helpful for clarifying development options and for communication amongst different disciplines and participants from science, politics and civil society.

Procedural components are the description of scenarios, a comparison of scenario results, and an evaluation of their consequences. When new scenarios have to be developed a careful analysis of the problem situation, the targets, the technical, economic and administrational options, and the framework conditions is crucial. Several suggestions and well-proven techniques are available to give support to that task. In assessments scenario analysis helps, on the one side, with the organisation of information and with keeping decisions transparent and open, especially by allowing various actors to participate in the preparation of decisions (the organisational aspect). On the other side, it guarantees the consideration of variations of driving forces and framework conditions which might have serious effects on the outcomes and effects of suggested strategies and problem solutions (the aspect of content). In general, a careful communication on the time characteristics of the scenario, its main features, considered step-wise changes and the included driving forces is necessary.

Characteristic methodologies are, for example, the von Reibnitz approach (Reibnitz, 1999) where eight working steps are described reaching from task analysis to scenario transfer (or application) or the story and simulation approach (developed by J. Alcamo) with nine steps starting with the establishment of the scenarios group and ending with publication and distribution of the results of the analysis (Alcamo, 2001).

Evaluation: When building on new scenarios than a certain control is guaranteed over the conceptual base, the scenario definitions and the criteria for evaluation. Therefore, more transparency exists and a perhaps better fitting to the problem at hand might be achievable. Due to the possibility to include qualitative as well as quantitative information scenario analysis provides a good opportunity for not overlooking crucial aspects which are not formulated in certain standardised data formats or accepted as hard data.

Weaknesses are related to the partly subjective character of the procedure due to the already mentioned lack of a standardised, well accepted method. Thus, also unrealistic, biased and unproven assumptions might find their way into the scenarios, e.g. triggered by political settings. It is the responsibility of the scenario group to avoid as far as possible such shortcomings. When new scenarios are to be developed the procedure is time consuming, needs significantly more man-power than the use of existing scenarios and will take also some time for iterations where scenario definitions and results have to be criticised by external experts and lead to modifications of the scenarios by the working group. In some circumstances those resource requirements cannot be met in ongoing political processes with completely different time constants.

4.7 Multi-Criteria Analysis tools

By: J. David Tàbara (IEST), Nadja Kasperczyk (IfLS), Karlheinz Knickel (IfLS), Gregor Meerganz (IEST), Daniela Russi (IEST), Marjam Van Herwijnen (IVM)

Multi-Criteria Analysis (MCA) is a tool that allows taking into account simultaneously a wide range of concerns in complex decision-making processes. A variety of different techniques, i.e. methods, often with rather similar sounding titles, is available, such as multi-criteria decision analysis, multi-attribute utility theory, the analytic hierarchy process, or fuzzy set theory. MCA approaches can take into account a wide range of concerns, and each concern is usually being expressed by one or more criteria which can simultaneously be assessed.

4.7.1 Tool: MCA

The core of a MCA is formed by methods that can be used to rank the alternative (policy) options. The outcome of carrying out an MCA is a ranking of the options, on the basis of the criteria and each option's score. An MCA is therefore useful when a number of alternative options need to be compared with one another in an assessment.

MCA helps to structure a policy problem on the basis of the recognition of *plurality* of alternatives and preferences. Understood in a wide sense, MCA is not simply a 'tool' or a method where some computer programs can be applied in a non-reflexive way. Rather it can be seen as a philosophy or an attitude in the understanding of policy problems, where the consideration of only value positions or policy options is discarded. In this sense, the current developments of MCA methodologies tend to focus on the MCA *process* rather that only the final outcome. For instance, the principal interest in Multiple-Criteria Decision Aid (MCDA, Roy, 1985) is not to discover one solution, but rather lies in its constructive or creative approach. Therefore, every aspect of the entire decision process, leading to a given outcome, very strongly contributes to the final decision's quality and success.

In MCA policy problems are seen as complex systems constituted by various functions (objectives or policy alternatives) which depend of various criteria, which means that no single solution would optimise all the criteria simultaneously. MCA methods are well suited in elaborating compromise solutions because various dimensions or criteria are incorporated into the decision process and the aggregation procedures used do not exclude dominated alternatives a priori. This is why MCA is now becoming widely used to structure complex decision problems, which include multiple and also conflicting goals. MCA can help to overcome problems of aggregation of incomparable values which are typical in the assessment of sustainability issues.

Three main types of methods for ranking alternative options can be distinguished in a MCA:

1. Compensatory methods: based on the aggregation of criteria to form unique meta-criteria. Compensation among criteria is allowed. Examples that will be evaluated in *SustainabilityA-Test* are:
   - MAVT
   - weighted summation
   - AHP, and
   - Fuzzy MCA;
2. Outranking methods: based on the pair-wise comparisons of alternative policy options and their outranking relations. Partial compensation among criteria is allowed. Examples that will be evaluated in *SustainabilityA-Test* are:
   - PROMETHEE
   - NAIADE
   - Electre
   - REGIME;
3. Non-compensatory methods: based on the assumption of no compensation between the criteria. Examples are:
   - Dominance method (evaluated within *SustainabilityA-Test*)
   - Conjunctive and disconjunctive selection procedure (not evaluated)
   - Lexicographic ordering (not evaluated).

Each method listed here – except the fuzzy MCA, Electre and the last two – is shortly discussed in what follows.

4.7.2 Method: MAVT

Multiple- attribute value theory (MAVT) is a compensatory technique, that is, it considers acceptable to trade-off between different criteria, so that low scores on one criterion may be compensated by high scores on another. The options' performance across all the criteria can be aggregated to form an overall assessment. MAVT assumes the 'mutual independence' of preferences between policy alternatives (although in the Keeney and Raiffa (1976) approach such assumption is not held). This means that the judged strength of preference for an option on one criterion will be independent of its judged strength of preference on another.

MAVT can be used in those cases where two or more alternative policies are involved that have to be evaluated and where two or more possibly conflicting objectives for the desired state of the system are involved. Two objectives are conflicting whenever getting closer to one objective implies having a worse performance in relation with the other one. One or more different attributes/criteria are used to measure the performance in relation to that objective. It does so by assuming that in every decision problem a real value function U exists that represents the preferences of the decision maker. This function aggregates for each alternative the criteria that are under consideration by the decision maker. The best alternative is then the maximisation of the set of value functions.

A closely related theory to MAVT is multiple-attribute utility theory (MAUT). The latter theory requires stronger assumptions to ensure additivity, but in practice it is rather similar to MAVT (Winterfeldt and Edwards, 1986; French, 1988).

4.7.3 Method: Weighted summation

Weighted summation is a compensatory method, which means that 'bad' criterion scores can be compensated by 'good' ones. Weighted summation can be used to address problems that involve a finite and small set of alternative policies that have to be evaluated on the basis of conflicting objectives.

For any given objective, one or more different attributes or criteria are used to measure the performance of each alternative option in relation to that objective. These aspects, the impacts of all alternative options for all attributes, are presented in a so-called evaluation table. The attributes are usually measured on different measurement scales and therefore cannot be compared with each other. Weighted summation makes the attributes comparable, prioritises them and finally reduces the amount of information in order to provide comprehensibility, the strengths and weaknesses of the policies and ultimately a ranking of the alternatives policies.

Applying weighted summation is only possible if information about priorities of criteria is available. The decision maker has to give his/her preferences with respect to the evaluation criteria incorporated into the decision model. These preferences are expressed in priorities or weights and indicate the importance of the criteria relatively to the other criteria. The most straightforward method to assess weights is the *direct estimation* of their relative importance by assigning a value to each criterion. This method assigns weights to criteria using a scale such as 0 to 10 or 0 to 100. If used carefully, direct estimation can be an effective methodology. Because the direct estimation method appears to be a very difficult task for a decision maker, especially if the number of criteria is large and the criteria are very different in character, different methods to estimate the relative importance of the criteria have been developed.

4.7.4 Method: AHP

The Analytic Hierarchy Process (AHP) has been developed by T. Saaty (1977, 1980, 1988, 1995) and is one of the best known and most widely used MCA approaches. It allows users to assess the relative weight of multiple criteria or multiple options against given criteria in an intuitive manner. In case quantitative ratings are not available, policy makers or assessors can still recognise whether one criterion is more important than another. Therefore, pair wise comparisons are appealing to users. Saaty established a consistent way of converting such pair wise comparisons (X is more important than Y) into a set of numbers representing the relative priority of each of the criteria.

The basic procedure to carry out the AHP MCA consists of the structuring of a decision problem and selection of criteria, priority setting of the criteria by a pair wise comparison (weighing), a pair wise comparison of options on each criterion (scoring) and obtaining an overall relative score for each option.

4.7.5 Method: Regime

Regime is a MCA qualitative method based on the possibility of partial compensation among the different criteria which affect the evaluation of the various policy alternatives. MCA qualitative methods are used when some or all data are not available in quantitative terms, and qualitative criteria and measurements must be applied. In this regard, decision makers working in government are frequently faced with circumstances where the information in the performance matrix, or about preference weights, consists of qualitative judgements. A number of methods exist to respond to this (NERA 2004) and among those, one can find the following:

- The *Regime method* can handle qualitative information on scores and priorities. The method provides a complete ranking and information on the relative certainty of the results (Hinloopen et al., 1983; Israels and Keller, 1986; Nijkamp et al., 1990; Janssen, 1992).
- The *permutation method* can handle qualitative information on scores and priorities. However, the method provides a ranking that is not necessarily complete (Paelinck, 1974, 1977; Ancot and Paelinck, 1982; Ancot, 1988).
- The *evamix method* has been especially designed to handle mixed qualitative/-quantitative information on scores and quantitative information on priorities. It provides a complete ranking and information on the relative qualities of the alternatives (Voogd, 1983; Nijkamp et al., 1990, Janssen, 1992).
- The *expected value method* can handle qualitative information on scores and priorities. The method provides a complete ranking and information on the relative differences between alternatives (Rietveld, 1980, 1984; Nijkamp et al., 1990; Janssen, 1992).

In *SustainabilityA-Test* only the *Regime* method will be evaluated.

Regime is a discrete method, that is, a method that compares a finite set of alternatives. It can use binary, ordinal, categorical and cardinal (ratio and interval rate), and also mixed information. Qualitative information is transformed into quantitative in order to be treated. It is a concordance analysis, meaning that it is based on pair wise comparison between alternatives according to some chosen criteria in order to establish a rank between them. It was developed by Nijkamp in 1982 (Hinlopen et al., 1983).

Regime uses as input an impact matrix and a set of weights. The first one resumes information about the various impacts of the alternatives in relation to the chosen criteria. The weights express the (politically determined) relative importance of the criteria.

4.7.6 Method: PROMETHEE

The method called Preference Ranking Organisation MeThod for Enrichment Evaluations (PROMETHEE) was developed by Brans (1982), further extended by Brans and Vincke (1985) and Brans and Mareschal (1994), is an outranking method, typical of the European (or French) MCA school.

Outranking methods are characterised by the limited degree to which a disadvantage on a particular viewpoint may be compensated by advantages on other viewpoints (Pirlot, 1997). The degree of dominance of one option over another is indicated by outranking (Vincke, 1992). Promethee is a well established Decision Support System (DSS) which deals with the appraisal and selection of a set of options on the basis of several criteria, with the objective of identifying the pros and the cons of the alternatives and obtaining a ranking among them. With PROMETHEE, as an outranking method, strong assumptions concerning the ‘true’ preference structure of the decision maker are avoided. In the evaluation of decision alternatives, the key question is whether there is enough information to state that one alternative is at least as good as another. On the basis of so-called outranking relations, build in a first step, a ranking of alternatives is derived.

4.7.7 Method: NAIADE

Naiade (Novel Approach to Imprecise Assessment and Decision Environments) is a MCA method able to deal with both qualitative and quantitative information concerning policy socio-environmental evaluation problems. NAIADE was developed by G. Munda and technical information on his theoretical background and application procedure can be found in Munda (1995). NAIADE can provide the following information: (a) Ranking of policy alternatives according to a set of evaluation criteria, (for instance, a compromise solution/s); (b) Indications of the distance of the positions of the various interest groups (e.g., possibilities of convergence of interests or of coalition formation; (c) ranking of the alternatives according to the actors’ impacts or preferences.

NAIADE also performs an equity and conflict analysis in order to identify those alternatives which could reach a certain degree of consensus or would provide a higher degree of equity among different interests groups. It is a very flexible method suitable for real-world applications, and in particular, for situations where *fuzzy uncertainty* or indeterminacy is recognised. In particular, fuzzy uncertainty regards not only to the difficulties to set probabilities about the occurrence of a particular event but especially to the difficulties to describe the event itself in an unambiguous manner.

The application of NAIADE entails the construction of an impact (evaluation) matrix which includes, on one axis, a limited set of given policy alternatives, and on the other, a limited set of different criteria by which such policy alternatives are to be evaluated. Inside the matrix measurements of all the policy alternatives with respect to each evaluation criterion are given and the different alternatives are assessed by means of pair wise comparison. This aggregation procedure takes into account both the number of the criteria in favour of each alternative and the intensity of the actors' preferences. In principle, the determination of the criterion scores is independent from the actors' preferences; for instance, impact on employment can be chosen or used as a criterion to evaluate a given set of policy alternatives but the score of each criterion may not depend on the actors' preferences

4.7.8 Method: Dominance Method

The dominance method can be used to address problems that involve a finite and discrete set of alternative policies that have to be evaluated on the basis of conflicting objectives. An objective is a statement about the desired state of the system under consideration. Two objectives are conflicting if improving one objective means that the state of the other objective is getting worse.

For any given objective, one or more different attributes or criteria are used to measure the performance in relation to that objective. These aspects, the impacts of all alternative options for all attributes, are presented in a so-called evaluation table. These attributes are usually measured on different measurement scales.

The dominance method is a non-compensatory method (ODPM, 2004). This means that the method does not allow compensation of weak performance of one criterion by a good performance of another criterion. The method does not aggregate the data but processes them in a different way. The method indicates one or more alternatives that perform better or equal on all criteria compared to the other alternatives. However, the likelihood of an option dominating or being dominated by all scores is very small so that in most cases the method will not supply a result.

The dominance method is a non-compensatory approach, that is, it does not allow any compensation among criteria. The set of alternatives are stepwise eliminated without trading off their deficiencies. The dominance method bases this elimination entirely on the criteria scores. An alternative is dominated and hence eliminated from considerations, if there is another alternative, which is better on one or more criteria and is equal on the remaining criteria. For further descriptions see Jankowski (1995).

4.7.9 Evaluation of all MCA methods

The MCA as a tool to support decision making is strong in clarifying the strengths and weaknesses of alternative policies along a set of criteria. When these criteria are agreed upon by stakeholders, MCA helps to streamline discussion and negotiations between different stakeholders and as such provides a (transparent) means of communication about the problem at stake and its possible solutions. MCA requires that those involved agree upon what impacts to include, how to assess or estimate these and how important each is. This can be a time consuming and difficult process.

An overview of the various methods that can be used to compare different criteria in an MCA, and the strengths and weaknesses of these methods, is given in the table below.

Table 4.4: An overview of the various methods that can be used to compare different criteria in an MCA, and the strengths and weaknesses of these methods

Method	Strengths	Weaknesses
MAVT	– Reduces the amount of information in order to provide comprehensibility	– Composing the value function is a very difficult task – Support of a decision analyst is needed – Criteria have to be mutually independent – Little experiences are known
Weighted summation	– Simple methodology, easy to use and explain – Clear distinction between objectivity and subjectivity – Software availability – Accommodates various types of information (quantitative and qualitative)	– Assessing/estimating impact scores is difficult and time consuming – Assigning weights is difficult – Criteria have to be mutually independent
AHP	– Users generally find the pairwise comparison form of data input straightforward and convenient. – AHP method has the distinct advantage that it decomposes a decision problem into its constituent parts and builds hierarchies of criteria. – AHP method supports aid group-level decision making through consensus by calculating the geometric mean of the individual pairwise comparisons.	– As a complete aggregation method, compensation between good scores on some criteria and bad scores on other criteria can occur. Detailed and important information can be lost. – The number of pairwise comparisons may become very large and therefore decision making may become a lengthy task – The artificial limitation of the use of the 9-point scale can be an important disadvantage – Ranking irregularities can occur when AHP or some of its variants are used. – AHP requires quite complex calculations, which would be undertaken in practice by the corresponding software.
Regime	– The most important advantage of Regime is that is can use different type of information. This flexibility is very important with real- world cases, were there is complexity and many data are not available in quantitative terms. – Able to deal with mixed information (both quantitative and qualitative of various dimensions: economic, environmental, social, etc.). – Ability to assess the impact of a given alternative in their original unit. – As all outranking methods, it is more realistic than the methods based on value functions, such as for example MAVT, because it does not reduce complexity to one single dimension. – The software makes the use of Regime user-friendly, so that it is not difficult to use for non-experts.	– Some level of technical expertise is required for its sound use and correct understanding. – As all outranking methods, Regime is less axiomatised than MAVT. – In the aggregation procedure, some information is lost. – The establishment of the weights might be problematic
PROMETHEE	– PROMETHEE supports to aid group-level decisionmaking through consensus. – PROMETHEE as an outranking method can simultaneously deal with qualitative and quantitative criteria; criteria scores can be left in their own units, which is important when they relate to different domains. – PROMETHEE (as outranking method) has the distinct advantage that it can deal with uncertain and fuzzy information.	– PROMETHEE does not provide the possibility to really structure a decision problem – This MCA tool does not provide any formal guidelines for weighing. – Ranking irregularities can occur when a new alternative is introduced. – The way in which the preference information is processed is complicated and hard to explain to non-specialists.
NAIADE	– Able to deal with mixed and incommensurable information (both quantitative and qualitative of various dimensions: economic, environmental, social, etc.). – Ability to assess the impact of a given alternative in their original unit. – Can use information affected by different types and degrees of uncertainty, including fuzzy data (where the phenomena to deal with cannot be defined in an unambiguous manner, e.g. indeterminacy). – Ability to attribute values to the alternatives' criteria expressed in the form of either crisp (exact and certain), stochastic, fuzzy numbers or linguistic expressions.	– Complexity of the model renders it little intuitive: many parameters need to be determined (aggregation and compensation coefficients like alfa, gamma, etc.) – Notable level of technical expertise is required for its sound use and correct understanding. – Fastidious participatory and social process: Four preference and indifference thresholds to be defined in agreement with all social actors involved. – No weights can be applied. – No possibility of directly using ordinal criteria (converted in qualitative or *crisp* criteria).
Dominance method	– Simple methodology, easy to use and explain – Accommodates various types of information (quantitative and qualitative)	– In practice, dominance is rare. Therefore, the extent to which DomM can help to support real decisions is limited. – Quite sensitive to errors in the data of the performance matrix.

4.8 Sustainability appraisal tools

By Måns Nilsson (SEI), Kerstin Ehrhart (SEI), Dirk Günther (ISF), Anneke Klasing (Ecologic), Kata Wagner (Ecologic), Gina Ziervogel (SEI-Oxford).

The tools in this group have different scopes, perspectives and methodologies, but they also have things in common. A first common characteristic is that the instruments are intended to be integrated as strategic decision support that can operate in decision-making processes. Second, they can all contain multiple elements of methods and as such are framework tools that can be filled with different analytical content. Third, they can all be applied on the strategic level (such as programmes and legislative proposals). Fourth, their scope is often connected to the general sustainability context (usually in connection with the three-pillars-model), although the assessor has considerable freedom in selecting what impact variables will be used. Fifth, due to the flexibility of the tools, and their role as framework tools connecting to policy processes, the sustainability appraisal tools as delineated here, with the possible exemption of Vulnerability Assessment, can act as a bridge between the other tool teams and the policy-making process. This also means that certain issues outlined as important in the review, such as specific tools, methods, models, including who developed them, is not completely relevant in this case.

4.8.1 (Sustainability) Impact Assessment (S)IA

With its communication COM (2002) 276 (CEC, 2002a) the European Commission developed a comprehensive approach to impact assessment. One motive behind the current initiatives in the EU is the establishment of more efficient and 'leaner' decision-making procedures. Another motive was for ensuring a coherent implementation of the EU's Sustainable Development Strategy is the commitment to include a sustainability impact assessment[34] for all major policy proposals, as set out in the Gothenburg Conclusions of the European Council[35]. Sustainability impact assessments also play a decisive role in contributing to an effective and efficient regulatory environment[36] in the framework of the governance process[37]. The system of EU impact assessments is outlined in a Commission communication. Impact assessments were gradually applied to the major initiatives presented by the Commission in its Annual Policy Strategy or its Work Programme 2003, which were expected to have an economic, social and environmental impact. The system is to be fully operational in 2004/2005.

Evaluation: The impact assessment as outlined by the Commission and the Strategic Environmental Assessment provide open tools which offer the opportunity to appraise various aspects of sustainable development. One of the main problems is that the selection of policy proposals to undergo an SIA/SEA follows no systematic procedure. Moreover, its application and the extent to which the aspects of sustainable development are covered strongly depend on the officer responsible for the SIA/SEA. Financial and time constraints and the availability of data appear to have a great effect on the thoroughness of the assessments. Critics also assert that issues such as cumulative impacts, long-term effects, irreversibility and limits to sustainability are not adequately addressed, and therefore the validity of the scoring mechanism for assigning 'significant' impacts is questionable.

[34] In the following the abbreviations SIA, IA, EU SIA are used synonymously.
[35] Presidency Conclusions of the Göteborg European Council, 15 and 16 June 2001.
[36] European Commission 2002: Communication from the Commission: Action Plan „Simplifying and Improving the Regulatory Environment". COM(2002)278 final.
[37] European Commission 2001: European Governance. A White Paper. COM(2001)428 final.

Despite these tools having a potential to link to the decision-making process, the practice of applying these tools have also shown a limited impact of the assessment itself in the decision making process. This could be caused by a limited range of policy options assessed, the lack of objectivity when presenting the results and the absence of a non-technical summary.

4.8.2 Strategic Environmental Assessment (SEA)

SEA is a decision-making support tool that aims at integrating the environmental aspects of decisions on programmes and plans that have long-ranging implications on broader aspects of society. SEA developed rapidly in the 1990s on the conceptual basis of environmental impact assessment (EIA). It was originally considered to be an application of EIA in strategic decisions, such as the formulation of policies, plans and programmes. SEA can be categorised as one of the administrative instruments for environmental policy integration into the highest levels of decision making, including proposed policies, legislation, plans and programmes (terms which mean different things in different countries). The basic idea for a strategic environmental assessment (SEA) on the European level was developed back in the late 1970s by a small network of experts. From the mid 1980s there have been several attempts of the Commission to adopt regulations on SEA. An official directive entered into force in 2001. The recent European Directive on environmental assessments of certain plans and programmes provides a legislative framework for SEA. The directive requests the competent authorities to elaborate an environmental statement and to perform consultations with the environmental authorities and the general public. However, the directive only addresses plans and programmes whereas SEAs in principle can be applied also to policy levels, and as such might overlap with Impact Assessments.

<u>Evaluation</u>: see SIA.

4.8.3 Vulnerability Assessment (VA)

VA has emerged as central concept in sustainability sciences in later years. A complete, systematic procedure for analysing vulnerability is not yet available. The choice of suitable tools depends on the environmental system of interest and socio-economic vulnerability determined by what aspects of vulnerability are being measured (the scale being assessed, vulnerability to what, over different time periods, qualitative or quantitative nature). The practical concerns of an assessment – the time and resources available, target audience, and required outcomes – are also relevant. The following five methods illustrating approaches in VA are distinguished:

- indicators: a common approach to representing complexity
- vulnerability mapping: extends indicators to considerations of spatial scale
- livelihood indicators: draw on the narrative behind vulnerability.
- syndromes and scenarios combine qualitative and quantitative methods.
- agent based social simulation: one way to represent the dynamism of vulnerability.

<u>Evaluation</u>: Vulnerability assessments allow a range of quantitative and qualitative elements of sustainability to be captured simultaneously, identifying clearly where policy interventions are needed, both on the short and long term. However, VA is also a concept requiring very detailed and rapidly changing information with respect to exposure to stresses and the capacity to cope with these, placing an enormous pressure on the data availability. Furthermore, it is less apt to operate within the framework of assessing the consequences of a

particular intervention such as a policy or a plan. Its role in current practice is more related to understanding the baseline and a situation assessment based on impacts/responses to existing stressors.

4.8.4 Indicators for sustainable development

Since the UN Conference for Environment and Development in 1992 in Rio de Janeiro Indicators are popular tools to assess sustainability in numbers. Chapter 40 of the Agenda 21 ask national and international institutions to contribute to the development of indicators for sustainable development (UNCED, 1992). They should define both, the state of a (eco- and/or socio-) system operationalised in different state characteristics (e.g. good ecological state operationalised by number of fish, habitat quality etc.) and a 'desirable' future state of a system. Furthermore indicators should measure how to achieve this future state and how far away from it the current situation is. Therefore indicators are measures and codes to document stages and transitory trends in different problem fields of Sustainable Development.

Evaluation: Indicator systems provide a simplified view of the status or development of a certain system and are therefore easy to communicate to the general public and to policy makers. However, each simplification is subjective, mostly due to the selection of one indicator (set) over the other. An indicator system therefore always is arbitrary and risks being repelled by decision makers. Furthermore, indicators are usually quantitative: if no data exist, no indicator exists, thereby risking overlooking crucial aspects. Indicator systems can often play an instrumental role in the ex post evaluation and follow-up if they are built into the implementation and follow-up of the policy.

4.9 Participatory tools

By: Matthijs Hisschemöller, Åsa Swartling, Marleen van de Kerkhof and Eefje Cuppen. Consensus conference: Eefje Cuppen; Focus Groups: Åsa Swartling. Repertory Grid Technique: Marleen van de Kerkhof. Backcasting: Matthijs Hisschemöller. Electronic Focus Groups: Gonçalo Lobo. Tools to Inform Debates, Dialogues & Deliberations: Ângela Guimarães Pereira.

Participatory methods can be defined as: 'Methods to structure group processes in which stakeholders play an active role in order to articulate their knowledge, values and preferences' (Van Asselt and Rijkens-Klomp, 2002: 168). The research field of PIA is characterised by an overwhelming variety of participatory methods and techniques, which stem from a broad range of scientific disciplines. Some participatory methods have been developed more recently, whereas others are rather classical. Depending on the objectives of a specific PIA research project, a participatory method is often used in combination with other (participatory) methods.

The participatory tools that will be evaluated in *SustainabilityA-Test* have all been designed to be applied in an interactive mode, claim to enhance creativity and learning among the participants involved, and been applied in participatory assessments with respect to sustainability issues. It concerns the following tools:

- consensus conference;
- focus groups;

- repertory grid technique;
- backcasting;
- electronic focus groups;
- tools to inform debates dialogues and deliberations (TIDDD)

Evaluation: All participatory methods contribute to increasing public awareness by actually involving the public in the discussion and through the media attention that could be a result of that involvement. Involving the public helps with putting the relevant points of view on the table. It also lets the experts, policymakers and lay men learn from these different points of view, which contributes to the accountability of the policy making process. Involvement of stakeholders furthermore improves the democratic design of policy making, which contributes to the quality of policy making. A general problem with all tools that aim to involve stakeholders is that the selection of stakeholders involved in the process is always 'just' a sample of the population. Making a generalisation on the basis of that sample, i.e. drawing conclusions for the whole population, is tempting, but not valuable.

Each of these tools will be explained in further detail in what follows.

4.9.1 Consensus conference

The consensus conference is a participatory method, which is aimed at involving the lay public in the policy making process. Hereby it can raise public awareness, may lead to better decisions, may increase the legitimacy and accountability of decision making and it may stimulate learning (as well for the public as for the decision makers and experts).

Important characteristics of this method are that the public determines the agenda for the conference and chooses which experts to consult. The consensus conference is usable for topics which are socially relevant, which can be delimited, which imply technological/ scientific knowledge and which have to deal with unclear opinions/points of view.

The consensus conference aims to give a voice to the lay public by forming a lay panel (or citizen panel). The panel (a group of 10-30 citizens) formulates the questions to be taken up and participates in the selection of experts to answer these questions. At the end a report is produced containing the consensus view (expectations, concerns and recommendations) of the panel regarding the issue at hand.

Evaluation: With sufficient media attention, consensus conference can contribute to increasing public awareness. It may furthermore increase making better decisions by enriching the process with relevant points of view present within the general public. Involving lay men in the discussion increases the legitimacy of decision making and thereby the accountability and democratic value of it. However, the effect of consensus conference on the policy making itself is still rather unclear. In addition, the desire to develop a shared position between lay men and experts could result in a loss of deviating points of view. The focus for the facilitation process on cooperation skills and expert information hides an inability to deal with the quality of argumentation that can be expected from citizens.

4.9.2 Focus groups

The focus group technique has become an important applied approach to integrate stakeholders' perspectives and knowledge into Integrated Assessment (IAs). Powell et al (1996:499) define a focus group as: 'a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research'.

There are many variations of the basic method, but generally, focus groups are a method for collecting qualitative research data through carefully planned group discussions with the purpose of obtaining perceptions of participants in a permissive and non-threatening environment (Morgan, 1988). The overall objective is to identify patterns and trends in people's perceptions about issues related to a certain research theme. The discussions are guided by a skilled moderator, who works from a predetermined set of questions. The group members influence each other by responding to comments raised in the discussions. The results are analysed with quantitative and qualitative social science methods. If conducted and analysed properly, focus groups are likely to stimulate learning and increased awareness among participants as well as promote more democratic and effective decision making.

<u>Evaluation</u>: This tool generates a rich understanding of stakeholder perceptions, experiences and believes, but it is a relatively expensive tool to set up. The fairly elastic boundaries of FG definitions often results in confusion about appropriate use and design of FGs, and runs the risk of being misused for sale attempts, educational seminars, therapy or consensus building exercises. The outcome of a FG is sometimes difficult to interpret and can usually not be generalised to a larger population. In addition, it will not tell us how people will actually respond or behave, no matter what they claimed.

4.9.3 Repertory grid technique

Repertory Grid Technique is a method from construct psychology that can best be characterised as a particular form of structured interview. It articulates people's individual 'construct systems', a construct system being defined as the set of qualities, or dimensions, that people use to interpret their experience of the world. The Repertory Grid Technique helps to better understand what meaning people give to a certain problem situation and what kinds of solutions they would prefer.

The Repertory Grid Technique unfolds people's categorisations and preferences by facilitating a systematic comparison and ranking of 'elements', i.e. the objects of people's thinking with regard to a certain topic or problem, to which they relate their concepts and values. In other words: Repertory Grid Technique gives insight into the way in which people make sense of the world around them, which may help to clarify people's perceptions of problems and solutions to these problems.

<u>Evaluation</u>: When used in a participatory process, the RGT has the capacity to enhance the quality of the argumentative process by facilitating the exploration of conflicting arguments and (underlying) claims on a specific topic (Van de Kerkhof, 2004). With a limited number of interviews (20 to 25) the RGT is able to elicit the true range of relevant constructs in a particular context (Dunn, 2001). RGT is able to develop the intersection between objective and subjective methods of assessment: it targets the articulation of deeply personal meanings and enables the comparison or compilation of these meanings vis-à-vis the meaning of others (Bannister, 1985, referred to in Neimeyer, 2002). The outcome of RGT is influenced by the chosen techniques for e.g. eliciting, sorting and rating, which places challenges on the researcher.

4.9.4 Backcasting

Already in 1976, Lovins introduced the ‘backwards-looking-analysis’ in his exploration of long-term energy policy in the United States (Lovins, 1976). In 1982, Robinson elaborated on this approach and developed the method of backcasting as an alternative approach to traditional forecasting and planning methods (Robinson, 1982). It has been especially designed as a tool for exploring sustainable policies. Backcasting exists of two parts: 1) develop the image of a desired future and 2) use models / scenarios to backcast the required pathways.

Originally, backcasting was not meant to be applied in an interactive mode. However, climate change projects in Canada and, more recently, in the Netherlands have developed interactive approaches to backcasting. Whereas traditional backcasting tends to take the desired end-state for granted and analyses how it maybe achieved, second-generation backcasting puts more emphasis on an iterative process to define the desired future result itself. In this process the stakeholder may learn about what is desirable through confrontation with the consequences of specific initial preferences. (Robinson, 2003: 849). From the perspective of research and modelling, backcasting provides an opportunity to involve the anticipated users of knowledge in the design and actual research itself.

An important feature of the tool is that it enables to analyze alternative images of the future, thoroughly analyzed as to their feasibility and consequences. Each alternative must appear as coherent and the analysis of consequences for social life must be credible. Backcasting should identify strategic choices for society and for decision makers of all kinds.

4.9.5 Electronic focus groups

Electronic focus groups (eFG) are internet platforms where virtual debates and discussions can take place. They have a similar role in assessment processes as standard focus groups (e.g. explore stakeholder perspectives on sustainability issues regarding a certain policy). However, in eFG the participants can be located in any place in the world with internet access and the discussion is normally done in a written format. In essence, electronic focus groups, are private internet chats where participants and moderators are invited people and with a maximum of 10 to 15 persons present. The discussion chat can be complemented with other collaborative features such as a white board or a file exchange area.

Evaluation: Electronic focus groups are relatively cheap, and offer the opportunity to reach many people within a large geographical area. Transcription is obviously not needed. However, the electronic counterpart of focus groups is only open to people with computers, with good typing skills and with the capability to express their opinions by writing instead of talking, without any visual contact.

4.9.6 Tools to Inform Debates Dialogues and Deliberations (TIDDD)

TIDDD is a general concept tool that aims at fostering debate among stakeholders in a governance process characterised by conflict or other dynamics that entail specific contexts for debate, dialogue or deliberations. It is an evolution of the Decision Support System concept, except that its aim is not to backup decisions, but to provide context for initial dialogues among relevant stakeholders or other social actors in governance issues. TIDDD is conceived to be used in participatory processes; they are informatics tools that mediate relevant knowledge of different types to inform debates in extended governance processes. As such they are meant to be platforms for organisation, communication and exchange of knowledge coming from different sources, expressed in different formats including associated uncertainties. TIDDD deploy methods that account for diversity of knowledge representation, such as multi-criteria evaluation and multi-dimensional visualisation tools.

Evaluation: TIDDD enhances the knowledge base necessary to initiate a debate in terms of organisation communication and exchange. It can have different roles depending on the context in which is applied; apart from it being a platform for debate, it can also have didactic roles; the principle of ‘progressive disclosure of information’ ensures this. The knowledge provided is socially robust in the Gibbons sense, since it is checked against the possible users of the ‘system’. But, building a TIDDD is very much time consuming and requires specific skills (social and informatics).

5 Experiences with the evaluation framework, the tool overview and grouping

This chapter describes the experiences with the evaluation framework for setting out the plans for future work on this framework and for reflecting on the tool overview and grouping.

5.1 Experiences with the evaluation framework

In the ‘evaluation of the evaluation framework’ we start with a few general observations on the framework itself followed by a description of some of the more specific issues related to the evaluation criteria.

5.1.1 General observations

Most of the feedback received on the tool evaluation framework consisted of questions on the applicability of the evaluation criteria to all tools. Whereas a criterion like *data needs* is relevant for models, it is irrelevant for a tool like *consensus conference*, for example. In short, the evaluation criteria are not equally relevant for all tools. No serious problem is posed to the evaluation in a situation in which the criteria are irrelevant to a certain tool; one can just simply consider the criteria to be ‘irrelevant’. However, in a situation where a relevant aspect of a tool is not captured by the evaluation criteria, a more serious situation exists, since in this case, the evaluation fails to address the most relevant characteristics of each tool. This may mean that for certain tool teams, additional criteria have to be formulated in order to be able to address the relevant characteristics of the tools. In the subsequent phases of the project, therefore, further attention should be directed toward the development of complementary criteria for certain tools, if needed.

For some tool groups it remains to be seen if the level at which the tools are currently being evaluated will render the sort of information that will be useful to the eventual users. An example of this is the tool group on models: whether a model would be useful in certain assessments depends on the model itself. If we evaluate at the level of model categories (e.g. *partial economic sector models* instead of PRIMES), it remains unclear whether such information would be useful to someone searching for the best way to assess the possible impacts of a certain proposal on energy demand, as in this example. The level of detail required for a useful evaluation is of particular importance for the following tool groups:

- Monetary assessment tools: evaluate at either the level of CBA or of methods to estimate costs and/or benefits;
- Models: evaluate at either the level of model categories or of the particular models themselves;
- Scenario analysis: evaluate the role of existing scenarios or new scenarios in assessments, or evaluate at the level of specific examples from existing scenarios or specific methodologies to develop new scenarios.

Lastly, for some procedural tools (such as *strategic environmental assessment* and *sustainability impact assessment*), the scoring of certain evaluation criteria fully depends on other tools being deployed within the context of these procedural tools. Consequently, the evaluation of these procedural tools for such criteria (e.g. the coverage of aspects for sustainable development) becomes difficult – if not impossible.

5.1.2 Evaluation criteria

Policy processes: The evaluation with respect to the policy processes supported by the tools works well. For most tools it appears to be feasible to assess which policy processes can be supported with them and which not.

Aspects of sustainable development: The two-level approach with respect to topics that can/cannot be addressed by each tool is not being used. The evaluation is usually carried out at the 'environment', 'social' and 'economic' levels, without further specifying what topics within each pillar can actually be addressed with the tool. The tool's ability to address the cross-cutting aspects of sustainable development is usually specified, although it appears that some toolevaluators find it difficult to understand what exactly is meant by each aspect.

Operational aspects: The operational aspects are usually specified, although not all operational aspects are relevant for each tool. For most of the tools it appears to be difficult to assess criteria such as costs and time needed to apply the tool. This problem can be partly solved by asking for a minimum and maximum value for these criteria.

5.2 Experiences with the tool overview and grouping

The tool overview is now almost complete, with all of the tool overview papers. Whether the tool inventory is complete or not – i.e. whether all of the tools commonly used in (sustainability) assessments are included – remains to be seen. The review of national level assessments in phase 2 and an external peer review may be helpful in this respect.

The grouping of tools is functional as it separates the project team into manageable units. We have to keep in mind, however, that tools have been assigned to certain tool teams on the basis of project management criteria (e.g. efficiency), which have little to do with the characteristics of the tool itself. Examples of tool grouping that need critical reflection:

- Models as scenario-building tools: currently partly covered in the tool group on modelling and also mentioned in the tool group on scenario analysis; this overlap is no problem as long as we are certain that models used for scenario-building are covered sufficiently in either one of the two groups.
- Participatory modelling: currently not being addressed (although preparations are being made to include it in the tool group on participatory tools);. we have to decide which is the most logical group to discuss this tool.

These two examples are also related to combining tools; this aspect will be analysed in more detail in the case study.

5.3 Conclusions on the evaluation framework, the tool overview and grouping

The experiences with the evaluation framework shows the evaluation framework to work fairly well for the tool group on physical assessment tools and multi-criteria analysis tools. However, for the other tool groups (monetary assessment tools, modelling tools, scenario analysis tools, sustainability appraisal tools and participatory tools), the combination of the evaluation framework with the level at which the tools are currently being evaluated will, most likely, not provide the usable information needed to actually develop ‘an evaluation matrix capable of showing which tools can be part of which methods or broader instruments or approaches to measure, and assess the three pillars of sustainable development and strategy definition’. Two different causes can be identified:

1. The evaluation framework does not contain the most relevant evaluation criteria; i.e. it does not ask the right questions. This applies in particular to the procedural tools, *sustainability impact assessment* and *strategic environmental assessment* within the group on sustainability appraisal tools and to all participatory tools. Additional evaluation criteria will most likely be needed to address the *procedural* role these tools have in assessments.
2. The evaluation framework is not being applied at the right level of the tools; i.e. the ‘unit’ that we evaluate is not the most efficient ‘unit’ to evaluate. This is particularly true for the monetary assessment tools (where monetisation methods rather than the tool cost-benefit analysis is the level at which the evaluation framework can render specific information), the model tools (where information on the specific models rather than on model types is needed to know which pillars of sustainable development can be addressed with them) and the scenario analysis tools (where the different existing scenarios and methods to build new scenarios are probably found at the level at which the evaluation framework will deliver the most useful information). However, in all three cases, adjustment of the evaluation criteria could also be a solution.

Further effort is therefore needed to overcome the difficulties sketched above (see below):

1. Critically analysing each tool group to determine if we are evaluating on the most efficient level and subsequently introducing adjustments to that level, if necessary (I&S team in cooperation with the tool teams);
2. Critically analysing each tool to determine if we are addressing the most relevant aspects with the evaluation framework and subsequently formulating additional evaluation criteria if necessary (the tool teams in cooperation with the I&S team);
3. After making the necessary adjustment(s), continuing to fill in the evaluation framework for each tool, specifying how well each tool can address the various aspects relevant for sustainable development (the tool teams).

References

General

Brewer, DeLeon, 1983. Brewer, G. and P. DeLeon, The foundations of policy analysis. Dorsey, Homewood, USA.

Boersema J. J. and Reijnders, L. (eds), to be published. Principles of Environmental Sciences, chapter 13: Environmental modelling, H.J.M. de Vries. Kluwer Academic Publishers, the Netherlands

CEC, 2001. A Sustainable Europe for a Better World: A European Union Strategy for Sustainable Development (Commission's proposal to the Gothenburg European Council). Communication from the Commission of the European Communities (COM(2001)264 final). Brussels, Belgium. http://europa.eu.int/eur-lex/en/com/cnc/2001/com2001_0264en01.pdf

CEC, 2002a. Towards a global partnership for sustainable development. Communication from the Commission of the European Communities (COM (2002) 82 final), Brussels, Belgium. http://europa.eu.int/eur-lex/en/com/cnc/2002/com2002_0082en01.pdf

CEC, 2002b. Communication from the Commission on impact assessment. Communication from the Commission of the European Communities (COM (2002) 276 final), Brussels, Belgium. http://europa.eu.int/eur-lex/en/com/cnc/2002/com2002_0276en01.pdf

CEC, 2003. A handbook for impact assessment in the Commission – How to do an Impact Assessment. Brussels, Belgium. http://europa.eu.int/comm/secretariat_general/impact/docs/imp_ass_how_to_en.pdf

Council of the European Union, 2000. Presidency conclusions of the Lisbon European Council, 23 and 24 March 2000, Brussels, Belgium. http://ue.eu.int/ueDocs/cms_Data/docs/pressData/en/ec/00100-r1.en0.htm

Eurostat, 2004. The EU Sustainable Development Strategy: A framework for indicators. SDI Workshop 9–11 February 2004, Stockholm, Sweden (see 'SDI-TF-030 Rev. 5 SDI Framework' at http://forum.europa.eu.int/Public/irc/dsis/susdevind/library?l=/policysdocuments&vm=detailed&sb=Title).

High Level Group, 2004. Facing the Challenge. The Lisbon strategy for growth and employment. Report from the High Level Group chaired by Wim Kok, November 2004, p. 37, Brussels, Belgium. http://europa.eu.int/growthandjobs/pdf/kok_report_en.pdf

Hirschemöller, M. and Gupta, J., 1999. Problem-solving through international environment agreements: The issue of regime effectiveness. International Political Science Review 20(2): 151–174.

IPCC, 2000. Special Report on Emission Scenarios. Intergovernmental panel on climate change (IPCC). http://www.grida.no/climate/ipcc/emission/index.htm

Meadows, D. et al., 1972. The Limits To Growth. Universe Books, New York, USA.

MNP, 2004. Kwaliteit en toekomst. Verkenning van duurzaamheid. [Quality and future. Exploring sustainability]. The Netherlands Environmental Assessment Agency (RIVM-MNP), Bilthoven, the Netherlands.

MNP/TNS-NIPO, to be published. Analyse enquêtes wereldbeelden en maatschappelijke vraagstukken [Analyses of surveys on worldviews and social dilemma's]. Bilthoven/Amsterdam, the Netherlands.
UN, 1992. Report of the United Nations' conference on environment and development. United Nations (UN) meeting in Rio de Janeiro, Brazil, 3-14 June 1992. http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm
UN, 2000. United Nations Millennium Declaration. Adopted by the General Assembly at United Nations Headquarters in New York (6 to 8 September 2000). http://www.un.org/millennium/declaration/ares552e.htm
Vries, B., 2004. Who says beta should also say gamma. Presentation by B. de Vries (RIVM-MNP, Copernicus Institute University of Utrecht), NOW-NOVEM conference. Utrecht, the Netherlands. http://arch.rivm.nl/ieweb/ieweb/publications/presentations/NWONOVEMBetaGammaMaart04/sld001.htm
WCED, 1987. World Commission on Environment and Development ('Brundtland Commission', Our Common Future, Oxford University Press, United Kingdom.

Physical assessment tools

Ayres and Kneese 1969. 'Production an consumption, and externalities', American Economic Review 59, 282-96.
Ritthoff, M.; Rohn, H. ; Liedtke, C, 2002. Calculating MIPS : resource productivity of products and services**.** Wuppertal, Germany. http://www.wupperinst.org/Publikationen/Wuppertal_Spezial/ws27.pdf
Bringezu, S., H. Schütz and S. Moll, 2003. 'Rationale for and Interpretation of Economy-Wide Material Flow Analysis and Derived Indicators', Journal of Industrial Ecology 7(2), 43-64.
EEA, 2000. Environmental Signals 2000. EEA regular indicator report. Chap. 16. Development in indicators: Total Material Requirement. Environmental assessment report No 6, Copenhagen, Denmark.
Eurostat, 1997. 'Materials Flow Accounting – Experience of Statistical Offices in Europe', Eurostat (Statistical Office of the European Communities), Directorate B: Economic Statistics and Economic and Monetary Convergence, Luxembourg, Luxemburg.
Eurostat, 2001. 'Economy-wide material flow accounts and derived indicators. A methodological guide', Eurostat (Statistical Office of the European Communities, 2001). Luxembourg, Luxemburg. http://europa.eu.int/comm/eurostat/Public/datashop/print-product/EN?catalogue=Eurostat&product=KS-34-00-536-__-I-EN&mode=download
Fischer-Kowalski and Haberl, 1993. 'Metabolism and colonisation – modes of production and the physical exchange between societies and nature', Schriftenreihe Soziale Ökologie, Band 26, Vienna, Austria.
Frankl, P. and Rubik, F. (eds.), 2000. 'Life Cycle assessments in Industry and Business, Adoption Patterns, Applications and Implications', Berlin, Heidelberg, New York: Springer, USA.
Gofman, K., Lemeschew, M. and N. Reimers, 1974. 'Die Ökonomie der Naturnutzung – Aufgaben einer neuen Wissenschaft [Economics of utilisation of nature – tasks for a new science]'
ISO, 1996. International Standardisation Organisation (ISO) (1996), 'Environmental management – Life cycle assessment – Principles and framework – ISO 14040', Paris, France.

Japanese Environmental Agency, 1992. Japanese Environmental Agency (1992), 'Quality of the encironment in Japan 1992', Tokyo, Japan.
Keuning, S.J. and de Haan, M., 1996. What's in a NAMEA? Recent results of the NAMEA-Approach to Environmental Accounting, Occasional Papers NA-080, Statistics Netherlands, Voorburg, the Netherlands.
Keuning, S.J. 1997. SESAME: An Integrated Economic and Social Accounting System, International Statistical Review 65, 111-121
Keuning, S.J., 1998. Accounting for Welfare with SESAME, in: Household Accounting: Experiences in the Use of Concepts and Their Compilation, Studies in Methods, Series F, No. 75, United Nations, New York, USA.
Moll, S., Bringezu, S., Schütz, H., 2003. Resource Use in European Countries - An estimate of materials and waste streams in the Community, including imports and exports using the instrument of material flow analysis (Zero Study). European Topic Centre on Waste and Material Flows, Copenhagen, Denmark. http://www.europa.eu.int/comm/environment/natres/index.htm
Rogich, D.G. et al, 1992. 'Trends in material use- implications for sustainable development', paper presented at the conference on sustainable development: energy and mineral resources in the circum-pacific region and the environmental impact of their utilisation, 9-12 march 1992, Bangkok, Thailand.
Schütz, H., Moll, S., Steger S., 2003. Economy-wide material flow accounts, foreign trade analysis, and derived indicators fort he EU. Study for the European Commission, DG Eurostat/B1, Contract no. 200241200011. Wuppertal Institute, Wuppertal, Germany.
Schütz, H. and S. Bringezu, 1993. Major material flows in Germany, Fresenius Environmental Bulletin 2, 443-8.
Society of Environmental Toxicology and Chemistry (SETAC), 1993. 'Guidelines for Life-cycle Assessment: A "Code of Practice"', SETAC workshop in Sesimbra, Portugal 31 March-3 April, Brussels, Belgium.
Steenge, A.E., 1997. On Background Principles for Environmental Policy: "Polluter Pays", "User Pays" or "Victim Pays", 121-136 in: Boorsma, B., Aarts, K. & Steenge, A.E. (eds.): Public Priority Setting: Rules and Costs, Dordrecht, the Netherlands.
Steenge, A.E., 1999. Input-Output Theory and Institutional Aspects of Environmental Policy, Structural Change and Economic Dynamics 10, 161-176
Steurer, 1992. 'Stoffstrombilanz Österreich 1988 [Material flow accounting Austria 1988]', Schriftenreihe Soziale Ökologie, Band 26. Universitäres Institut für Interdisziplinäre Forschung und Fortbildung (IFF) [Institute for intersisciplinary research and education]. Vienna, Austria.
United Nations, 1976. 'Draft guidelines for statistics on materials/energy balances', UN document E/CN.3/493. New York, USA.
Van de Ven, P., Kazemier, B., and Keuning, S.J., 1999. Measuring Well-Being with an Integrated System of Economic and Social Accounts, Occasional Papers NA-090, Statistics Netherlands, Voorburg, the Netherlands.
Wackernagel, M., Onisto, L., Bello, P., Linares, A. C., Falfan, I. S. L., Garcia, J. M., Guerrero, A. I. S. and Guerrero, C. S., 1999. "National natural capital accounting with the ecological footprint concept." Ecological Economics **29**(3): 375-390.
Wenzel, H., M. Hauschild and L. Alting, 1997. 'Environmental Assessment of Products – Volume 1, Methodology, tools and case studies in product development', Chapman & Hall London, United Kingdom.
Wernick, I.K., 1996. 'Materialisation and dematerialisation – measures and trends', Daedalus: The liberation of environment 125, 171-98.

Monetary assessment tools

Freeman A.M., 1982. Air and Water Pollution Control – A Benefit-Cost Assessment, John Wiley & Sons, New York, USA.

Nordhaus, W. D. and J. Tobin, 1972. "Is Growth Obsolete?", In: Economic Growth, National Bureau of Economic Research, General Series No. 96, pp.1-80, New York, USA.

Pearce, D.W., G. Atkinson and K. Hamilton, 1998. "The Measurement of Sustainable Development", In: J.C.J.M. van den Bergh and M.W. Hofkes [eds.], Theory versus Implementation of Sustainable Development Modelling, Kluwer, Dordrecht, the Netherlands

Modelling tools

None in the inception report – all references to specific models can be found in the preliminary tool overview and evaluation paper at www.SustainabilityA-Test.net.

Scenario analysis tools

Alcamo, J., 2001. Scenarios as Tools for International Environmental Assessments. Environmental Issue Report No. 24. Experts Corner Report: Prospects and Scenarios No. 5. European Environment Agency (EEA), Copenhagen, Denmark.

Reibnitz, 1999. SCENARIOS+VISION. Managing and Planning in Turbulent Times. How Scenario Techniques Help you Plotting a Successful Path into the Future. Lecture held at the SBM Conference – Cannes, France.

Multi-criteria analysis tools

Ancot, 1988. Micro-Qualiflex: an interactive software package for the determination and analysis of the optimal solution to decision problems. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Ancot and Paelinck, 1982. Recent experiences with the Qualiflex multicriteria method. In: J.H.P. Paelinck (ed.), Qualitative and Quantitative Mathematical Economics (pp. 217-266), Martinus Nijhoff, The Hague, the Netherlands.

Brans, J.P., 1982. L'ingéniérie de la décision. Elaboration d'instruments d'aide à la décision. Méthode PROMETHEE. In: Nadeau, R., Landry, M. (Eds.), L'aide à la decision: Nature, instruments et perspectives d'avenir, Presses de l'Université Laval, Québec, Canada, 183-214.

Brans, J.P., Mareschal, B., 1994: The PROMCALC and GAIA decision support system for MCDA. Decision Support Systems 12, 297-310.

Brans, J.P., Vincke, Ph., 1985: A preference ranking organisation method: The PROMETHEE method. Management Science 31, 647-656.

French, S., 1988. Reading in Decision Analysis. Chapman and Hall, London, United Kingdom.

Hinlopen E., Nijkamp P., Rietveld P, 1983. Qualitative discrete multiple criteria choice models in regional planning, Regional Science and Urban Economics 13, pp.77.102

Israels,A.Z. and W.J.Keller, 1986. Multicriteria analyse voor kwalitatieve data. Kwantitatieve methoden, Vol. 7, No. 22. p. 49-74.

Jankowski,P., 1995. Integrating Geographical Information Systems and Multiple Criteria Decision-Making Methods. International Journal of Geographical Information Systems 9(3), 251-273

Janssen,R., 1992. Multiobjective decision support for environmental management, Kluwer Academic publishers, Dordrecht, the Netherlands.

Keeney, R. and Raiffa, H., 1976. Decisions with multiple objectives: preferences and value trade-offs. Wiley: New York, USA.
Munda, G., 1995. Multicriteria evaluation in a fuzzy environment. Theory and applications in ecological economics, Physica-Verlag, Heidelberg, Germany.
NERA, 2004. Department of the Environment Transport and the Regions - Defra, UK - Environmental Protection 2000- Economics and Appraisal. NERA for UK WIR (Water Industry Research). Multicriteria Decision Analysis and Environmental Appraisal. UK. www.defra.gov.uk/environment/economics/
Nijkamp, P., Rietveld, P. and Voogd, H. 1990. Multicriteria evaluation in physical planning. North-Holland, Amsterdam, the Netherlands.
ODPM, 2004. DTLR multi-criteria analysis manual. Office of the Deputy Prime Minister (ODPM), Corporate Publication, Government UK. Internet: www.odpm.gov.uk/stellent/groups/odpm-about/documents/page/odpm_about_608524.hcsp
Paelinck, J.P., 1974. Qualitative multiple criteria analysis, environmental protection and multiregional development. Papers of the Regional Science Association, Vol. 36, pp. 59-74.
Paelinck, J.P., 1977. Qualitiative Multiple Criteria Analysis: An application to airport location. Environment and Planning, Vol. 9, pp. 883-895.
Pirlot, M., 1997. A Common Framework for Describing Some Outranking Methods. Journal of Multi-Criteria Decision Analysis, 6: 86-92.
Rietveld, P., 1980. Multiple Objective Decision Methods in Regional Planning, North Holland, Amsterdam, the Netherlands.
Rietveld, P., 1984. Public choice theory and qualitative discrete multicriteria evaluation. In: G. Bahrenberg (ed.), Recent developments in spatial data analysis, Gower, Aldershot, pp. 409-426.
Roy, B., 1985. Méthodologie multicritère d'aide à la décision, Economica, Paris, France.
Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15: 59-62.
Saaty, T.L., 1980. The Analytical Hierarchy Process, McGraw Hill International, New York:, USA.
Saaty, T.L., 1988. The Analytical Hierarchy Process, McGraw Hill International, New York:, USA.
Saaty, T.L., 1995. Decision Making for Leaders, The Analytic Hierarchy Process for Decisions in a Complex World. RWS Publications, Pittsburgh, USA.
Vincke, Ph., 1992. Multi-Criteria Decision Aid. Wiley: New York, USA.
Von Winterfeldt, D. and W. Edwards, 1986. Decision Analysis and Behavioural Research. Cambridge University Press: Cambridge, United Kingdom.
Voogd, H., 1983. Multicriteria Evaluation for urban and regional planning. Pion, London, United Kingdom.

Sustainability appraisal tools

CEC, 2001. European Commission 2001: European Governance. A White Paper. COM(2001)428 final. Brussels, Belgium.

CEC, 2002a. Communication from the Commission on impact assessment. Communication from the Commission of the European Communities (COM (2002) 276 final). Brussels, Belgium. http://europa.eu.int/eur-lex/en/com/cnc/2002/com2002_0276en01.pdf

CEC, 2002b. Commission of the European Community: Communication from the Commission: Action Plan ‘Simplifying and Improving the Regulatory Environment’. COM(2002)278 final, Brussels, Belgium.
Council of the European Union, 2001. Presidency Conclusions of the Göteborg European Council, 15 and 16 June 2001.
European Parliament and Council of the European Union, 2001. Directive 2001/42/EC of the European Parliament and of the Council on the assessment of the effects of certain plans and programmes on the environment, June 27, 2001
Hey, C., 1998. Nachhaltige Mobilität in Europa: Akteure, Institutionen und politische Strategien. Opladen,Wiesbaden: Westdeutscher Verlag. Germany.
Jacob, K and Volkery, A, 2003. Potentials and Limits for Policy Change Through Governmental Self-Regulation – The Case of Environmental Policy Integration, paper presented at the 2002 Berlin Conference on the Human Dimension of Global Environmental Change, Berlin, Germany. (See: http://www.fu-berlin.de/ffu/akumwelt/bc2002/proceedings/bc_2002_jacob_volkery.pdf)
Sadler, B., 2001. A Framework Approach to Strategic Environmental Assessment: Aims, Principles and Elements of Good Practice, paper presented at the International Workshop on Public participation and health Aspects in Strategic Environmental Assessment, November 23-24, 2000, Szentendre, Hungary
Therivel, R., Wilson, E., Thompson, S., Heaney, D. and Pritchard, D., 1992. Strategic Environmental Assessment, Earthscan Publications Ltd., London, United Kingdom
UN, 1992. Report of the United Nations’ conference on environment and development. United Nations (UN) meeting 3-14 June 1992, in Rio de Janeiro, Brazil. http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm

Participatory tools

Dunn, W., 2001. Using the method of context validation to mitigate Type III errors in environmental policy analysis. In: Hisschemöller, M., R. Hoppe, W. Dunn and J. Ravetz (eds.). Knowledge, power and participation in environmental policy analysis, 417-436. Transaction Publishers. New Jersey. USA.
Lovins, A.B., 1976. Energy strategy: the road not taken? In: Foreign Affairs 55 (1): 63 – 96.
Neimeyer, G.J. and Hagans, C.L., 2002. More madness in our method? The effects of repertory grid variations on construct differentiation. In: Journal of Constructivist Psychology 15: 139-160.
Morgan, D., 1988. Focus Groups as Qualitative Research. Sage Publications. Newbury Park, CA, USA.
Powell R.A. and Single H.M., 1996. Focus Groups. In International Journal of Quality in Health Care, Vol. 8, No. 5: 499-504.
Robinson, J., 1982. Energy backcasting: a proposed method of policy analysis. In Energy Policy 10 (4): 337 – 344.
Robinson, J., 2003. Future subjunctive: Backcasting as social learning. In Futures 35: 839-856.
Van Asselt, M.B.A. and Rijkens-Klomp, N., 2002. A look in the mirror. reflection on participation in integrated assessment from a methodological perspective. Global Environmental Change, Vol. 12: 167 – 184.
Van de Kerkhof, M., 2004. Debating climate change. A study of stakeholder participation in an integrated assessment of long-term climate policy in the Netherlands. Lemma Publishers. Utrecht, the Netherlands.

Annex 1 List of project partners

Participant name	Participant short name	Country
Institute for Environmental Studies/Vrije Universiteit	IVM-VUA	the Netherlands
The Tyndall Centre for Climate Change Research	Tyndall-UEA	United Kingdom
Institute of Environmental Systems Research/ University of Osnabrück	USF-UOS	Germany
Institute for Rural Development Research at Goethe	IfLS	Germany
National Institute for Public Health and the Environment	RIVM	the Netherlands
Czech Environmental Institute	CEI	Czech Republic
Potsdam Institute for Climate Impact Research	PIK	Germany
Joint Research Centre	EC-JRC	Italy
Stockholm Environment Institute	SEI	Sweden
International Centre for Integrative Studies	ICIS	the Netherlands
Center for Environmental Systems Research	CESR	Germany
Institute of Environmental Sciences and Technology	IEST-UAB	Spain
Ecologic, Institute for International and European Environmental Policy	Ecologic	Germany
Fondazione Eni Enrico Mattei	FEEM	Italy
Wuppertal Institute	WI	Germany
Institute for Environmental Science and Management	IESAM	Latvia
University of Twente	UT	the Netherlands
Sustainable Development Research Initiative/British Columbia	SDRI-UBC	Canada

Annex 2 The project teams and their members

I&S team
Leader: RIVM (Bart Wesselink and/or Wouter de Ridder)
Members: Ecologic (Anneke Klasing)
EC-JRC (Angela Guimaraes Pereira and/or Gonçalo Lobo
IfLS (Karl-Heinz Knickel)
IVM (Marjan van Herwijnen)
Tyndall (Alex Haxeltine and/or John Turnpenny)
UOS-USF (Claudia Pahl-Wostl and/or Dirk Günther)

Peer group
Leader: IVM (Marjan van Herwijnen)
Members: EC (Daniel Deybe)
EC-JRC (Angela Guimaraes Pereira and/or Gonçalo Lobo)
EEA (Ann Dom)
IfLS (Karl-Heinz Knickel)
RIVM (Bart Wesselink and/or Wouter de Ridder)
Tyndall (Alex Haxeltine and/or John Turnpenny)
UOS-USF (Claudia Pahl-Wostl and/or Dirk Günther)

Tool team 1: physical assessment tools
Leader: Philipp Schepelmann (WI)
Members: Stefan Bringezu (WI)
Matthias Nerger (WI)
Stephan Moll (WI)
Helmut Schütz (WI)
Karl-Heinz Simon (CESR)
Anne van der Veen (UT)

Tool team 2: monetary assessment tools
Leader: Anna Alberini (FEEM)
Members: Julia Bartos (CEI)
Benjamin Gorlach (Ecologic)
Onno Kuik (IVM-UVA)
Anne van der Veen (UT)

Tool team 3: modelling tools
Leader: Hermann Lotze-Campen (PIK)
Members: Reyer Gerlach (IVM)
Alex Haxeltine (Tyndall)
Tom Kram (RIVM)
Gonçalo Lobo (JRC).
Pim Martens (UM-ICIS)
John Robinson (UBC-SDRI)
Karl-Heinz Simon (UNIK-CESR)
John Turnpenny (Tyndall)
Pieter Valkering (UM-ICIS)

Tool team 4: scenario analysis tools
Leader: Karl-Heinz Simon (UNIK-CESR)
Members: Alexa Matovella (CESR)

Tool team 5: multi-criteria analysis tools
Leader: Joan David Tabara (ICTA)
Members: Marjan van Herwijnen (IVM)
Nadja Kasperczyk (IfLS)
KarlHeinz Knickel (IfLS)
Gregor Meerganz von Medeazza (ICTA)
Daniel Russi (ICTA)

Tool team 6: sustainability appraisal tools
Leader: Måns Nilsson (SEI)
Members: Kerstin Ehrhart (SEI)
Dirk Günther (USF-UOS)
Anneke Klasing (Ecologic)
Kata Wagner (Ecologic)
Gina Ziervogel (SEI-Oxford)

Tool team 7: participatory tools
Leader: Matthijs Hisschemöller (IVM)
Members: Eefje Cuppen (IVM)
Dirk Günther (UOS)
Marleen van de Kerkhof (IVM)
Gonçalo Lobo (EC-JRC)
Angela Guimaraes Pereira (EC-JRC)
Åsa Gerger Swartling (SEI)

Tool team 8: transition management tools
Leader: Pim Martens (ICIS)
Members: Mita Patel (ICIS)

Annex 3 Details of the list of environmental, social and economic topics

The following pages contain the detailed list of impacts of which the main categories are listed in Table 3.2 on page 33. These main categories are listed in **bold** text.

1. Economic

1.1 Economic growth
- 1.1.1 Encourage or discourage fixed-capital investment, including in clean technologies?
- 1.1.2 Increase or decrease investment in human capital?
- 1.1.3 Affect endowments of human and physical capital?
- 1.1.4 Influence the labour participation rate?
- 1.1.5 Improve the functioning of capital markets?
- 1.1.6 Price levels and stability
- 1.1.7 Affect the level or composition of aggregate demand?
- 1.1.8 Create upward or downward pressures on prices?
- 1.1.9 Influence the costs of production factors?

1.2 Effects on public authority budgets
- 1.2.1 Require public expenditure or entail future budget commitments?
- 1.2.2 Affect tax rates?
- 1.2.3 Impact on the public sector budget balance or the quantity of government debt?

1.3 Human capital formation and employment
- 1.3.1 Affect education services?
- 1.3.2 Affect the training of workers?
- 1.3.3 Change the level of employment?
- 1.3.4 Change the composition of employment?
- 1.3.5 Affect working conditions?
- 1.3.6 Affect unemployment?

1.4 Economic cohesion
- 1.4.1 Change the geographic distribution of economic activity?
- 1.4.2 Change the geographic distribution of infrastructure?
- 1.4.3 Impact on economic integration?
- 1.4.4 Affect rural and slow growth areas?
- 1.4.5 Have effects on the black economy?

1.5 Innovation
- 1.5.1 Increase or decrease R&D investments?
- 1.5.2 Lead to technical innovation?
- 1.5.3 Lead to organisational innovation?
- 1.5.4 Lead to institutional innovation?

1.6 International performance
- 1.6.1 Influence international trade or price competition?
- 1.6.2 Increase or decrease foreign direct investments?
- 1.6.3 Potentially affect international agreements and alliances?
- 1.6.4 Have an impact on enlargement countries?
- 1.6.5 Have an impact on developing countries?

1.7 Market structure
- 1.7.1 Increase or decrease single market trade and cross-border investments?
- 1.7.2 Improve price convergence across single market?
- 1.7.3 Change the sectoral distribution of economic activity?
- 1.7.4 Result in increased market concentration?
- 1.7.5 Have specific impacts on SMEs?
- 1.7.6 Change the degree of market competition?
- 1.7.7 Result in or facilitate firm entry or exit?

1.8 Microeconomic effects on enterprises, non-profit organisations etc.
- 1.8.1 Increase or decrease sectoral fixed-capital investment (e.g. in production plant, buildings, technology or equipment), including in clean technologies?
- 1.8.2 Increase or reduce operating costs (e.g. for raw materials, labour, other recurring production costs, licence fees, periodic inspections)?
- 1.8.3 Increase or reduce administration costs (e.g. on formalities and paperwork)?
- 1.8.4 Influence the cost or availability of firm financing?
- 1.8.5 Encourage or discourage innovation and R&D?
- 1.8.6 Affect output and turnover?
- 1.8.7 Change the input and output prices?

1.9 Effects on households
- 1.9.1 Affect household income and wages?
- 1.9.2 Impact on consumer prices and financing costs?
- 1.9.3 Increase or decrease consumer purchasing power and choice?
- 1.9.4 Affect consumer protection?
- 1.9.5 Affect pensions or asset holding?

1.10 Global partnership
- 1.10.1 Imports of fair trade goods
- 1.10.2 EU financing to sustainable development as part of development aid
- 1.10.3 Material flows (waste/resources) to and from developing countries

2 Environment

2.1 Air, water soil or climate
- 2.1.1 Emissions of acidifying, eutrophying, photochemical or harmful air pollutants?
- 2.1.2 Unpleasant smells and odours?
- 2.1.3 Decrease or increase the quality or quantity of surface and groundwater?
- 2.1.4 Raise or lower the quality of waters in coastal and marine areas (e.g. through discharges of sewage, nutrients, oil, heavy metals and other pollutants)?
- 2.1.5 Affect drinking water resources?
- 2.1.6 Affect acidification, soil erosion rates, contamination salinity of soil?
- 2.1.7 Lead to loss of available soil (e.g. through building or construction works) or increase the amount of usable soil (e.g. through land decontamination)?
- 2.1.8 Ozone-depleting substances (CFCs, HCFCs)?
- 2.1.9 Change emissions of greenhouse gases (e.g. carbon dioxide, methane etc)?

2.2 Renewable or non-renewable resources
- 2.2.1 Use renewable resources more quickly than they can regenerate?
- 2.2.2 Reduce or increase use of non-renewable resources?

2.3 Bio-diversity, flora, fauna
- 2.3.1 Reduce the number of species in any area (i.e. reduce biological diversity) or increase the range of species (e.g. by promoting conservation)?
- 2.3.2 Affect protected or endangered species or their habitats or ecologically sensitive areas?
- 2.3.3 Affect the scenic value of protected landscape?
- 2.3.4 Split the landscape into smaller areas?
- 2.3.5 Affect migration routes, ecological corridors, or buffer zones?

2.4 Land use
- 2.4.1 Bring new areas of land (‘greenfields’) into use for the first time?
- 2.4.2 Affect land designated as sensitive for ecological reasons?
- 2.4.3 Change land use?

2.5 Natural and Cultural heritage
- 2.5.1 Affect natural structures or sensitive areas, mountains, coastal areas or islands?
- 2.5.2 Affect cultural landscapes or heritage such as protected areas, buildings, sites, monuments or features?

2.6 Waste production/generation or recycling
- 2.6.1 Affect waste production (solid, urban, agricultural, industrial, mining, radioactive or toxic waste)?
- 2.6.2 Affect how waste is treated, disposed of, or recycled?

2.7 Human safety or health
- 2.7.1 Increase or decrease the likelihood of health risks due to substances harmful to the environment?
- 2.7.2 Change the amount of noise and health damage caused by noise?

- 2.7.3 Improve or reduce air quality in populated areas?

2.8 The likelihood or scale of environmental risks

- 2.8.1 Affect the likelihood or prevention of fire, explosions, breakdowns, accidents and accidental emissions?
- 2.8.2 Increase or decrease the risk of unauthorised or unintentional dissemination of environmentally alien or genetically modified organisms?
- 2.8.3 Increase or decrease the likelihood of natural disasters?

2.9 Mobility (transport modes), or the use of energy

- 2.9.1 Increase or decrease consumption of energy and production of heat?
- 2.9.2 Will the proposed policy increase or decrease the demand for transport (passenger or freight), or influence its modal split?
- 2.9.3 Increase or decrease vehicle emissions?

3 Social / cultural

3.1 Social Cohesion

- 3.1.1 Affect social integration
- 3.1.2 Affect the extent of extreme and/or persistent poverty?
- 3.1.3 Affect the risks of poverty or social exclusion?
- 3.1.4 Affect geographical social cohesion?
- 3.1.5 Affect long term unemployment?
- 3.1.6 Affect the accessibility of services of general interest?

3.2 Employment Quality

- 3.2.1 Affect the organisation of labour markets (public employment services; job matching etc.)?
- 3.2.2 Facilitate and protect labour transitions (in and out of education, inactivity or caring, training, and retirement)?
- 3.2.3 Affect occupational health and safety arrangements?
- 3.2.4 Affect workers rights and their effective recognition?
- 3.2.5 Affect the opportunity of employment and integration through employment?
- 3.2.6 Affect the balance between professional and personal life?
- 3.2.7 Affect the prospects of worker in industries undergoing restructuring?
- 3.2.8 Affect industrial relations (bargaining, strikes etc)?

3.3 Public health

- 3.3.1 Affect the health of the population including life expectancy, mortality, morbidity? Have special effects on risk groups based on age, gender, social groups, migrants, regions.
- 3.3.2 Impact on the socio-economic environment including working environment, income, education, occupation, nutrition? Focus on risk groups, consider for example age, gender, social groups, migrants, regions.
- 3.3.3 Affect the health of individuals? Focus on risk groups.
- 3.3.4 Change life style related determinants of health such as use of tobacco, alcohol, physical activity etc?

3.4 Health systems and security

- 3.4.1 Impact on health services including quality of health services, access to health services, health professionals education and mobility
- 3.4.2 Have a cross-border dimension: provision of services, referrals, co-operation in border regions?
- 3.4.3 Affect the financing of the health system?
- 3.4.4 Change the organisations within the health system?
- 3.4.5 Affect the potential for bioterrorism?

3.5 Social Protection and Social Services

- 3.5.1 Affect levels of social protection (risks and rights covered)?
- 3.5.2 Affect accessibility?
- 3.5.3 Have effects on the demands for social protection and services (e.g. demography, reproductive behaviour, etc)?
- 3.5.4 Have effects on the inputs of social protection and services (costs, long term sustainability, professional and institutional arrangements)?

3.6 Consumer interests

- 3.6.1 Affect consumer safety (including food safety)?
- 3.6.2 Change consumer information?

- 3.6.3 Affect consumer choice?
- 3.6.4 Impact on animal health and welfare?
- 3.6.5 Change the effectiveness of public controls?

3.7 Education

- 3.7.1 Affect the provision of education?
- 3.7.2 Encourage or discourage participation in education?
- 3.7.3 Affect educational achievement?
- 3.7.4 Life-long learning

3.8 Social Capital

- 3.8.1 Reinforce or marginalise local, regional, national or community identities?
- 3.8.2 Encourage or discourage civic behaviour, active citizenry, and volunteering?

3.9 Liveable communities

- 3.9.1 Affect housing quality or provision?
- 3.9.2 Affect infrastructure?
- 3.9.3 Affect services?
- 3.9.4 Improve or worsen geographical access and transport opportunities?

3.10 Equality of opportunity and entitlement

- 3.10.1 Lead to the inclusion or exclusion of individuals or groups?
- 3.10.2 Affect poverty and dependence?
- 3.10.3 Impact on human rights?
- 3.10.4 Have an ethical dimension?
- 3.10.5 Impact on xenophobia?
- 3.10.6 Affect racial or ethnic discrimination?
- 3.10.7 Affect sexual discrimination?
- 3.10.8 Affect others forms of discrimination based on religion, language, genetic characteristics, opinions, birth, handicap, (see article 21 of the Charter of Fundamental Rights)
- 3.10.9 Does the initiative provide or result in reasonable arrangements for improving accessibility for handicapped people?
- 3.10.10 Impact on child labour?

3.11 Culture

- 3.11.1 Affect cultural diversity?
- 3.11.2 Affect cultural heritage?
- 3.11.3 Impact on sports, arts and entertainment?

3.12 International co-operation

- 3.12.1 Affect overseas development assistance?
- 3.12.2 Impact on international relations and trade?
- 3.12.3 Affect public health in non-EU countries?
- 3.12.4 Impact on education in non-EU countries?
- 3.12.5 Affect liveable communities in non-EU countries?
- 3.12.6 Impact on equality of opportunity and entitlement in non-EU countries?
- 3.12.7 Affect culture in non-EU countries?
- 3.12.8 Affect governance and participation in non-EU countries?
- 3.12.9 Affect international migration flows

3.13 Governance and participation

- 3.13.1 Lead to more or less stakeholder involvement?
- 3.13.2 Affect social dialogue (resources, risks, opportunities and rules at local, national and european level)?
- 3.13.3 Affect transparency and the balance of powers?
- 3.13.4 Improve accountability and democracy?
- 3.13.5 Affect the empowerment of disadvantaged groups?
- 3.13.6 Affect the empowerment of local and regional actors?
- 3.13.7 Impact on citizenship and citizens rights?
- 3.13.8 Influence civil dialogue (resources, risks, opportunities and rules at local, national and European level)?
- 3.13.9 Policy coherence (including harmful subsidies, share of council conclusions that are followed up, number of infringement procedures and cases, and share of policies, programmes and plans for which an impact assessment has been decided and undertaken
- 3.13.10 Institutional learning/knowledge

3.14 Fundamental human rights

3.14.1 Support and sustain the principles of fundamental human rights?

3.15 Security, crime or terrorism

3.15.1 Improve security arrangements?

3.15.2 Facilitate or hinder crime or terrorism?

3.16 Ageing of society and pensions

3.16.1 Pensions adequacy

3.16.2 Demographic changes

3.16.3 Financial stability

3.16.4 Exit age

Annex 4 Detailed list of principles for sustainable development in Europe

The following pages contain the detailed list of impacts that are considered relevant in the context of sustainable development in Europe. The list is constructed from the Sustainable Development Indicators (indicators in font 'times new roman') and the Structural Indicators (in font 'arial'). The main categories, which are also listed in Table 3.4 on page 35 are listed in **bold** text.

SDI 1. Economic development
1. GDP per capita in Purchasing Power Standards
2. Total consumption expenditure as % of GDP
3. Net disposable income received as a % of GDP
4. Inflation rate
5. National savings as % of GDP, by source of funds
6. Total R&D expenditure as a % of GDP
7. Labour productivity per hour worked
8. Unit labour cost growth, for total and industry
9. Life-long learning
10. Total employment growth
11. Total unemployment rate, by gender, by age group, and by highest level of education attained
12. Labour productivity per person employed

SI 1. Employment
1. Employment and productivity development in the EU
2. Employment rate – total and by gender
3. Employment rate of older workers – total and by gender
4. Long-term unemployment rate – total and by gender

SI 2. Innovation and Research
1. GERD (Gross domestic expenditure on R&D)
2. Evolution of R&D spending
3. Youth educational attainment level – total and by gender
4. Evolution of youth educational attainment level

SI 3. Economic Reform
1. Comparative price levels
2. Business investment
3. Evolution of business investment

SDI 2. Poverty and social exclusion
1. At-risk-of-poverty rate, by gender, by age group and by highest level of education attained
2. Relative at-risk-of-poverty gap
3. Inequality of income distribution (Income quintile share ratio)
4. Poverty mobility (i.e. probability to enter or exit poverty)
5. Gender pay gap in unadjusted form
6. Very long-term unemployment rate
7. People living in jobless households, by age group
8. At risk-of-poverty rate after social transfers by most frequent activity
9. Persons with low educational attainment, by age group
10. Adequacy of housing conditions

SI 4. Social Cohesion
1. At-risk-of-poverty rate after social transfers – total and by gender
2. Evolution of the at risk of poverty rate

3. Dispersion of regional employment rates – total and by gender

SDI 3. Ageing society

1. At-risk-of-poverty rate for persons aged 65 years and over
2. Projected theoretical replacement ratio
3. Total fertility rate
4. Net inwards migration, by main age groups
5. Average exit age from the labour market: Total
6. Current and projected public (and private) pensions expenditure as % GDP
7. Current and projected public expenditure on care for the elderly as % of GDP

SDI 4. Public health

1. Health care expenditure as % of GDP
2. Cancer incidence rate, by gender and by type
3. Suicide death rate, by gender and by age group
4. Percentage of present smokers, by gender and by age group
5. Work with high level of job strain/stress
6. Accidents at work - Serious accidents
7. salmonellosis incidence rate
8. Occupational diseases caused by certain chemical agents
9. Proportion of population living in households suffering from noise and from pollution

SDI 5. Climate change and energy

1. CO_2 intensity of energy consumption
2. Losses caused by extreme weather conditions (insurance payouts)
3. CO_2 removed by sinks
4. Share of renewable energy (including indivcative targets), by source
5. Combined heat and power generation as % of gross electricity generation
6. Energy intensity of manufacturing industry
7. Energy tax revenue vs. GDP at constant prices
8. Consumption of biofuels, as a % of total fuel consumption in transport
9. High-level radioactive waste and spent nuclear fuel awaiting permanent disposal
10. Energy intensity of the economy

SDI 6. Production and consumption patterns

1. Components of Domestic Material Consumption
2. Domestic Material Consumption, by material
3. Municipal waste treatment, by type of treatment method
4. Generation of hazardous waste, by economic activities
5. Household number and size
6. Meat supplies per capita
7. Share of consumption of products with EU ‘flower’ or similar labelled products
8. Livestock density index
9. Share of organic farming
10. Use of selected pesticides
11.Ethical financing
12. Eco-label awards, by country and by product group

SDI 7. Management of natural resources

1. Change in status of threatened and/or protected species
2. Effective fishing capacity vs. quotas, by specific fisheries
2a Size of fishing fleet
3. Structural support to fisheries and % allocated to promote env. friendly fishing practices
4. Population connected to wastewater treatment systems
5. BOD loading of rivers
6. Index of pesticide risk to aquatic environment
7. Total area at risk of soil erosion
8. Total area of soil contamination
9. Percentage of forest trees damaged by defoliation

SDI 8. Transport

1. Volume of freight transport vs. GDP at constant price
2. Energy consumption by transport mode
3. Access to public transport
4. Fragmentation of habitats due to Transport
5. People killed in road accidents, by age group

SDI 9. Good governance

1. Share of key legal acts for which an impact assessment has been decided and undertaken
2. Transposition of Community law, by policy area
3. Voter turnout in EU parliamentary elections, by gender, by age group and by highest level of education attained

SDI 10. Global partnership

1. Total EU imports from developing countries by income group
2. Total EU imports from developing countries by group of products
3. Total EU financing for development, by type
4. ODA and FDI to developing countries, by income group and geographical area
5. Share of untied ODA in total bilateral ODA commitments
6. ODA per capita, in EU donors vs. in recipient countries
7. Contribution of the Clean Development Mechanism (CDM) to GHG emission reductions in developing countries

7a. CO_2 emissions per capita in the EU vs. in developing countries

8. Environment-adjusted EU imports of materials from developing countries, by group of products, income group and geographical area