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

2.1 Introduction

In which settings should the combination of CBA and MCA be applied? Because evaluation is a very broad research field, it may be clear to the reader that this thesis is not developing an 'all purpose' evaluation approach. Evaluation literature indicates that before making claims about the applicability of a method, one should try to be specific about both the setting in which a method is to be used and the purpose of the evaluation (Rossi et al., 1999, p. 94).[1] This chapter will sketch the challenges to an accountability-enhancing project evaluation tool that has to be used in a sustainability context.
In order to specify the characteristics required by the evaluation tool, section 2.2 will discuss the purpose of the evaluation and section 2.3 will treat what are, generally speaking, the types of projects to be evaluated. Section 2.4 synthesises the essential aspects necessary for the evaluation tool, while section 2.5 provides a brief summary.

2.2 The judgement perspective in evaluation

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

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

The choice of perspective goes hand-in-hand with the intended purpose of the evaluation and with the types of questions being posed.
The first perspective is the *judgement* perspective. The aim of evaluation from this perspective is judging merit, worth or value. The evaluation is focussed on measuring results or efficiency.
The second perspective is the *enlightenment* perspective. Evaluation here aims at generating knowledge. The evaluation tries to acquire a more profound understanding in some specific area or field; its aim is generalised knowledge and ideas, often drawing from a body of existing evaluations. Enlightenment is, however, not always the primary aim, but it may simply occur as a by-product or accidental spin-off of evaluations done from the other two perspectives (Weiss, 1988 and 1990, p. 22).
The third perspective is the *development* perspective, which has the role of facilitating improvement of existing activities and organisations. Evaluation literature in general regards this facilitation in the light of providing (evaluative) help to strengthen organisations or institutions.

[1] Compare Fischhoff, 1991 about processes that stimulate exaggerations of applicability.
[2] Although the following terminology is more commonly used (respectively): the accountability, the knowledge and the development perspective. In this thesis judgement will be used, because the word accountability is already used here in a somewhat different, although related, meaning. The word enlightenment will be used (Weiss, 1988, p.176) because knowledge seems to be the purpose of every type of evaluation, while enlightenment has more the connotation of "*just* for knowledge's or understanding's sake".

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

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

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

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

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

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

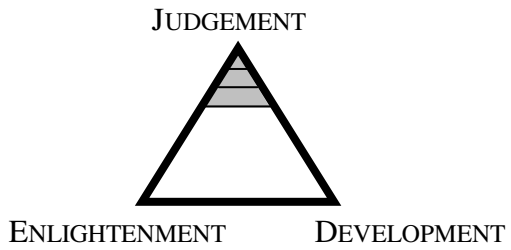

Figure 2.1: Three perspectives in the evaluation process

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

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

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

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

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

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

2.3 Type of projects

2.3.1 General characteristics

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

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

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

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

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

Ad d)

[5] The thesis examines evaluation experience and theory in order to derive a new improved methodology, which is drawn foremost from the knowledge perspective.
[6] Understanding logically also implies understanding the limitations to the evaluation.
[7] Although – other than evaluators in general – economists seem rather flexible in defining policy options as project alternatives.

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

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

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

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

[8] The author's experience in project evaluation is largely in the field of spatial economics: land-use changes, spatial redistribution of economic activity, infrastructure development, and sustainable corporate performance.
[9] Not least because one can build up experience over several projects in using a standardised evaluation approach. Compare Farrow and Toman (1999).

2.3.2 Evaluating on sustainability

2.3.2.1 The concept of sustainability

Although public concern about environmental damage caused by economic activity has existed for a long time (Edwards-Jones et al., 2000, p. 26), concern has progressively heightened in recent decades regarding the Earth's capacity to sustain economic activity. When the UN established the World Commission on Environment and Development in the 1980s, ecological impacts resulting from economic issues have since been seen in a broader sustainability context.[10] Many governments, NGOs, citizens, and businesses have committed themselves to achieving greater sustainable development[11], and it has become "a high profile objective" (Pearce, 1998a, p. 69). But what exactly is sustainable development?

Defining sustainability and sustainable development is often seen as a difficult task, as in many contexts: 'a more sustainable world' has come to mean 'a better world', which can have many (normative) connotations and dimensions, from the economic and environmental to the social, cultural, historical, and ethical vantage points. The concept can sometimes be very vague (Sillanpää, 1998) or "somewhat obscure, often unique to the speaker" (Sarkar, 1997), which does not help to facilitate an evaluation.[12]
However, for our purposes, the essentials of sustainability will become clear enough. Furthermore, in discussing these essentials and the possibilities of different evaluation tools in handling these essentials, the reasons for the vagueness – that may occur sometimes – will be explained.[13] The sections below will therefore define sustainability, or perhaps better, give it operational content, by identifying key elements of the concept (compare Pannell and Schilizzi, 1999, who follow a similar approach). These key elements together comprise 'the context of sustainability' for project evaluation.

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

[10] See the website of the UN Commission on Sustainable Development: www.un.org/esa/sustdev/csd/csd.htm.
[11] See for instance the Millenium goals of the United Nations (www.un.org/milleniumgoals), or the Council of the European Union (2006), or WBCSD, 2000.
[12] "Sustainability is at once extremely important and practically useless" (Pannell and Schlizzi, 1999, p. 65). Bromley (2001) sees sustainability as both "a fine idea" as it brings the fate of future generations to our attention and as "a hopeless concept", because it lacks operational content. Compare Little and Mirlees (1994, p. 213): "Sustainability has come to be used in recent years in connection with projects. This is more of a buzzword … than a genuine concept. It has no merit. Whether a project is sustainable … has nothing to do with whether it is desirable."
[13] Running ahead a little: these reasons seem to lie in stakeholder involvement identifying mere 'concerns' and a lack of causal analysis preventing sustainability in a specific evaluation setting.
[14] See for instance Feyerabend, 1975 and compare Sheldrake, 1997 for a reading on several interesting blind spots in science.

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

2.3.2.2 A long-term view (f1)

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

> *"Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs."*

The Brundtland definition emphasises the future-oriented aspect of sustainability. There seems to be little disagreement between science and politics in this respect. Following the explanation of Pearce (1998a, p. 69), one may say that sustainable development can 'endure', it can 'last', which should be possible to 'keep in being'.[17] In business contexts a long-term view is sometimes seen as more than 10-15 years forward. Although one need not rule out this possibility in any specific evaluation context, many authors have much longer time periods of analysis in mind. For instance, Pearce (1998a) argues that 100 years might be taken as a minimum (!) time horizon.
Evaluation should (be able to) show long-term future impacts, which can be understood as strengthening the characteristics of point 'a' in section 2.3.1 above (ex ante evaluation, future oriented); and in point 'b' (impacts carry on over time).

2.3.2.3 A global (and local) perspective (f2)

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

[15] Compare Meppem and Bourke (1999) for an analysis of difficulties in overcoming paradigmatic fixations in the 'communicative turn' that is needed in sustainability matters. Compare further, for instance, the EU difficulties in making progress in its Sustainable Development Strategy (SDS) and especially in balancing sustainability demands with the demands from the Lisbon agenda for growth and jobs (Council of the European Union, 2006).
[16] See Munasinghe et al. (2001).
[17] Compare Spash 2000; Spash sees two types of 'economic' definitions of sustainability: Ends-based definitions (e.g., non-declining consumption or utility) and means-based definitions (e.g., a non-declining stock of capital for "producing" well being).

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

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

2.3.2.4a Three elements: Triple E

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

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

[18] In the next sub-section a third reason will be given.
[19] It may be surprising to some readers that analysing at too low a level may lead to *over*estimation of impacts (apart from obvious chances for underestimation) as feedbacks may not be traced (Lakshmanan et al., 2001).
[20] See for instance Vogtländer (2002, p. 25): "Only when [emission] norms are set for the whole world, will problems like 'export of environmental problems' and 'levelling the commercial playing field' be resolved definitely."
[21] However, at below global levels the matching of Triple E scales will be a point of concern (Musters et al., 1998; Boisvert et al., 1998).
[22] Compare Rischard (2002) for an analysis showing the need for improving the currently ineffective and slow global-problem solving on pressing global issues.

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

The notion that sustainability is about finding a good mix of economic development, environmental quality and social quality has widespread approval (Daly and Cobb, 1989; Elkington, 1997; Munasinghe, 2001; GRI, 2002; Pope et al., 2004 among others). There is, however, wide difference in the emphases given, and the social dimension in particular is treated quite differently by various authors. The interpretation of this thesis which, following Brundtland, focuses on extreme poverty and basic needs may show a 'bias', that some regard as common to the economist's perspective[23]. However, this interpretation can also be seen as choosing 'minimum' areas of concern whose importance seems undisputed.
The sustainability terminology, which has gained widest acceptance most quickly among business managers and consultants is Elkington's 'Triple P' bottomline of "people, planet and profit" (Elkington, 1997). Elkington mainly addresses businesses, and argues that if corporations want to become sustainable, they should improve performance in the three areas of people, planet and profit. However, it seems that 'people' has a broad connotation, and the interpretation of 'profit' may be somewhat narrow. When we discuss here the three elements of sustainability as noted above, we will sometimes use the term 'Triple E' – paraphrasing Elkington – because it is conveniently short. To reiterate, Triple E stands for: 1) Economic development, 2) Extreme poverty[24] and 3) Environmental degradation. Evaluation in a sustainability context has to show Triple E impacts.

2.3.2.4b Triple E relations: separate concerns or closely related concerns?
The introduction of separate Triple E elements into a sustainability evaluation is tricky, however, as confusion is likely, unless the interrelation between the Triple E elements is clarified.[25] In the sustainability context there is obviously an interest in economic development that can be sustained over generations, but how do extreme poverty and environmental degradation relate to it? Are they simply 'added concerns' or are they strongly dependent upon each other?
Figure 2.2 illustrates the basic options. The Triple E elements could be unconnected (I) and could therefore warrant interest in their own right. Conversely, they can be very strongly connected (IV): an interest in one necessarily leads to interest in the others. All three elements can be somewhat related (III), or the situation can be different for each link (II).

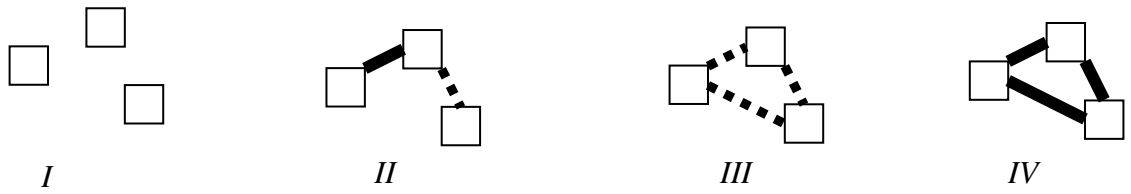

Figure 2.2: *Four basic options of connections between the Triple E elements*

[23] As Serageldin (1993) argues. See Munasinghe (1993), Pezzey (1992) and Hanley (2000) for a number of economists' views on sustainability. See Cernea (1993) and Rees (1993) for examples respectively of a sociologist and ecologist's perspective.
[24] The focus on the extreme poor rather than people in general provides a clear link with Rawlsian justice ideas (see for instance Langhelle, 2000).
[25] This situation mirrors the confusion around the Triple P approach to business performance.

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

2.3.2.4c Triple E – partly elements in their own right

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

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

2.3.2.4d The three relations between the Triple E elements

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

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

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

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

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

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

Again, causality here as such need not be too complex, as can be seen in many examples of the 'tragedy of the commons' type.[29] The cumulative size of many small-scale economic choices becomes too large and leads to over-exploitation of environmental resources (Ophuls and Boyan, 1992; Gordon, 1998; Schultz and Holbrook, 1999). For this type of problem adopting a higher spatial or organisational level of analysis or using a longer time perspective is enough to clarify the simple causal relationship.[30]
However, causality can be complex too, if one for instance addresses the problem of greenhouse gases (Van den Bergh, 2004).[31] This causality may then be uncovered by analysing developments over longer periods of time, by examining impacts on higher spatial levels, by finding the mechanisms of environmental change, or by a combination of these, and in the end, even this may not be sufficient to depict the relations.[32]

The relationship between *Extreme poverty and environmental degradation* is often very sharp and clear on a local or regional scale. Extreme poverty, or "short-term survival pressure", is often the basis for strong population growth, while strong population growth leads to pressure on the environment. The literature mentions, for instance, heightened pressure nowadays towards deforestation for energy, with general disregard for the longer term consequences for water regulation. Wood shortages also lead to the burning of dung for fuel, which, according to Hart, is "one of the greatest – and least well known – environmental hazards in the world today" (Hart, 1997, p. 69).
For evaluation purposes we can emphasise that the first situation here rather resembles the former link where a more distant perspective may clarify the relation; whereas the burning dung for fuel case shows that extra knowledge about problems of differing complexity may be involved. Both examples also show that, although causality may be clear, this has little bearing on the availability or viability of alternative courses of action.

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

[29] See for instance Lundgren and Burnet (1999) and references therein for an introduction and some recent debate – included is also a short commentary by the author of the influential 1968 article 'The Tragedy of the Commons': Garrett Hardin. See Gordon (1998) for a brief discussion of an example of the tragedy of the commons type: the example of the North American codfish.
[30] Formulated the other way around: causality may be blurred when a low level and a short-term view is adopted in evaluation.
[31] Compare Aguilera-Klink et al. (2000) for the subtle socio-economic complexities in defining and experiencing water scarcity due to various economic developments.
[32] A complete account of these possibilities is not appropriate to this thesis, but it may be fruitful to specify a number of important problems. According to Ayres (2001), the most immediate problems where the environmental limits hamper economic development are the supply of water (both for irrigation and drinking in cities in several regions in Asia) and wood (for fuel and charcoal production in Africa and Asia). The shortage of wood leads to uncontrolled deforestation, which in turn leads to soil erosion and increasingly catastrophic floods. Furthermore, Ayres sees a major problem in the reduction of toxic-waste assimilative capacity of the earth (due to topsoil loss and degradation and loss of biodiversity) while the demand for this 'service' of the environment is rising. The last problem posing binding limits is climate warming and its associated dangers.
[33] The subsistence economy is not necessarily a situation with extreme poverty. In this distinction the market economy consists of developed and emerging economies. The survival economy is the traditional, village-based way of life in rural areas of developing countries. According to Hart, about one-sixth of the world's population (one billion people) live in developed countries of the market economy. They account for more than 75% of the

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

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

2.3.2.4e Triple E relations: facts and values

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

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

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

2.3.2.4f Conclusions about Triple E relations

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

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

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

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

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

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

there is no relationship or the question of their relationship is not seen as relevant (see figure 2.2 - I).
Second, some of the links between Triple E elements are really rather direct and simple in terms of the mechanism involved.
- In some cases the links are direct and simple, and visible or clear to everyone involved.
- In other cases these direct and simple relations are not always visible on a local scale or low spatial level,[37] but may rather easily become clear when a (somewhat distant) perspective is adopted. This state of affairs strengthens the importance of evaluation at different spatial/organisational levels, and of the inclusion of the global level. Likewise in another set of cases these direct and simple relations are not always visible from a short-term perspective, but may rather easily become clear when a long time frame is considered. Figure 2.3 illustrates this situation. With this type of link one should be aware that, although clear links may be identified at high spatial levels or from longer time perspectives, this fact as such does not imply the existence of an agreement on possible mechanisms of change. This fact underlines the importance of 'economic realism' on a micro level, as the 'distant' understanding has to relate with the real-life mechanisms involved. It has to relate with the alternatives for actions that are available or absent at the micro level.

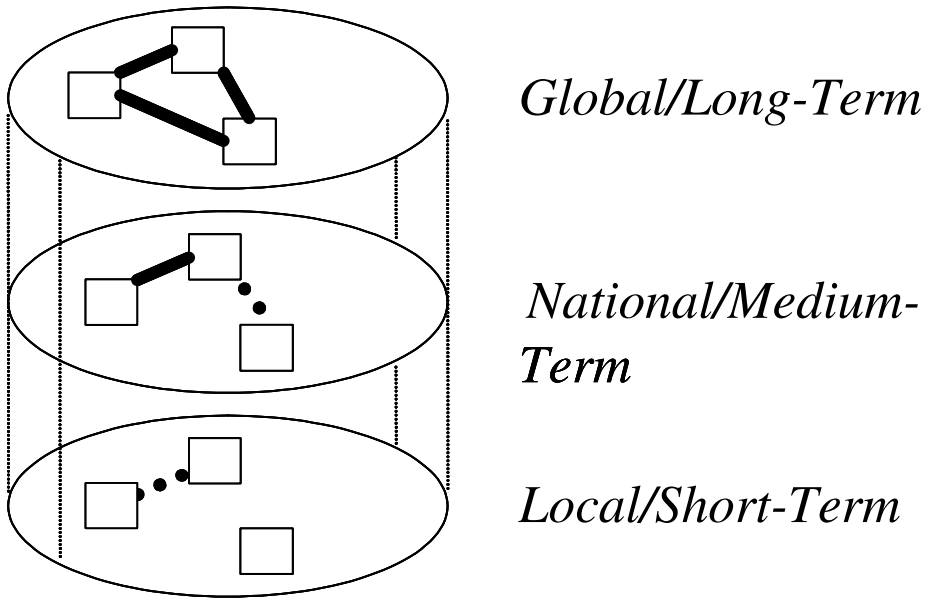

Figure 2.3: Illustration of how a higher spatial level and a longer term perspective may give a clearer view on the interrelatedness of Triple E elements.

Third, some of the links are complex and uncertain. Limited knowledge may indicate important risks in the interrelatedness of Triple E elements, or positive feedback relations, but the relationship remains complex and uncertain. This type of relation may occur in both the socio-economic and the ecological-environmental systems[38.] (Figure 2.2. III).

[37] Although of course at regional or local levels we may see *some* strong links; and when we think of traffic congestion or noise, the links may be especially strong at the local level.
[38] Several authors believe that system thinking may be helpful as tools of analysis (Clayton and Radcliffe, 1996). System thinking would stress that social-economic and environmental-ecological systems are complex adaptive systems that are living and learning. These systems have numerous thresholds, feed-back structures, capacities to accrue and interpret information (Schütz, 2000). Over and above that, the interrelations between the several co-

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

2.3.2.5 Stakeholder participation (f4)

2.3.2.5a Two reasons for stakeholder participation

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

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

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

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

2.3.2.5b Various groups of stakeholders

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

---

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

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

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

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

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

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

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

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

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

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

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

2.3.2.5c Minimum requirements for stakeholder participation

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

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

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

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

2.3.2.5d Conclusion

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

2.4 Synthesis of the essential characteristics

2.4.1 Four key items

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

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

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

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

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

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

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

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

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

content to the type of judgement being aimed at, as it is additional to the most relevant decision-making level(s). The global level can also be interpreted as an increasingly common impact level, as projects tend to have more impacts across the globe.
III Over (a long) time – Here the element 'long-term view' strengthens the 'time dimension' element.
IV Which increases understanding of all (also about limitations) – The evaluation should increase the understanding of all parties concerned.
Not only should experts or decision-makers have increased understanding, but broad ranges of stakeholders should too. The ideal project evaluation tool will allow for fruitful stakeholder participation. The elements 'accountable and open to stakeholders' and 'stakeholder participation' strengthen each other and lead to this extra challenge for the evaluation tool. The element 'large and complex: formal project evaluation' specifies the general background. The increased understanding relates to the limitations as well. Evaluation in the context of accounting for sustainability will have various limitations and uncertainties that should be clarified. The elements 'ex ante' and sometimes 'complexly related systems' – provide a general warning about possibilities and limitations: ex ante assessments of complex systems are performed about which there is limited knowledge. This general warning is greatly strengthened by the stakeholder participation elements.

Accountability and sustainability = sustainability (for short)

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

2.4.2 More on limitations, judgement and uncertainty

2.4.2.1 Limitations and judgement

Four key items were identified in the discussion above that are important for a project evaluation tool used in a sustainability context of decision-making. In order to increase the understanding of the key items, we notice that, generally speaking, these four items can relate to two central concerns: limitations and judgement.
Every evaluation technique and every estimate of impacts has its limitations, but as seen above, evaluation in the accountable-sustainability context may have more than the average number of limitations. These should be (actively) clarified, as it is as least as important to know what is neglected as it is to know what is taken into account (Fischhoff, 1977).
Relating this discussion back to the three evaluation perspectives – in section 2.2 – it may be clear then that the evaluation context at hand is still judgement-oriented, but that already at this stage there may be some doubt about whether the upper parts of the shaded triangle of figure 2.1 can be attained.

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

2.4.2.2 Uncertainty

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

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

The terminology largely follows Vercelli (1999).[47] First, there is certain knowledge, about which little needs to be said, except that the evaluation process may *establish* what is certain,[48] and that evaluation will involve a *structuring* of certain knowledge.[49]
Soft uncertainty is defined as the area of uncertainty in a familiar world (theoretically, roulette games or horse races function as examples). The area of soft uncertainty partly consists of the area which in formal decision theory is called risk: uncertainty with known probabilities. However, it may also include risk in a more real life setting. Risk in a real life setting will be frequently somewhat more loosely defined and may involve several probabilities.[50] In the sustainability context a complete empirical risk assessment may typically require several steps, which may necessitate estimates of probability.[51]
However, lack of certainty not only relates to empirical matters as such, but also to *beliefs* about empirical matters (Vicelli, 1999, p. 238). Different stakeholders may have different beliefs; experts and non-experts may differ (Pollak, 1998), but experts may also disagree among themselves (OECD, 1983). Because of the importance of beliefs about empirical matters, it seems useful to define risk in a more general and less formal way by making it refer to a situation in which there are either known probabilities or where there is a 'familiar world'. Familiar is defined here as a situation where there is substantial knowledge about the mechanisms involved.
Hard uncertainty thus refers to knowledge areas where there are either unknown probabilities or in which the mechanisms involved are unfamiliar.[52] Still, in this uncertainty area, there is awareness. Decision-makers and stakeholders know that they have incomplete knowledge.
The final category that Vercelli distinguishes is unawareness of ignorance. This category seems to be especially useful if the evaluation process is considered. Different stakeholders may have different levels of unawareness and may be unaware of different areas, and

[46] Compare a quote of Funtowicz and Ravetz in 1991 (quoted in Froger and Munda, 1998, p. 173): " the facts are uncertain, values in dispute, stakes high, and decision urgent. In this way it is 'soft' scientific information which serves as inputs to the 'hard' policy decisions on many important environmental issues."
[47] However, our definition of soft uncertainty is less formal than Vercelli's.
[48] It need not be obvious at the start, as evaluation may bring awareness of this certain knowledge.
[49] It is not a mere collection of certainties; they are structured, which is the reason why in figure 2.5 we have given this area some structure too.
[50] E.g., Covello and Merkhofer (1993) define risk as a characteristic of a situation or action with more than one possible outcome; it is unknown which outcome will occur and at least one outcome is undesired.
[51] Covello and Merkhofer (1993) in this respect make use of the 'risk-chain', which analyses first the risk source release processes (e.g., the likelihood of radioactivity release from a nuclear power plant); second, the exposure processes (how many people live in the vicinity of the plant); and third, the consequence processes (what effects may be experienced by exposed individuals).
[52] Naturally there may be a grey area between these categories. For instance, risk defined in a real life way (following Covello and Merkhofer, 1993) may refer to both soft uncertainty or hard uncertainty, mostly depending on the familiarity of the risks involved.

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

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

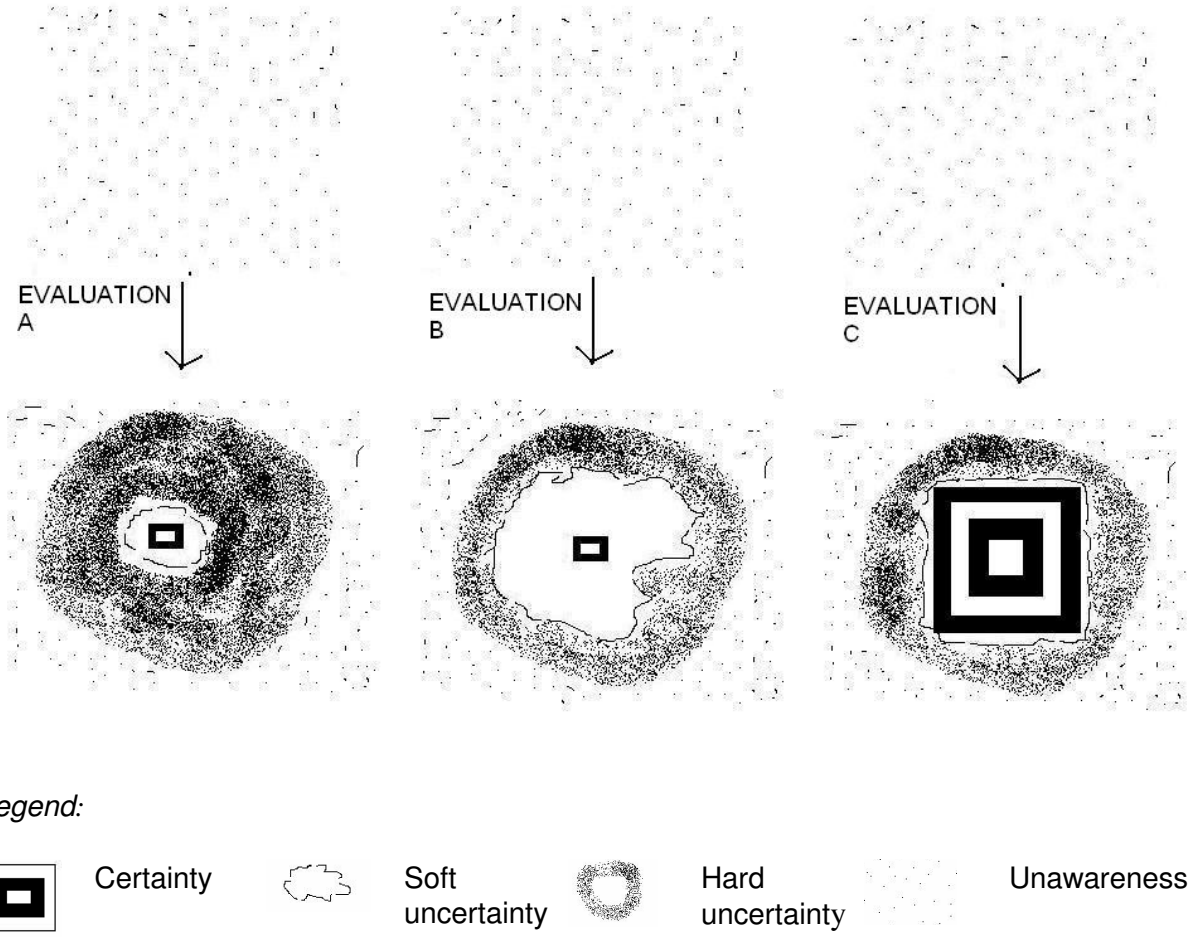

Figure 2.5: Illustration of how, from an initial state of unawareness, three different evaluations may result in various divisions of knowledge areas.

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

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

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

2.5 Short summary of evaluation demands in a sustainability context

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

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

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

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

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