



# **EASY-ECO 1**

**EvAluation of SustainabilitY**

**EuroCOnferences**

**Conference Proceedings**

**Editors: Ursula Kopp,  
André Martinuzzi and Uwe Schubert**

**May 23 – 25, 2002 Vienna / Austria**



# **EASY-ECO I**

## **EvAluation of Sustainability**

### **EuroCOnterences**

May 23 – 25, 2002 Vienna / Austria

## **Conference Proceedings**

**Editors: Ursula Kopp,  
André Martinuzzi and Uwe Schubert**

Project Coordination and Conference Management:  
Professor Uwe Schubert  
Institute für Economic Geography, Regional Development and Environmental Management  
Department of Environmental Economics and Management  
Research Focus “Managing Sustainability”

E-Mail: [easy@wu-wien.ac.at](mailto:easy@wu-wien.ac.at)  
<http://www.sustainability.at/easy>

Vienna University of Economics and Business Administration  
P.O.Box 37  
A-1091 Vienna  
Austria

August, 2002



## Table of contents

### MARKET DEVELOPMENT

**Uwe Schubert, Austria**

Evaluation of Sustainability – New Challenges for Evaluators and Methods of Evaluation

**André Martinuzzi, Austria**

The Market for Evaluation of Sustainability – Trends and Institutionalisation

**Tomasz Zylicz, Poland**

The Role of Science in Sustainability Policies

**Ernst Max Nielsen, Denmark**

Sustainability in CEE

### STANDARDS AND

### METAANALYSIS

**Thomas Widmer, Switzerland**

Evaluation Standards in the Context of Sustainability

**Astrid Kuffner, Austria**

Evaluation of Sustainability – Sustainable Evaluation? – lessons learned from a survey of evaluation reports

### EVALUATION

**Jon Lovett, United Kingdom**

Ecological Sustainability: A Policy Oxymoron?

**William Lafferty, Norway**

Political evaluation: Premises, Approaches and Methods

**Wolfgang Meyer, Germany**

Sociological evaluation – Methods and Limits

**Clive Spash, United Kingdom**

Valuing Society and Environment: Economic Methods and Limits

**Jürgen Freimann, Germany**

Evaluating the Impacts of Corporate Environmental Management Systems. A Comparison between EMAS and ISO 14.001



## **SUSTAINABILITY**

**Edwin Zaccai, Belgium**

From Conceptions of Sustainability to Indicators

**Thiemo Eser, Germany**

Evaluation and the Appropriate Policy Level under the Demands of Sustainability

**Markus Langer/Aloisia Schön, Austria**

An Integrated Referential Framework for Sustainable Development

## **FIELDS OF APPLICATION**

**Maria-Angeles Diez, Spain**

Current Approaches used in the Evaluation of Regional Policies: a Critical Review

**Alain Thierstein, Switzerland**

An Evaluation Framework for the Quality Management of Projects on Sustainable Development

**Diane Warburton, United Kingdom**

Evaluating Participation in Sustainable Development

**Clive George, United Kingdom**

Sustainability Impact Assessment

**Kate Theobald, United Kingdom**

LASALA: Self-assessment of Local Agenda 21 for Social Authorities



**MARKET DEVELOPMENT**

**Uwe Schubert, Austria**

Evaluation of Sustainability – New Challenges for Evaluators and Methods of Evaluation

**André Martinuzzi, Austria**

The Market for Evaluation of Sustainability – Trends and Institutionalisation

**Tomasz Zylicz, Poland**

The Role of Science in Sustainability Policies

**Ernst Max Nielsen, Denmark**

Sustainability in CEE



## Uwe Schubert

### EVALUATION OF SUSTAINABILITY - NEW CHALLENGES FOR EVALUATORS AND METHODS OF EVALUATION

#### Sustainability poses new challenges for methodology and evaluators

The evaluation of Sustainable Development poses new challenges to evaluators and methods of evaluation which cannot be solved through the application of methods developed in other fields of evaluation. This can be illustrated by highlighting some of the principles of Sustainable Development:

#### Integrated Approach

- Sustainable Development aims at an integrated solution of ecological, economic and social problems. For the evaluation of Sustainable Development this means the use of multidimensional target systems and the aggregation of data from different scales and indices. For specific decisions between different options there is also the problem of aggregating data, measured on different scales, which can only partly be expressed in monetary terms. This implies the use of sophisticated and well tested methods for the evaluation and aggregation of data, the further development of which seems one of the most interesting tasks for the future.
- From this follows, however, that the evaluation of Sustainable Development should only be conducted by interdisciplinary teams, which have a common understanding of the problems involved and use a common metadisciplinary language. For example, the evaluation of a communal programme may include the analysis of the environmental impact (natural science aspect), the assessment of the technologies used (technological aspect), an evaluation of how public funds were spent (economic aspect), a role analysis (social science aspect) and changes in power constellations resulting from it (political science aspect).
- Furthermore, the integrated approach offers a chance to control longer chains of cause and effect (for example using material flow analysis), and to find solutions which would be inconceivable with a purely sectoral approach (for example through networks of agents). For evaluations, this means a significant increase in complexity of the matter investigated, the deliberate inclusion of indirect effects, and a special focus on the delineations of the system evaluated. For example, when evaluating Sustainable Development in a region, one has to clarify whether the focus is to be on direct resource use only (e.g. tons of steel) or whether the side effects arising from the production of these resources (e.g. the use of energy and materials in mining) are to be taken into account as well.

#### Emphasis on the Long Run

- The deliberate focus on a long-run perspective in Sustainable Development offers the chance to tackle environmental problems not only after they have arisen in the sense of damage repair, but to prevent them from arising in the first place. At the same time, however, there are the problems of a comparative scarcity of data on long-term developments (for example global climate change) and of the comparability of solutions involving different time scales. (The latter problem is only insufficiently tackled by economists applying discount rates.) Evaluations of Sustainable Development can therefore only be done if firmly based on the findings of natural sciences, must include long term effects and should be conducted repeatedly to assess the latter.

#### Participation

- Sustainable Development not only addresses ecological questions, but also tries to give equal weight to social issues. This is made especially difficult by the fact that, so far, no common understanding has been developed of Social Sustainability and there is no established set of indicators to measure it. The further development of such indicators would therefore be a prerequisite for its evaluation.



- The Evaluation of Sustainable Development therefore calls for a participatory approach to evaluation in general. This could be done through participatory contracting of projects and project design (through programme steering or Local Agenda initiatives), task sharing between internal and external evaluators, or increased involvement of stakeholders in the assessment of results.
- As a result, evaluation also shares some characteristics with mediation and can contribute to a change in lifestyle patterns, increased awareness, the strengthening of problem-solving capacities of various stakeholders in society and to develop vision statements of Sustainable Development.

**Implicit understanding of the concept of Sustainable Development raises the need for a meta-disciplinary dialogue**

Evaluation of Sustainability requires a common understanding and a shared usage concerning the concept of Sustainable Development. In evaluation practice, this common understanding is often assumed, although literature research shows that there are vastly different scientific positions. The following examples show the extreme positions held in the discussion about Sustainable Development:

Examples	Extreme Position A	Extreme Position B
<b>Normative approach</b>	Sustainable Development means to identify natural limits and to take them into account in all societal decisions. Within this approach, ecological limitations put normative limits on economic and social activity.	The discussion of ecological problems and natural limitations is a societal process and thus to be observed without any value judgements. The normative approach of natural science cannot be justified since it is the goals of societal agents that are important.
<b>Dimensions of goals</b>	Sustainable Development aims at an equally weighted and simultaneous representation of ecological, economic and social goals (three-pillar model).	Different authors have extended the three-pillar model (adding, for example, cultural and institutional sustainability) or changed the weights given to the individual pillars.
<b>System boundaries</b>	Regionality is a natural principle which should be reflected in the concept of Sustainable Development. Therefore, all claims resulting from the guiding principles of Sustainable Development have to be fulfilled on a regional basis. Violations of system boundaries (for example through imports and exports and far-reaching indirect effects) are to be minimised. Taken to the extreme, this concept advocates regional autarchy.	Through regional, national and international division of labour, higher efficiency can and should be attained (e.g. through specialisation and economies of scale). Sustainable Development therefore is a global concept, from which no demands for regional sustainability can be deduced.
<b>Substitutability</b>	Non-regenerative resources and energy sources must not be used, as their rate of regeneration (e.g. fossil fuels such as oil) exceeds the expected life span of mankind by far. A shift to renewable resources and regenerative energy sources is therefore indispensable.	The use of regenerative resources and energy sources can build up human capital (for example better education and higher social security) which justifies its use and can also lead to the development of new resources or the more efficient use of existing ones.
<b>Ability to innovate</b>	Most technological developments have led to a higher rate of use of natural resources and energy, and higher efficiency could only marginally compensate for the overall growth. Since the risk potential of new technologies (nuclear power, genetic engineering) is disproportionately high Sustainable Development is associated with "soft" and decentralised technologies or even implies the banning of certain technologies.	Only through the development of new technologies can a solution to the high use of energy and other natural resources be found. Today's technologies only show the way to future technologies. To renounce the use of present technologies would reduce the ability to innovate and thus prevent progress towards Sustainable Development. The scarcity of resources itself and the urgency of current environmental problems will create pressure to raise the ability to innovate.
<b>Political Background</b>	Sustainable Development is a political concept which includes questions of distribution (e.g. between current and future generations or between less developed and industrialised countries). Therefore, methods are needed to balance interests of different social groups and to critically assess their wants and needs (for example the principle of subsistence).	Sustainability can best be achieved by decoupling the satisfaction of needs from energy and resource use (e.g. the factor-4 concept and eco-efficiency). Technological innovation may enable needs to be satisfied through qualitative growth and may thus also solve distributional problems.



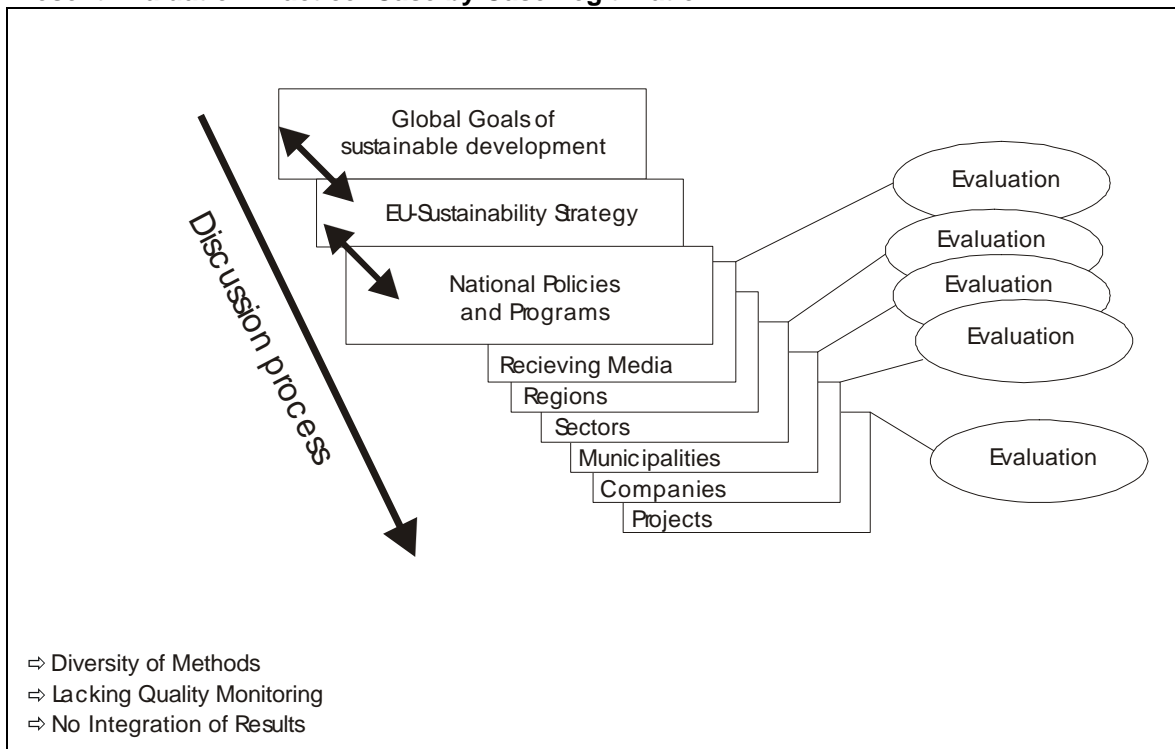
In view of this diversity of positions we can expect evaluation results to be strongly biased, depending on which concept of Sustainable Development is used in the evaluation. Therefore the same evaluandum may be assessed positively or negatively according to the assumptions underlying the evaluation. At present this problem is not adequately addressed, nor is it discussed sufficiently on an academic level. In evaluation reports, these paradigmatic questions are rarely addressed and the commissioning agencies are largely unaware of the consequences of different paradigms. In light of this situation, further research is clearly required.

### Evaluation of Sustainability must be seen as a part of a Sustainability Management System

Sustainable Development aims at the integration of environmental protection with other policy areas, which requires the integration of the aims of different sectoral policies. For this purpose, a cross-sectoral goal system is needed, the development and harmonisation of which calls for additional coordination. Since the number of agents involved is large, flexibility is needed and the principle of subsidiarity, in this case hinging critically on the assumption of better know-how in special agencies, has to be taken into account. The process of defining these goals cannot be centralised, but local and regional agents have to be involved in the implementation of Sustainable Development plans.

A policy of Sustainability must therefore concentrate on setting overall goals and rough guidelines and leave it to decentralised agents to decide how to attain them, as they usually have the best technical information and know-how. Through the setting of macro goals, a framework for orientation is created for the decentralised agents, which allows closer monitoring of decentralised activities with regard to their attainment of goals (Target base decentralised planning). Therefore, evaluations are not only assessments of the “success“ of programmes and policies, but part of a larger “Sustainability-Management-System“ on supranational, national, regional, communal and corporate levels, which consists of agreed goals, operationalisation of concepts, programme conceptualisation, programme implementation and feedback (=evaluation). It should lead to a more objective environmental policy, guarantees an adequate use of resources to attain goals and reduces the need for tedious, detailed coordination on the part of the agents. Environmental policy thus becomes more rational, goal-setting becomes more similar to corporate management systems, the organisational prerequisites for the attainment of goals are created and the attainment of the goals itself is measured.

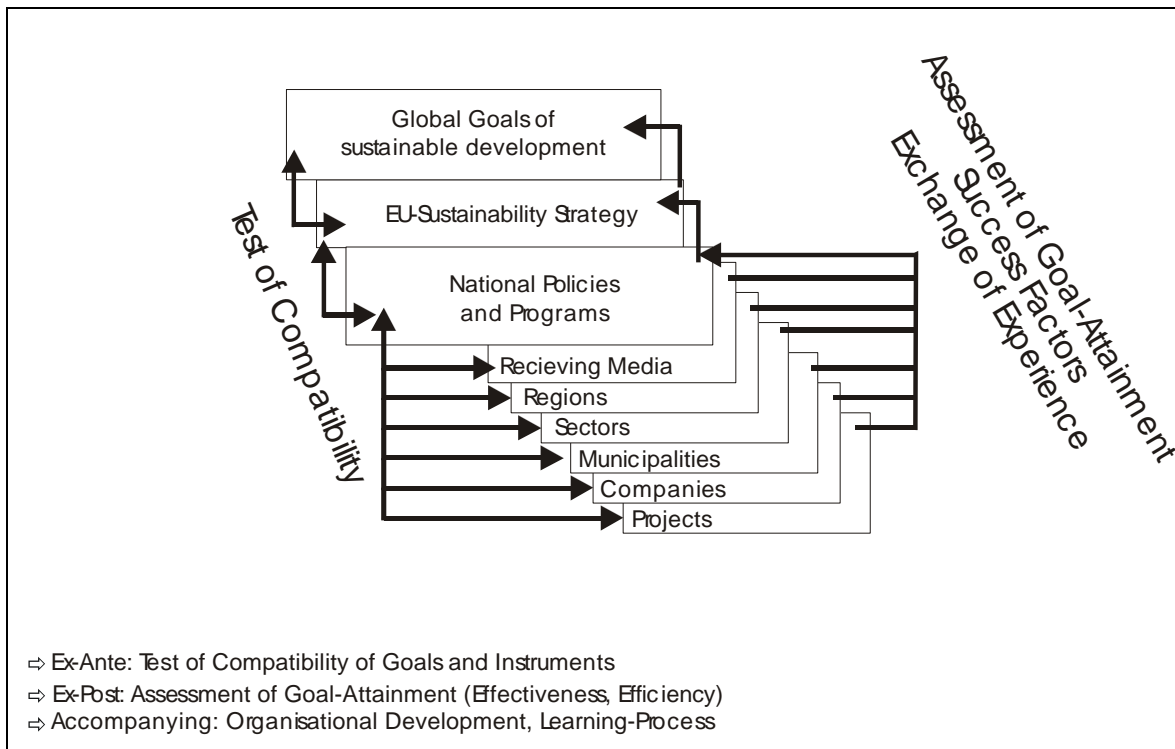
### Present Evaluation Practice: Case by Case Legitimation







## Evaluation in the Future: Integral Part of a Sustainability-Management-System



Evaluation of Sustainability is therefore not an isolated task to measure the effectiveness and impact of individual projects, but an organised feedback process for decision-makers in politics, business and society. The sustainability strategy which the EU is currently developing, and which is a top-down process designed to integrate environmental concerns into other policy areas, particularly requires a feedback process which...

- secures learning effects and enables the transfer of experience within and between sectors, and between different tiers of administration,
- guarantees decentralisation and subsidiarity, and therefore efficient decision-making structures,
- ensures the attainment of the goals following the hierarchy of tiers ,
- provides a reference framework for the evaluation of smaller units (individual companies or projects)

In the context of Sustainable Development, evaluations are therefore not just ex-post outcome analyses, but instruments to co-ordinate and harmonise the goals of different agents and part of a feedback loop designed to provide information about the success of implementation activities.

## Conclusions

Evaluation of Sustainability could play a significant role in the implementation of an international, national and regional sustainability policy. Therefore the following tasks will have to be taken care of:

- Theoretical discussion of the paradigms of Sustainable Development, reflection on the basic assumptions underlying evaluation projects and their consequences for the evaluation of Sustainable Development.

This discussion should lead to a generally recognised operational concept of "Sustainability", directly applicable for evaluation projects. To overcome the semantic confusion between different disciplines involved a referential framework for a meta-dialogue is warranted.



- Development of an interdisciplinary framework to systemise present know-how available from the various disciplines dealing with the evaluation of Sustainability.

The prominent purpose of this endeavour is to show the contributions as well as the limitations of the different disciplines involved and bridging the gap between them. Only a modular combination of the know-how of the various disciplines can meet the challenge posed by the postulate of sustainable development plans.

- Critical reflection on current evaluation practice via meta-evaluation, comparative methods analysis, country analysis and dialogues between scientists and practitioners.

The aim of this effort is to increase market-transparency in the field, more extensive exchanges of experiences as well as ascertaining the quality of evaluation analyses. Surveys, meta-evaluations and a continuous dialogue between scientists and practitioners is indispensable to this end..

- Fundamental research to further develop methods for the evaluation of Sustainable Development with a view to creating a modular toolkit.

Theoretical concepts need to be developed which facilitate the integration of different evaluation approaches by creating common concepts, which should make cause-effect analyses viable. It is only on this basis that evaluations can make assessments of impacts of policy interventions.

- Further development of the Evaluation of Sustainability into an integral part of a Sustainability-Management-System which shows successful implementation on different levels and contributes to a transfer of learning effects.

The final goal is to establish evaluation as a social learning and feedback process , which enables a wider range of actors in development policy to benefit from the experiences made in the process of sustainable development. This requires professionalism of agencies commissioning evaluation projects, raising the awareness of commissioning agencies of the need for an early conception of evaluations instead of mere ex-post analyses and measures to improve the market transparency.

For the attainment of these objectives monitoring and analysis of current demand trends on the international, national and regional levels, collection of evaluation reports, conducting meta evaluations and comparative country analyses are necessary. EASY-ECO constitutes an attempt to contribute significantly to this end.



## **André Martinuzzi**

### **THE MARKET FOR EVALUATION OF SUSTAINABILITY - TRENDS AND INSTITUTIONALISATION**

In the last 25 years Sustainable Development has evolved from a political vision into an essential part of international and national policy and politics. The resulting request to integrate environmental, economic and social goals makes the goal-setting-task of policy makers more complex. Ascertaining the consistency of policy signals and the monitoring of goal-attainment are becoming an integral element of implementation. The tasks mentioned can be assisted significantly by evaluation.

A move in environmental policy from the sectoral and problem-oriented application of command-and-control instruments to incentive-based instruments can be observed. From this new style of policy making stems the necessity to choose instruments with the intention to further the co-operation of the different actors and to provide incentives for innovation at a decentralised level.

This new combination of the vision of sustainability and a new style of policy warrants the creation of a market for evaluation analyses of sustainability. During a period where rising institutional demand at all levels is being established, the theory and practice of evaluating sustainable development has still not been fully understood: the academic discourse is split between many different disciplines. Quality control, institutionalisation and dialogue about the methods of evaluation hardly provide a guide in this expanding area. Based on a survey of evaluation reports from Germany, Austria and Switzerland, this paper will discuss the driving factors of the market, trends and institutionalisation of the evaluation of sustainability.

---

Acknowledgement: The author wishes to thank Astrid Kuffner and Wolfgang Meyer for the valuable discussions on this topic in the last years.



## The demand side – increasing need but little experience

### Driving factor 1 – Establishing a Sustainability Policy

The gradual establishment of the model “sustainable development“ can be understood by the following milestones:

- 1987 „**Our Common Future**“ – The „Brundtland Report“ of the World Commission on Environment and Development (WCED) establishes sustainable development as a new model for long-term environmental, economic and social policy making. The challenge, to balance the needs of the people of today without restricting the possibilities of future generations, has from now on been accepted as a principle; an appeal to the awareness of responsibility and inter-generational justice was clearly made. Until now there have been a lot of definitional and operational investigations which have emphasised different issues.
- 1992 **International Conference for the Environment and Development (Rio-Conference)** – The targets of Brundtland Report are set in the real political context of global development (climate convention, convention about biological variety, the Rio declaration about the common responsibility of states for the environment of the planet, Agenda 21)
- 1997 The **Amsterdam Treaty** declared „sustainability“ within the EU as an overarching goal. The basis of sustainable development was adopted in the preamble and in the goals of the EU treaty as well as in Article 2 of EG-Treaty; on the same level as economic growth and social solidarity. This may result in the need for evaluation and for a concrete practical realisation of these aims. In addition, the new article 6 in the main chapter of the Treaty takes over the clauses concerning the inclusion of environmental protection. Indicators for monitoring and assessments of achievement are to be developed. This alone can result in an increase of demand for evaluation analyses of sustainability and therefore an established methodology. Article 12 further requires that projects and programmes funded by the EU-structural funds make a contribution to the protection and improvement of the environment. This rule implies an implicit obligation to evaluate with respect to sustainability.
- 2001 At the Council of Göteborg the EU **Sustainability Strategy** was presented, which contains premises for the monitoring, evaluation and development of national sustainability strategies. Within the sixth environmental action program the integration of environmental law into other political areas represents a new focus. Accelerated implementation of environmental law in the member states as well as the participation and information of citizens of Europe constitutes a natural national follow-up to take into account.
- 2002 From the **Johannesburg Summit** a stronger impulse for global and regional sustainability is expected.

In the last 25 years Sustainable Development has evolved from a political vision into an essential part of international and national policy and politics. The resulting request to integrate environmental, economic and social goals makes the goal-setting-task of policy makers more complex. Ascertaining the consistency of policy signals and the monitoring of goal-attainment are becoming an integral element of implementation. The tasks mentioned can be assisted significantly by evaluation.

### Driving factor 2 – Programmes as an instrument of a new political style

The second driving factor is a move in environmental policy from the sectoral and problem-oriented application of command-and-control to incentive-based instruments which aim to develop integrated solutions. Numerous actors and political fields address these problems and strive for an integrated solution. From this new style of policy making stems the necessity to choose instruments with the intention to further the co-operation of the various actors and to provide incentives for innovation at a decentralised level. While the environmental policy for European states (and in particular the German speaking areas) in the last 20 years has been based on command and control, with the resulting flood of environmental laws, along with a performance deficit and high co-ordination expenditure, the last few years have seen a gain in importance of the market-conforming and incentive-oriented instruments of environmental policy. Coming from the Anglo-American countries, these well-tested, straight-forward programmes for the promotion and practical realisation of voluntary environmental protection measures have resulted in a rapid diffusion at different levels of environmental policy.



- On the **European level** there are, above all, the structural funds, the framework programmes of research and technical development as well as the implementation of the EMAS regulations and the accompanying pilot projects. Present examples of evaluation of sustainability at the EU-level are the compulsory evaluation of the structural funds' programmes, the newly developed Sustainability Impact Assessment (SIA) and the Strategic Environmental Assessment (SEA) of plans, projects and programmes as well as the discussion about evaluation in the development of the sixth environmental action programme and in the implementation of the EU sustainability strategy.
- On the **national level** national sustainability strategies are developed, numerous funds and subsidies as well as voluntary agreements are implemented (e.g. "Umweltpakt Bayern"). Most of the examples on the national level are evaluations of subsidies and funding-programmes (e.g. renewable energy sources, ecologically oriented agriculture, development aid).
- On a **regional and communal level** numerous programmes are to be found which are initiated by established regional actors (for example industry and chambers of commerce, research and technology centres).<sup>1</sup> Furthermore the evaluation of Local Agenda 21 projects is becoming more and more important.

As a result sustainability policy concentrates on goal-setting and leaves the application of instruments and funds to decentralised agents. On the one hand, the use of public funds creates a need for public accountability and legitimisation, which results in increased demand for evaluation. On the other hand, the success of these new instruments of environmental policy depends to a much higher extent on individual decisions taken by those affected (for example the decision if a company should join EMAS or not) so that evaluations can be used to optimise projects, support organised learning and disseminate and share implementation experience. Evaluation can be used as a feedback-loop for conception and implementation of sustainability policy. Therefore a significant increase of demand for evaluation projects is to be expected within the next few years. **Driving factor 3 – clients lack experience**

Most of the out-contractors of evaluation projects are international organisations or public authorities at federal, regional and communal levels. Private clients (associations, companies) have until now only rarely demanded evaluation services and these predominantly as ecological or economic expertises of selected aspects.

Many clients have only little experience in contracting out such projects. They rarely have the methodological knowledge required to choose the best evaluation concept and they typically have no overview of the present market of evaluators as well as different methodological approaches, their advantages and disadvantages. Since standards for evaluation (developed in the US) are not commonly used in Europe, a minimum quality standard is not guaranteed. Evaluating projects are therefore typically commissioned under circumstances of great uncertainty and without objective criteria. As long as there are no corresponding standards of quality control or the form of inquiries are unknown, there is a fear that the increase of demand will lead to a deterioration of quality. Therefore it is especially important that potential clients actively participate in the development of evaluation standards and in the institutionalisation of quality control. Moreover, professional training of and networking between clients of evaluations are necessary.

---

<sup>1</sup> Selected cases of such programs are:

- The „EcoProfit-Programm“ (ÖKOPROFIT): After considerable success in Austria is now spreading rapidly to Germany and other European countries. Many variants of the original concept are now being considered and implemented (e.g. EcoProfit –Tourism in Graz and Vorarlberg; EcoBusinessPlan Vienna, Environmental Program "Linz lebt Umwelt"). Up to the recent past more than 1000 companies at more than 50 locations have participated in the EcoProfit-Programme all over Europe.
- „The-Natural-Step“-Programme developed in Sweden is oriented to the principles of Sustainable Development. More than 100 companies and more than 50 municipalities have participated in this training- and consulting-programme.
- Various subsidy-programmes to support the diffusion of EcoManagementSystems such as EMAS have been implemented. These contain also elements of group-consultations (z.B. EcoPartnership Munich) and additionally image-effects by public statements of politicians (e.g. „50 EMAS-companies for Frankfurt“, „Trierer EG-Eco-Audit Model“).



## The supply side – poor institutionalisation and hardly any interdisciplinary exchanges

Above all it is the public or private research institutes who are commissioned to perform the evaluations. In any case, in the states of the European community, there is not as yet **any clear differentiated professional image of evaluators**. Since the use of the title of “evaluator” is not protected and there are no established “brands”, entering the market requires little investment in terms of both capital and personnel. This entails the risk that evaluations not reflecting the state of the art may ruin the reputation of the whole profession.

Practice has shown that evaluators can assume three types of roles:

- **„Evaluator as technical expert“:** Evaluation is exclusively understood as measurement of effects and serves above all the legitimization of political decisions. Most of these evaluations are conducted ex-post. The expertise of the evaluator determines the rationality of the client, the decision about the evaluandum is in the centre of the analysis. The dominant perception is of ‚right and wrong‘. Important examples are projects which result from excessive public expenditures.
- **„Evaluator as moderator“:** Evaluation is used as ex ante comparison between several alternatives and also serves as a solution of conflicts in contested projects. Neutrality and competence in process moderation determines the rationality of the contract. The dominant perception is ‚finding a solution together‘. Important examples are projects characterised by intensive societal conflicts.
- **„Evaluator as coach“:** Evaluation accompanies the practical realisation of an organised process of learning. This role serves the continuous improvement of the evaluandum and the exchange of know-how between all actors involved. Orientation of supervision, degree of innovation and procedural knowledge determines the rationality of the contract. The dominant perception is ‚What can we do better‘. Important examples are projects with a high degree of innovation.

In Europe **the degree of institutionalisation among evaluators is still very low**. Although associations of practitioners of evaluation exist in several European states (e.g. DeGEval, SEval, EEA etc.), there are no professional representatives and supervisory institutions yet, which could guarantee a minimum quality standard. In particular, in the field of the environment, the participation of the practising evaluators and clients in the newly founded Evaluation Associations is currently very low. Taking this into account, these associations’ ability to take over the function of quality control still requires considerable efforts.

A further deficit is in relation to the scientific development of evaluation concepts and methods. There still is a lack of interdisciplinary working groups, specialised research institutions, publication media and other platforms of information exchange as well as an integration of different scientific disciplines. An analysis of evaluation reports and research projects shows that **a multitude of different disciplines, approaches and methods are used**, but none of them can, individually, cover the complete task of evaluating sustainable development policy. The following disciplines or sub disciplines work on particular aspects of the Evaluation of Sustainability:



Discipline	Current State of Discussion	Contribution to the evaluation of Sustainable Development
<b>Instruments of Environmental Policies and Environmental Economics</b>	The scientific discussion about the use of instruments of environmental policy and environmental economics has a long tradition. However, until now little knowledge has been gained about the compatibility of incentives of different policies (consistency evaluation) and the specific reactions of stakeholders to the signals of environmental policies.	The increased use of market and incentive- based instruments increases the decision-making power of decentralised agents (companies, households, communities, regions). The measurement of the impacts of environmental actions must therefore be complemented with detailed knowledge about the effects of the instruments of environmental policy.
<b>Indicators of Sustainable Development</b>	Broad scientific debate about indicators (for example area-based versus mass- based indicators, multitude of individual indicators versus integration of indicators into indices). Until now, no standard set of sustainability indicators has been established. The central issue is how indicators can be better integrated into political decision-making processes.	On the one hand, indicators provide the factual basis which makes the impacts of policies, programmes and projects measurable and easier to demonstrate. On the other hand, indicators rarely give explanations as to <i>why</i> the effects or developments observed have occurred.
<b>Assessment methods (CBA, MCA)</b>	Cost benefit analyses (CBA) and multi criteria analyses (MCA) are used to aggregate multidimensional data to decide between various alternatives. Present knowledge in this area is largely quantitative.	Especially for the evaluation of Sustainable Development, the integration of data from different descriptive domains is necessary. CBA and MCA are mainly used in the preparation of decisions and ex-ante evaluations of large projects.
<b>Evaluations by Social Sciences</b>	The most important methods of social sciences used in evaluations are interviews and opinion surveys. For this purpose, researchers can fall back on detailed methodological knowledge. Methods of qualitative social research (for example narrative interviews) could provide explanations for the effects observed, but are rarely used in evaluations because the results are not representative.	Phenomena observed by social sciences (like attitudes, value judgements, opinions, decision-making bases, problem-solving capacity) are central factors determining the progress made on the way towards Sustainable development. The consideration of these phenomena is the second important data source for the evaluation of Sustainable Development, complementing indicator-based technical and (natural) science data. Furthermore, the methods of qualitative social research could give an insight into the causes of developments and effects. Also, the further development and operationalisation of the concept of Social Sustainability is a necessary prerequisite for the evaluation of Sustainable Development as a whole.
<b>Public Administration</b>	Evaluations always take place in a political context (through the awarding of the evaluation contract, the results of evaluations and the recommendations based on them). Know-how from political sciences is mainly used in the evaluation of policies. In sciences of public administration, the use of business principles is discussed within the framework of New Public Management.	Taking the political context into account, the (natural) science-based approach to evaluations of Sustainable Development is complemented with the context of practical applications and the micro-political strategies of the stakeholders involved. The know-how of new public management is particularly necessary for the development of better policy recommendations in order to increase the efficiency of policies, programmes and projects.
<b>Regional and communal development</b>	In communal development (visions for development, land use management) as well as in regional planning (transport routes and other infrastructure planning), evaluations already have a long tradition.	The compulsory evaluation of the Sustainability of EU structural funds programmes and the application of the Strategic Environmental Assessment (SEA) both increase the importance of this area of research for the evaluation of Sustainable Development. Also Local Agenda 21 projects create additional demand for evaluation.
<b>Participation, Moderation, Mediation</b>	Environmental Protection has a long tradition of conflict (mainly through disputes about atomic power stations and large projects). As a result, some methods of conflict resolution and involvement of stakeholders have been developed.	Sustainable Development is only possible through participation of those affected. As mentioned earlier, methods of participatory evaluation have to be developed.
<b>Environmental Management Systems</b>	Setting up an environmental management system requires an assessment of environmental impacts, the formulation of environmental policies and the declaration of an environmental programme. External evaluators assess the system and its implementation.	The external assessment of environmental management systems is an established part of corporate environmental evaluation and evaluations of Sustainable development can draw on this experience. Moreover, corporate controlling systems share many functions with evaluations (measuring performance against objectives, identification of weak spots, feedback). The experience thus gained can be used for other evaluation projects.
<b>Standards and Guidelines for Evaluation</b>	Originating from the U.S., the establishment of standards and guidelines for the evaluation of Sustainable Development is now also being discussed in Europe. Scientists expect this to lead to minimum quality standards, professional representation of and image protection for evaluators and higher credibility with clients.	Most of the standards are independent of the evaluandum and therefore process-oriented. Their application can lead to a better image and relationship with commissioning agencies.
<b>Meta Analysis and comparative analysis of methods</b>	In contrast to the U.S., only few meta analyses and meta evaluations have been done in Europe. Systematic comparative analysis shows the methods used, the state of the art in evaluations and the strengths and weaknesses of the methods used. Meta analyses can be used to integrate the results of individual projects into a common framework.	The evaluation of Sustainable Development should not be confined to isolated projects, but should exploit the full strategic benefits of evaluations. This can be achieved by integrating evaluations already when designing policies, programmes and projects and by coordinating the implementation of individual evaluation projects.



## Conclusions - weak points and challenges

Since 1999 the research focus “Managing Sustainability” at Vienna University of Economics and Business Administration has collected and assessed over 60 evaluation reports from Germany, Austria and Switzerland according to two criteria:

1. Do the evaluation reports treat all aspects of Sustainable Development (e.g. integrated approach to the three dimensions of sustainability, emphasis on the long run, participation) ?
2. Do the evaluation reports satisfy the standards for evaluation (so far can be answered through documentary analysis) ?

As a result of this meta-evaluation the following weak-points and challenges for the Evaluation of Sustainability policy have been worked out:

- implicit understanding of Sustainable Development determines the study design and the results
- weak theoretical models backing many evaluations
- insufficient consideration of system boundaries (hardly any attention to indirect effects)
- not many inter-disciplinary evaluation teams
- most reports only measure consequences and only very rarely are causal models applied
- insufficient data availability, bad timing
- faulty causal attribution and self-promotion as biasing factors are often neglected
- evaluation reports only seldom published
- legitimisation stands in the foreground, organised learning process is neglected.
- neutrality postulate for evaluators are neglected
- lack of participation concepts as an integral part of evaluations

To cope with the expected strong increase of demand, Evaluation of Sustainability as a political and societal feedback-function is to be established and the weak points detailed above can be remedied quickly, in the opinion of the author, by the following necessary steps.:

- **Setting up qualified training possibilities and further education for evaluators**  
University curricula are to be created at least at some European universities. EASY-ECO at the moment still seems to be the best educational option available.
- **Professional training of clients**  
Sensitising for basic assumptions and paradigms underlying evaluation projects and their consequences for the evaluation of Sustainable Development, training in evaluation project-management.
- **Transparency of evaluation results**  
Publication of all evaluation reports, realisation of surveys and meta-evaluation to promote the increase of market transparency and quality control, exchange of experiences between evaluation projects, evaluators and locations. **Establishment of evaluation as a societal feedback process**  
Creation of a climate of learning, openness and trust in evaluation, use of knowhow from controlling and organisational development to build up Sustainability-Management-Systems.





**Tomasz Zylicz**

## **THE ROLE OF SCIENCE IN SUSTAINABLE DEVELOPMENT POLICY**

### **Abstract**

Sound sustainability policies are based on the classic Brundtland definition of sustainable development, i.e. "to meet the needs of the present without compromising the ability of future generations to meet their own needs". On the other hand, there are ambiguities with respect to what development patterns do "compromise the ability of future generations". As a result, a number of academic studies and sustainability assessments carried out seem to be inconclusive. In many countries a typical study demonstrates the environmental friendliness of a given project without addressing the question of whether the project is sustainable if continued and/or replicated by others.

The paper draws examples from academic assessments in Poland in order to identify areas where scientific expertise is strong and abundant, and those where studies suffer from omissions or inadequate methodologies. In particular, domestically funded assessments, typically focused on technical and engineering aspects, are contrasted with EU-commissioned ones covering social and political contexts. The author concludes by calling at wider co-operation between domestic and international academic institutes so as to address all salient issues relevant for sustainability.

### **1. Introduction**

The popularity of the sustainable development concept does not come along with the commonness of its understanding. For some it is simply the successor of the idea of environmental protection at a time when it has become clear that the traditional "pure" protection is no longer politically appealing. Others see it as a way to improve the weakened reputation of the industry which strives to maintain a leading role in the forming of economic policy. Natural scientists, for their part, stress the ecological aspect of sustainable development thus creating a tendency to substitute the English term "sustainable development" with the Polish "eco-development". Only sporadically does the original concept laid down in the Brundtland Report (WCED 1987) appear, the concept which was the starting point for the *Agenda 21* accepted in Rio di Janeiro in 1992. The classic Brundtland definition identified sustainable development with "meeting the needs of the present without compromising the ability of future generations to meet their own needs" (Pezzey 1998).

The classic definition of sustainable development consists of few words, being at the same time very significant and substantial. It may be criticised for excessive generality but that is precisely what makes the definition non-controversial and hard to challenge. In other words, nearly everyone should agree with it. Yet many practical corollaries may be derived from it. All chapters of the *Agenda 21* are, in a way, an extension of the classic definition, trying to find a solution for overcoming poverty and other contemporary problems without making the future generations pay the debt.

There exist two completely different concepts of securing a sustainable base for future development, described as "strong sustainability" and "weak sustainability". The first concept calls for keeping natural resources as well as, additionally, any other resources related with human activities. According to the second concept, it is necessary to maintain an unchanged total amount of resources. That is to say that using up environmental capital may be in accordance with the concept of weak sustainability as long as it is compensated by the accumulation of other resources, that is human capital (i.e. knowledge, technology, institutions and behavioural norms) as well as man-made capital (work tools and other accessories).

Many researches have shown that particular kinds of resources may not always be freely substituted one by another. Of course within some limits the loss of one resource may be compensated by the growth of another. For example, the insufficiency of human or man-made capital may not be strongly felt if accompanied by an abundance of natural capital. The functioning of the economy and society, however, requires all three capital components and even ordinary people will find it obvious that it is impossible to go on substituting for example natural capital with other capital types forever (Daly 1990).



The limitations to substituting particular basic kinds of capital are an argument against the weak sustainability concept. Leaving the future generations with an unchanged total capital amount can reduce the chance to meet their own needs if the total lacks any of the key components. The economic theory can solve this problem only by creating a new one. Namely, if one of the capital components turned out to be crucial for meeting the needs, its value should resemble it. But then since it was substituted by another, "defective" one, obviously their summed value diminished, which contradicts the initial assumption. Looking realistically at the case though, it would be hard to expect anyone to have at their disposal a fully adequate assessment of biological diversity, natural landscape or even clean water. Furthermore, this hypothetical assessment would have to be made according to the preferences of the future generation and not the present who are responsible for current economic decisions. The possibility of practical operationalisation of the notion of weak sustainable development is therefore highly controversial.

Yet the notion of strong sustainability is just as controversial. Natural capital comprises both renewable resources which may constantly be exploited and resources that are depletable. The latter - by definition - cannot be passed on to the next generation undiminished unless the present generation gives up using them at all. As it is impossible to imagine a world of today without crude oil and other raw materials, it is also impossible to leave the strong sustainability principle in its strict form. A somewhat milder version of the principle has been formulated, one that permits using depletable resources at the price of creating their renewable substitutes (Daly 1990). This principle refers to the so-called John Hartwick rule (1977) in the economic growth theory which calls for the investment of scarcity rents obtained from the exploitation of depletable resources. The Hartwick rule does not specify the destination of these capital expenditures, though; they may build up any of the capital types. If the rents are consumed and not invested, economic growth will not be sustained.

There is actually no obvious logical justification for Herman Daly's principle. It assumes the possibility to substitute the two components of natural capital - renewable and depletable resources - although the relation between them may in reality be more complicated. It is nevertheless quite eagerly acknowledged as a principle that allows moderating the strong sustainability principle without resorting to the weak principle which "sacrifices" natural capital for other kinds of capital too easily.

Irrespective of the doubts that accompany the attempts to operationalise the notion of sustainable development, it is something that may be subject to empirical testing. What turns out to cause much more trouble is the rational justification of the principle of equal opportunities - the very fundament of the sustainability concept.

The principle of fair division of welfare between generations is derived usually from Kant's categorical imperative, especially the way John Rawls expresses it (1971). The core of such deliberations is the idea of a hypothetical convention where representatives of different generations meet to decide upon the division of welfare between them. It is of great importance that the representatives not know which generation they belong to, in other words they should act "behind the veil of ignorance".

Needless to say it is impossible to make people act as if they did not know who they were; in some circumstances, however, negotiating an agreement "behind the veil of ignorance" is possible indeed. One example is the arrangement and approval of the convention on preventing climatic changes caused by burning fossil fuels before the parties had unquestionable knowledge as to how much each one would lose due to these changes. (The convention was accepted in Rio in 1992.) Usually, however, parties are aware of their own position and interests so it is unlikely that sustainable development instruments, e.g. taxing energy sources or subsidising clean technologies, should be introduced on a scale that is indeed economically justified.

## 2. Key Research Areas

The scientific dilemmas outlined above show the breadth of the area of research that is crucial for the implementation of the sustainable development concept. It includes mainly traditional environmental protection that comprises both natural science and technical disciplines. There is no sustainability without guaranteeing the environment elementary protection from attacks by human activity. Scientific research should therefore encourage people to recognise the natural components of the environment as well as the relations between them. It should also help natural scientists and engineers design equipment that reduces the negative influence of the economy on the environment and find ways to repair damages once these have occurred.



Lastly, scientific research ought to indicate the relations between public health and the state of the environment thus increasing the effectiveness of medical care.

Traditional environmental protection is merely a small part of the sustainable development mechanism. As the above review shows, sustainability requires solving a number of problems that have to do with the economy itself as well as the relation between man and the environment and man and the products of man's activity.

Within the frames of economic sciences there are controversial and unsolved dilemmas, among which the most important one is the substitution and complementarity of different kinds of capital, namely human, natural and man-made capital. In Poland, the theoretical aspect of the problem has so far attracted very few researchers while no empirical work whatsoever has been done. Meanwhile determining to what extent the loss of capital may be compensated by investing in other capital types is substantial for the sustainable development policy. The case is similar when regarding the possibility of substituting depletable resources with their renewable substitutes. It is impossible to conduct an efficient and effective state policy without such deepened analyses.

Another research area that allows to understand sustainable development is natural resource valuation including the damages caused by the degradation of the environment. Although some scientific research in this field is being done in Poland, it is rather occasional and has little to do with international science. It is worth adding that nature valuation methodology has improved considerably in the world during the last two decades, allowing more and more precise estimations of the economic value of some resources. The results of such analyses are of great importance when planning the environmental protection strategy as they explain the nature of the benefits that may be expected after keeping or restoring particular elements of natural environment (Zylicz 2000).

Apart from economics, other humanities are also essential for the implementation of sustainable development. Since intergenerational equity is one of the important and at the same time controversial elements of the concept, as it was pointed out earlier, not only sociology but also philosophy should be applied in policy design. Intergenerational equity is the subject of economic inquiry although its practical meaning seems to be rather insignificant (Zylicz 2001). What appears to be more important is subjecting the notion to sociological and philosophical analysis. Thanks to philosophy it would be possible to deliberate on the subject of values in man's life and their role in economic and social development, whereas sociology is the appropriate discipline on the basis of which to compare theoretical concepts with the actual social and political life organisation.

### 3. The involvement of science

There are many scientific programmes whose aim is to help implement sustainable development. The *Global Change* scientific network is probably the oldest one of them. It originated as a response to the challenges resulting from the disturbing changes in global ecological systems caused by human activity. Due to its spatial and time extent, the network is in fact devoted to sustainable development issues. The network's main interest area are intergenerational issues. Human capital management and institutions whose goal is to maintain social order in the process of global changes are beyond its interests, though.

The social aspect of sustainability lies in the heart of another scientific network's interests, namely the *International Human Dimensions Programme*. The projects carried out within this network deal exactly with the social perception of global ecological concerns as well as the way the economy and political institutions respond to those.

State and regional governments try to steer scientific research so as to find support for sustainable development policy. We shall illustrate the outcome of these efforts with the example of projects financed by the European Union (EU) funds included in the *Fifth Framework Programme* and also - during the last decade - by the Polish Committee for Scientific Research.

In the EU, sustainable development has become a nearly constitutional principle. However, the same standing has been given to free trade with which it is, at times, incompatible. As practice shows, in conflict situations priority is given to free trade or other "non-ecological" principles. Sustainable development has nevertheless been ratified in many official documents and was also given priority treatment in the competition for scientific research budgets. As a result, the notion frequently appears in EU-financed research.



And so in the *Community Research and Development Information Service (CORDIS)* data base there are as many as 1153 records that contain the key word "sustainability". It appears that there are 8 long-term programmes dealing exactly with sustainability. Out of the programmes that are financed by the EU, in 71 the word "sustainability" appears in the title. In the next 274 programmes, "sustainability" is listed among the key words characterising the contents of the project. All this proves the great consideration of the notion in the EU. Scientists are, in a sense, "programmed" to associate the subject of their research with sustainability.

Among the projects that declare a direct relation to the notion of sustainability there is, of course, ecological research, e.g.:

- Phosphorus cycling and sustainability in agroforestry systems in the humid tropics
- Biological criteria for sustained development in natural degenerate forests of Mediterranean Europe
- The impact of forests and silvicultural practices upon the extreme flows of rivers
- Low-input agriculture and soil sustainability in Eastern Europe

Other projects are limited to technical sciences, as indicated by the following titles:

- Technological paradigms and transition paths: the case of energy technologies
- Durable binders from renewable sources for paints and coatings in the building industry
- Demonstration plant for steam gasification of PVC rich plastic waste
- Sonar technology for monitoring and assessment of benthic communities
- Product development and innovation in shipbuilding
- Improvement of the storage process of electricity in remote photovoltaic installations

A different category comprises a number of projects which involve social disciplines, especially economics:

- Application of non-monetary procedures of economic valuation for managing a sustainable development
- Tourism eco-labelling
- Economic and social implications of moving towards environmental sustainability through fiscal reform
- Resource-based growth of new, technology-based firms - new theory and its empirical application
- New elements for the assessment of external costs from energy technologies

Clearly interdisciplinary projects that bring together specialists in diverse disciplines such as economics, ecology, political sciences, engineering etc., should also be listed:

- Land, culture and crisis. From productionist success to fiscal and environmental impasse on European farms (1940-1990)
- Sustainable new housing in Ireland
- European network for management of arthropod resistance to insecticides and acaricides
- Strategies towards the sustainable household
- Climate, hydrochemistry and economics of surface water systems
- Baltic basin case study
- Achieving sustainability in transport and land use

Lastly, some projects intend to determine or assess sustainability indicators, e.g.:

- Operational indicators for progress towards sustainability
- Urban lifestyles, sustainability and integrated environmental assessment
- Development and application of soil productivity indexes for central America
- Implementation of EMAS (Eco Management and Audit Scheme) in Turkey
- Promoting action for sustainability through indicators at the local level in Europe



The review of Polish research related to sustainable development has been prepared on the basis of the reports done by the Polish Committee for Scientific Research. It turns out that of all projects carried out since the beginning of the Committee's work, only 8 - i.e. less than one per year on average - contain the word "eco-development" (which, in Polish, typically stands for sustainable development) in the title, namely:

- Analysing characteristics of red soils from the perspective of eco-development of rural areas
- An eco-development strategy for Zakopane and its impact zones including its recreational and sport functions
- Revitalising degraded settlement structures in the context of sustainable development (eco-development)
- Foundations of eco-development programming in local government units
- Environmental management in eco-development with particular emphasis on agriculture and forestry
- Natural science foundations of an urban eco-development theory and their practical implementation in spatial planning
- Conditions for the existence of organogenic soils in an agricultural landscape managed according to eco-development principles
- Environmental protection considerations in regional assessments for siting and modernisation of transport infrastructure according to eco-development principles. A methodology - a model design for the Cracow area

These projects are mostly rather small, with an average budget only slightly exceeding PLN 80,000 in current prices. Since all but two projects started in the latter part of the nineties, the budget may approximately be treated as if expressed in the 2000 prices. With such a budget it is possible to hire a team of several people for only a few months (in terms of full-time jobs). This proves that eco-development has not become a popular research topic in Poland.

A more detailed review of the 907 projects that deal with economics and various environmental protection domains, where the sustainability issue is likely to appear, does not alter this view significantly. Among those where this issue does appear, the majority are small projects limited to one single discipline. Large projects, especially ones that assemble teams from various disciplines, are very rare. Here are some examples of the latter category:

- Costs and benefits of public and private transport in cities
- Euroregions as a factor of economic development and environmental protection in border regions in Poland
- Sustainable development in indicators and statistical reporting systems
- Studies of recreational centres ("filters") clustering mass tourist flows in national parks of the Tatry region (Tatry, Babia Gora, Pieniny and Gorce National Parks)

Other projects that may be of interest are those connected with financial institutions as they combine advanced economic issues with sustainable development. Even though these projects are usually small, they are - potentially - very important given the role of money in the economy and in steering social processes.

- Banks and investment funds *versus* the environmental protection problem. Ecological criteria in financial transactions
- Pro-ecological reorientation of Polish banks in view of the UN declaration "Banking and environment"
- Using security market instruments in financing environmental investments

The majority, however, are narrowly outlined topics. It is very likely that the authors would not want their work to fall into the sustainable development research category; the Polish Committee for Scientific Research is far from encouraging such classification, anyway. As a result, Polish scientific research is nowhere near the model promoted in the EU.

Cases of partnership of Polish research teams in sustainability-related programmes financed by the EU are rare. "Baltic Drainage Basin Project" carried out in 1994-95 may serve as an example. Thanks to this study, Polish scientists were able to gain experience in environmental valuation in association with the leading centres in Sweden and England. It is nevertheless all too common for Polish sustainable development research teams to have very loose ties with international science.



#### 4. Recommendations for the future

Sustainability research comprises many disciplines. It requires the involvement of economists, ecologists, engineers and representatives of other sciences. In part 2 above, significant (as far as working out and promoting sustainability principles is concerned) research areas have been identified. These include: social sciences, especially economics, environmental sciences, especially ecology, and technical sciences, especially environmental engineering.

With so broadly outlined a research area, some key issues of particular meaning ought to be identified. This group should contain natural capital valuation research as it is irreplaceable at integrating different science disciplines. Such research is a good opportunity to develop the uncommon in Poland modern economic analysis techniques. Other than that, natural capital valuation research allows for a detailed recognition of the current role of particular natural capital elements, including the possibilities of its substitution. Last but not least, it forces people to seriously consider the future role of natural capital: the anticipated social preferences, technical abilities and ecosystem resilience. A truly responsible attempt to value a small part of natural capital requires the cooperation of economists, ecologists, engineers and sociologists.

Research on social aspirations and preferences as well as society's functioning mechanisms is also essential given its role in implementing sustainable development. This kind of a development path may not be decreed; it must result from a choice made on the basis of economic and political conditions but also from the awareness of natural capital inherited by the society and from the sense of intergenerational unity. Research projects in this field should assemble - alongside natural scientists - other humanities' representatives, including philosophy and pedagogy. It is the only way to competently analyse decision processes which encourage sustainability.

Looking at projects financed by the Polish Committee for Scientific Research from this perspective, it is clearly visible that they already include many of the above mentioned issues. What the projects lack is thorough economic research on sustainability fundamentals. Moreover, one may observe the dominance of narrow topics, limited usually to a small geographic region. The projects are mainly small and assemble teams of few scientists representing related disciplines.

The outlined diagnosis suggests the need for stronger ties between Polish and EU research programmes. This postulate is fortunately in agreement with the Polish government's political decision to allocate part of the science budget to international project selection procedures by participating in the Fifth and later in the Sixth Framework Programme. Independently of that, there are other, substantive reasons to encourage Polish research teams to cooperate with European centres more closely.

First of all, some scientific disciplines in Poland are still at initial development stages, which is why cooperating with leading foreign partners is highly desirable. Secondly, the fragmentary research conducted in Poland is valuable and often deserves wider popularisation, most easily achieved by preparing a synthesis in cooperation with a renowned research centre. Lastly, carrying out interdisciplinary research - i.e. the kind of research that is most desired in the field of sustainability - requires the involvement of large teams and, consequently, large budgets. Since the Polish science budget is very limited, and will certainly remain such for some time, there is no better strategy than taking advantage of international funds.

There is an additional reason why taking part in EU programmes can be attractive. While the selection of projects by the Polish Committee for Scientific Research is carried out by narrowly specialised internal commissions, the projects submitted to the European Commission have much higher chances of interdisciplinary assessment. And so, since sustainability research by its very nature requires an interdisciplinary approach, the authors of the best projects will receive kinder assessment by the Commission than by the Polish Committee for Scientific Research.

The research offer presented by the European Commission in the VI Framework Programme for the years 2002-2006 is very interesting from this point of view. The topics related to sustainable development may appear in different chapters of the Euro 12.5 billion budget undertaking; the most appropriate one, however, is the priority area entitled "Sustainable development and global change". The anticipated budget for this field is Euro 1.7 billion, i.e. nearly 14% of the total sum. Unlike Polish science administrators, the European Commission supplies its offer with rather detailed indications as to preferred research projects. Assuming that energy and transport are in the greatest conflict with sustainability, priority has been given to these areas.



The "Sustainable development and global change" area has been divided into two parts: *Technologies for sustainable development* and *Global change*. Five research priorities have been singled out in each of them, namely:

- Renewable energy sources, more efficient and clean use of energy, especially in urban areas, new concepts of energy efficient and cleaner transport
- Sustainable transport
- Fuel cells
- Hydrogen
- Solar photovoltaic technologies and biomass

and

- Impact and mechanisms of greenhouse gas emissions on climate and carbon sinks (oceans, forests and soil)
- Water cycle
- Biodiversity, protection of genetic resources, functioning of terrestrial and marine ecosystems and interactions between human activities and the latter
- Mechanism of desertification and natural disasters connected with climate change
- Global climate change observing systems

The priority topics above may at first seem narrow and technical. A closer analysis shows, however, that the proposed projects are expected to be complex and interdisciplinary. The introduction to the first part of the "Sustainable development and global change" area contains an indication that "proposed technological solutions [... should address] technical but also organisational, institutional, financial and social issues". Each topic is additionally described in a few sentences and supplemented by an indication as to what project authors should pay particular attention to. Here is an example of such indications concerning biodiversity: "research will focus on: assessing and forecasting changes in biodiversity, structure, function and dynamics of ecosystems and their services; relationships between society, economy, biodiversity and habitats; integrated assessments of drivers affecting biodiversity and mitigation of biodiversity loss; risk assessment, management, conservation and rehabilitation options.

Independently of research projects focused on fundamental sustainability issues, there is demand for scientific analysis of minor issues with which particular economic sectors or local governments have to deal. Such cases ought also to be seriously treated though few of them will succeed in securing funds from a science budget, be it Polish or European. In Poland, there are a number of institutions that can provide financial support for research projects of this type. The most important one of these institutions is the National Environmental Protection and Water Management Fund.

The basic requirement for the success of sustainability-devoted research is ensuring cooperation of specialists in different fields who represent the highest scientific level. The need for an interdisciplinary approach is hard to challenge. The quality of such interdisciplinary cooperation, however, is sometimes problematic. In a team of, for example, an economist, ecologist and engineer we have interdisciplinarity but the team members may not be equally competent. If the results of their work are discussed by a group consisting of representatives of one single discipline, e.g. ecology, the quality of the results cannot possibly be properly assessed; in this case the economist's and engineer's role may be assessed very superficially while the rank of the results will remain unclear. Therefore, providing competent support for sustainability research requires great intellectual and organisational effort related to project assessment.



## Literature

- Daly, H. E. 1990, "Sustainable Development: from Concept and Theory to Operational Principles", in: K. Davis and M. S. Bernatam (eds.), *Resources, Environment and Population*, New York: Population Council and Oxford University Press, pp. 25-43
- Folmer, H., L. Gabel, H. Opschoor (eds.) 1996, *Ekonomia œrodowiska i zasobów naturalnych*, Wydawnictwo Krupski i S-ka, Warsaw
- Hartwick, J. M. 1977, "Intergenerational Equity and Investing the Rents from Exhaustible Resources", *American Economic Review* 66, pp. 972-974
- Pezzey, J. 1989, *Economic Analysis of Sustainable Growth and Sustainable Development*, Environment Department Working Paper No. 15, The World Bank, Washington, D.C.
- Rawls, J. 1971, *A Theory of Justice*, Harvard University Press, Cambridge, Mass.
- WCED, World Commission on Environment and Development 1987, *Our Common Future*, Oxford University Press, Oxford
- Zylicz, T. 2000, *Costing Nature in a Transition Economy. Case Studies in Poland*, Edward Elgar, Cheltenham
- Zylicz, T. 2001, "Wiara i nauka a sprawiedliwoœæ miêdzypokoleniowa", in: A. Bia³ecka and J. J. Jadacki (eds.), *U progu trzeciego tysiąclecia. Cz³owiek - Nauka - Wiara*, Wydawnictwo Naukowe Semper, Warsaw, pp. 163-170

Information available at the following internet sites has also been used:

[www.kbn.gov.pl](http://www.kbn.gov.pl)

and

[www.cordis.lu](http://www.cordis.lu)





## Ernst Max Nielsen

### SUSTAINABILITY IN CEE

#### Notes View excerpts

My presentation is going to look at Sustainability in Central and Eastern Europe (CEE) and raise different questions

1. What is sustainability? And why do we measure "it"?
2. Which indicators are used to measure sustainability in the "West"
3. CEE is adopting "Western" perspectives and methods: what are the issues involved? Is the transport policies of CEE governments sustainable? How about the Common agricultural policy in CEE? These are some of the interesting questions we should answer.
4. Some indicators are easily measured quantitatively by using scientific methods ("number crunching") but most of the contested ones can only be measured by qualitative methods. This raises a question about how to do that and who does it. The answer lies in management of multi-stakeholder dialogue processes.

The concept of Sustainability is actually a term which has come into use only recently and was seemingly first raised by the Bruntland Commission as a rather broad concept of philosophical nature (diplomacy rules!). In recent years many governments, NGOs and corporate businesses seem to have developed the Bruntland concept to the so-called Triple Bottom Line (TBL), "Bottom Line" of course referring to the traditional economical view of a "sustainable" business entity (ie. Capable of survival or a "going concern"). The two new dimensions added are an environmental and a social scoreboard or bottom line, measured by certain indicators. You can read some of the criteria in the three dimensions on the slide.

The Sustainability Concept is, however, not a value-free concept. It is a concept which makes sense only in a Christian/Jewish philosophical setting after the advent of industrial capitalism, where the "poor and dirty" nations (according to Amory Lovins and Ernst Ulrich v. Weizsäcker "Factor 4") go through a phase of being "rich and dirty", now wishing to become "rich and clean". In a basic sense, the term is central to eastern thinking e.g. in the concept of Tao or yin-yan.

The recent governmental efforts stem from preparation of participation in the "Rio+10" Johannesburg World Summit on Sustainable Development (WSSD). Many NGOs, both from civil society and from the corporate world have tried to interpret the Triple Bottom Line (WBCSD, WE Forum, IISD, REC, GRI to mention a few). And finally, some 120 multinational corporations have adopted a Triple Bottom Line approach to reporting.

It is also worth mentioning that many institutional investors have adopted a TBL approach. Even Dow Jones, the Index for corporate shares, has developed a Sustainability Index. And Portfolio Companies, such as the Swiss SAM Group develop an own Index and portfolio for clients.



## Why Sustainability?

Many governments from the "rich and clean" countries have political reasons - stakeholder management!?!- to show they are supporting development of the "dirty and poor", so there is a North-South dimension to the rise of the sustainability concern.

From the corporate side, it seems now clear that sustainability adds not only stakeholders' value but certainly shareholders' value. If you look on the chart, you'll see the added value of the Dow Jones Sustainability Index compared with the General Index. This means: the yellow line shows the development of the SI, whereas the blue line shows the development of the GI. Investors clearly trust more in businesses, which have a declared policy of sustainable development. Think of some of the examples, you know:

Philip Morris and the US cigarette industry

Shell and Brent Spar

Nike and child labour

The Mining Industry and Tisza River on the border between Romania and Hungary

I'll talk in some more depth about the SI.

This graph was made by M. Porter for the World Economic Forum (Davos) in 2000. On the x-axis it shows the ranking of countries with regard to the overall competitiveness of their industry. On the Y-axis, an indicator for Environmental regime has been made, i.e. indicating how severe environmental regulations prevail in a country and how severely these regulations are enforced.

The graph clearly shows a correlation (trend), i.e. that the most competitive industries are to be found in countries with the toughest environmental regime. If the anti-environmentalist lobby were right, this relationship would not be possible, as more regulations would supposedly cause mainly more expenses and less competitiveness.

Furthermore, we can see that selected CEECs are placed relatively low on the scale.

This drawing explains quite well the expanded criteria of the Triple Bottom Line, namely

1 we don't only look at the individual legal entity, the incorporation, but also at the supply chain and stakeholders (Nike eg.)

2 the traditional economical view is annual; now the perspective is multi-annual

3 and the information used is not only financial, but also very qualitative, such as the perception of the business in the public or by certain NGOs (Shell and Greenpeace)



I wish here to especially attract attention to the "stakeholder" concept and clarify: we often talk of multistakeholder dialogue and the methods to manage these processes are unscientific at best or clearly political.

Example: Bellagio Forum for Sustainable Development ([www.bfsd.org](http://www.bfsd.org)) and "the fighting in the streets"

What I wish to emphasize is this: we can collect and evaluate as many "hard" data indicators as we wish from a scientific point of view; their value will be interpreted according to agendas set in the political sphere. Even the hardest data are contested (Sustainable Mobility as example).

Example: EU's Common Agricultural Policy and the ESTO Enlargement project, [www.jrc.org](http://www.jrc.org)

Here is another chart showing how SAM (in cooperation with Dow Jones) evaluates their portfolio companies using the TBL approach. The evaluation is made according to a (web) questionnaire which is based on self-reporting. [www.sam-group.ch](http://www.sam-group.ch))

The trend of SAM is to actually independently monitor and check the validity of the information given.

SAM tried to establish a CEE portfolio, but the liquidity of the businesses in CEE traded on CEE stock exchanges, is too low, it seems. The attempt was given up.

Now, let us look at some governmental initiatives, like the UK Sustainability Strategy. The Scandinavian, Dutch, German and other governments have published similar strategies.

They often contain "pretty talk" about goals and then for each goal an indicator has been selected.

The examples I have chosen are clear:

GDP is an objective measure, it seems

But level of crime is certainly not. It was the German author, Hans Magnus Enzensberger, who showed us that the dividing line between crime and legal activity is defined by politics. Take the very different levels of crime in CEE and e.g. Denmark: Denmark is numerically much higher than e.g. Hungary but the Hungarians feel that crime is a serious issue because the crime rate rises. We also have examples of certain policing efforts which can raise the level in one crime domain and not in another.

Road traffic seems an objective measure, but is it? The "modal split" in CEE still gives a 50% preference for collective urban traffic, whereas EU can show a meagre 18%. In the near future, CEE citizens are going to buy more cars, but still stay well below the Danish level.

And finally, "satisfaction with life" is normally measured by respondents answering a questionnaire, often given via the telephone. I haven't seen any questions capable of getting to the core of what people (I) really think.



How is the situation in CEE? First of all, let me play down the homogeneity of the region. CEE consists of 15 different countries plus more Balkan candidates. There are at least three different sub-regions with specifically different histories, present trends, cultures etc: The Baltic Region, South East Europe (Balkan), and the Central Region.

In the Communist Era, Sustainability was not an issue or it was demeaned as a "Western trick" to show that the rich imperialist nations and governments are doing something to counter the economic effects of exploitation. Now the agenda is clearly post-communist, by which I mean that many efforts are taken to reverse that former period. This gives a certain preference for such policies as sustainability to become popular among politicians.

But in most ways, Accession to EU has effectively set the agenda and yardstick for most of the indicators. In the field of environment, for instance, this is a real pity, as CEE has a potential to use the "forwardness of being backward", i.e. Jump across the levels of development which are present in EU today, so as to set an example for the EU, but NO, the environmental acquis is the measure of everything in these years: "just reach the level of allowed pollution in EU and you'll be sustainable", seems to be the message. A pity!

A frequent concern has been that FDI in CEE would be a move of the most polluting industry from EU and USA to CEE, but this has not (always) been the case, so FDI becomes a driver of change in the right direction. We can thank several Cleaner Production initiatives for that

On the governmental and NGO level, many initiatives are taken:

Here are some examples of initiatives which have also been developed in CEE:

UNEP has supported governmental and NGO input to the Rio+10 process, [www.unep.org](http://www.unep.org)

The ABEI (Aarhus Business and Environment Initiative) has identified and worked with stakeholders in CEE about support to cleaner production and environmental management initiatives. This included the social agenda. [www.rec.org](http://www.rec.org)

The IISD (International Institute for Sustainable Development [www.iisd.org](http://www.iisd.org), headed by former REC director Petr Hardi) and the World Economic Forum ([www.weforum.org](http://www.weforum.org)) have developed their Indicator projects: Dashboard and Global Competitiveness (by the way, this was originally performed by Harvard University and Prof. Porter, but has now been taken over by Yale, after a row over the validity of indicators).

The Austrian NGO, ÖGUT, has developed an initiative with Henkel to support sustainable businesses. REC (with the support of Procter&Gamble!) made a similar initiative and also manages a "CEE sustainable cities" programme (the name was changed, prudently, to become the Award for Cities towards Compliance" (ie. to EU Accession criteria): actually, there is an unsettled dispute in EU who "owns" the sustainability agenda.

REC has participated in several national initiatives, eg. Slovakia, Slovenia and Hungary

The biggest issues for evaluation of sustainability in CEE concern the reliability of the indicators, not only quantitative but certainly the qualitative:

To me it is clear that the effective strategy is to copy the "Western" strategy even though the "West" and CEE have clearly different points of departure. From experience we know that the statistical offices of CEE governments are not always reliable sources of updated information.

I also don't feel convinced about the universality of the indicators. 1 meter is 1 meter wherever we measure; but as I have shown many other indicators are not unambiguous.

As for the qualitative indicators, they need to be evaluated and reported by multi-stakeholders, among them governments and independent NGOs and institutions, but such organisations hardly exist, nor have they fully developed their institutional capacity to "deliver the goods" - yet.

One can say that the single most important contribution in support of sustainable development in CEE is the development of such a capacity; which is exactly the remit and action program of REC.



Cont. from prev. page:

How about the importance of the difference of cultures measured by the same indicators? For instance, the existence of social networks such as families that stay together and support each other is much higher in CEE than in EU. Or the use of cars is measured against two totally different physical/urban/regional infrastructures.

Denmark is often seen as the most un-corrupt nation in the world, whereas many CEE countries are seen as being corrupt. But the concept of corruption is not unambiguous. In CEE, say in the Balkans, there is a tradition of "master-slave" and nepotistic relationships, which stem - among other sources- from Ottoman traditions (Turkey is extremely "corrupt"), but is this not a cultural phenomenon?

Environmental inspection as example: as fees and fines are given by environmental inspectors based on reported levels of pollution, there is an incentive to omit reports, in order to save local jobs.

Example of life cycle analysis (ISO 14040) difficulties of measuring the environmental impact of different substances (who decides what is worst: 0.5 mg/kg copper from the circuit print board of a scrape television set or the number of kg of CO<sub>2</sub> emitted from a power station to power its "standby" button for the remote control system?).

Example of debate between WBCSD and traffic NGOs in CEE (Sustainable Mobility project): [www.wbcd.org](http://www.wbcd.org)



## Thomas Widmer

### EVALUATION STANDARDS IN THE CONTEXT OF SUSTAINABILITY

#### 1. Evaluation Standards

In order to understand the thinking that underlies evaluation standards, it's a good idea to keep in mind what we do when we make a judgment or an assessment. Expressed simply, we relate a given object to a specific criterion. Obviously, defining the appropriate criterion is a crucial task in evaluation. If our goal is to judge the quality of an evaluation, we must choose the appropriate criterion to make this judgment. Because evaluation quality is a multidimensional concept, we need not only one, but several evaluation criteria that represent the expectations that evaluations must fulfill.

##### 1.1. Reference Levels for Evaluation Quality

One can now differentiate the demands made on evaluation according to two different assessment criteria, namely internal and external, and use them to assess the quality of an evaluation. Table 1 provides an overview of the respective reference levels with their corresponding assessment bases:

Table 1: Reference Levels (Overview)

Level	Object	Basis for Internal Assessment Criteria	Basis for External Assessment Criteria
First Order Constructs	Program, Project, Measure, etc.	Objectives of the Program, Project, etc.	Social Significance of the Program
Second Order Constructs	Evaluation	Objectives of the Evaluation	Evaluation Theory and Methodology
Third Order Constructs	Metaevaluation	Objectives of the Meta-evaluation	....

An evaluator can thus assess a program (or another evaluandum; a first order construct) based on the program objectives (internal assessment criteria). But an evaluator can also assess a program based on the social significance of the program (external assessment criteria). The same reasoning can be applied to an evaluation as a second order construct, where the relevant assessment criteria would therefore first be the objectives of the evaluation (internal assessment criteria) and second general evaluation theory and methodology (external assessment criteria).

By analogy, this argument can be pursued at the third (or subsequent) order, and if one undertakes an assessment (or evaluation) of an evaluation, then one speaks of a *metaevaluation*. But in answer to the question who evaluates a metaevaluation, I agree with Michael Scriven who has written that: "no infinite regress is generated because investigation shows it usually doesn't pay after the first metalevel on most projects and after the second on any" (1991:230).

At this point, since I have used the term metaevaluation, let me say something about wording in order to avoid confusion.



The term *metaevaluation*, noted above, needs elucidation as well as demarcation from other approaches such as *evaluation syntheses* or *meta-analysis*.<sup>2</sup> The latter two instruments are derived from the substantive contents or results of evaluations; the basis for the investigation is the thematic evidence the evaluations themselves can provide. A *meta-analysis* requires a sufficient number of existing evaluation studies to permit a quantitative answer to be given in answer to a specific question. This quantitative focus typically makes a *meta-analysis* more narrowly oriented than an *evaluation synthesis*. An *evaluation synthesis*, though it requires a thematic clustering among the evaluation studies upon which it is based, is, in its qualitative approach, far less narrowly circumscribed.

Table 2: Overview of Third Order Evaluation Instruments

Instrument	Description
Evaluation Synthesis	Content Synthesis of Various Evaluation Studies (largely qualitative) (global evaluation/cross-sectional analysis)
Meta-Analysis	Quantitative Integration of the Results of Various Evaluation Studies (research synthesis)
Metaevaluation	Evaluation of Evaluation(s): Systematic Assessment of the Quality of one or more Evaluation Studies

A *metaevaluation*, by contrast, has a fundamentally different goal, as it is an evaluation of one or more evaluations that intends to systematically establish the worth and merit of evaluation(s). In other words, a *metaevaluation* assesses the quality of an evaluation and for that it needs assessment criteria. As previously noted, both evaluation objectives (internal assessment criteria), and evaluation theory and methodology (external assessment criteria) need to be taken into account for this, and evaluation Standards allow for a more precise formulation of the external assessment criteria. As is true for evaluations themselves, metaevaluations can be fashioned in quite different ways: as self-evaluations or heteronomous evaluations, executed internally or externally, or they can fulfill formative or summative functions.

Let me now return to the issue of assessment criteria and reference level: The relevant point here is that if we want to investigate the quality of an evaluation, we essentially have two possible assessment criteria, the internal and the external. Because it is fundamentally impossible to generalize about objectives that are set for and in specific evaluations, we instead focus on the external assessment criteria for evaluations, that is, on evaluation theory and methodology.

## 1.2. Evaluation Standards

### 1.2.1. The Origin of the Standards

As part of the professionalization of evaluation, various parallel initiatives were launched in the U.S. during the late 1970s to define a set of criteria applicable to the many dimensions of evaluation quality.<sup>3</sup> The intent was to establish criteria capable of assessing the quality of an evaluation with respect to a general theory and methodology of evaluation – in effect to craft standards that would be external assessment criteria. Depending upon author, such parallel definitional efforts were sometimes based on thematic areas of evaluation, and sometimes on varying cognitive or epistemological basic premises. Various lists of suggested criteria were published at the time, many marked by more or less limited epistemological orientations.

The publication in 1981 of the "Standards for Evaluation of Educational Programs, Projects and Materials" by the "Joint Committee on Standards for Educational Evaluation" (henceforth "JC Standards"), was, by contrast, characterized by a comparatively open methodological orientation. This was fortunate, since the U.S. evaluation community was embroiled in a long and heated controversy during the 1980s over the question

<sup>2</sup> On this, and in following, see the classic texts of Scriven (1969), Stufflebeam (1974) and Cook/Gruder (1978), and, more recently, Stufflebeam (2000).

<sup>3</sup> On the process of professionalization, see the two five-step models of Harold Wilensky (1964), and Dreyfus/Dreyfus (1986) as well as the discussion and literature noted in Altschuld (1999a: 483-6), who refers directly to evaluation.



whether qualitative methods - and not just classical, quasi-experimental orientations or understandings of evaluation - were acceptable. Due to its methodological and epistemological openness, the JC Standards anticipated the resolution of this controversy, namely that the use (or existence) of alternative methodologies is (generally) acceptable.

The JC Standards are directed primarily at those who deal with education evaluation, and most of the organizations represented in the "Joint Committee" come from education. While the Standards have been widely adopted in American evaluation practice, it has increasingly been in realms and topic areas far removed from their originally intended application to educational evaluation. This development is reflected not only in the new name for the Standards themselves (reissued in a revised version in 1994 as "The Program Evaluation Standards" with the original connection to education only visible in the subtitle "How to Assess Evaluations of Educational Programs"; see Joint Committee 1994) but also in the fact that the organizations represented in the Joint Committee go well beyond those only concerned specifically with education (see the overview in Widmer/Beywl 2000: 250). Thus the name change only completes what, in my view, has already long changed in practice, namely that the Standards are widely employed outside of educational contexts.<sup>4</sup>

The JC Standards were not given much attention in German-speaking Europe until the mid-1990s (Widmer 1996a-h, Widmer/Rothmayr/Serdült 1996; for a relatively early exception, see Beywl 1988: 113-23), despite the gradual increase in the use of evaluations during the 1980s. To ease access to the Standards for German speakers, a translation was finally published in 1999 (Joint Committee 1999), and the fact that it soon had to be reprinted in a second edition testifies to how quickly these Standards were taken note of and used (Joint Committee 2000; see also Beywl/Taut 2000).

In its efforts to professionalize evaluation in Switzerland, the Swiss Evaluation Society SEVAL gave the task of developing evaluation standards appropriate to Switzerland to one of its Working Groups.<sup>5</sup> This Working Group, after carefully considering a wide variety of existing models and standards (see Beywl/Widmer 2000), decided to develop country-appropriate evaluation standards, using the basic structure of the JC Standards but revising and adapting specific individual Standards.

A first version of this effort was greeted with great interest in the spring of 1999 by the SEVAL General Meeting, and after the Working Group reworked and submitted a revised version, the SEVAL General Meeting in the spring of 2001 approved what are now known as the SEVAL Standards (Widmer/Landert/Bachmann 2000). In so doing, SEVAL acted as a pioneer, becoming the first European evaluation society to establish its own evaluation standards. As Table 3 indicates, Switzerland is not an international exception; rather, it is one example of the many efforts being made by national and regional evaluation societies to improve quality.

---

<sup>4</sup> The former chairman of the Joint Committee has recently assessed a whole series of general evaluation models (Stufflebeam 2001).

<sup>5</sup> This Working Group included members from quite different realms, including private sector evaluation firms, state administrative offices, and universities.





Table 3: The Development of Standards and Guiding Principles (selection)<sup>6</sup>

Organization	Guidelines and Standards	Source
African Evaluation Society	African Evaluation Guidelines (adapted from the Program Evaluation Standards of the Joint Committee)	<a href="http://www.geocities.com/afreval">www.geocities.com/afreval</a>
American Evaluation Association (AEA)	Guiding Principles for Evaluators and Program Evaluation Standards (Joint Committee)	<a href="http://www.eval.org">www.eval.org</a> (see AEA 1995)
Australasian Evaluation Society (AES)	Guidelines on Ethical Conduct of Evaluation and Program Evaluation Standards (Joint Committee)	<a href="http://www.aes.asn.au">www.aes.asn.au</a>
Canadian Evaluation Society (CES)	CES Guideline for Ethical Conduct	<a href="http://www.evaluationcanada.ca">www.evaluationcanada.ca</a>
German Evaluation Society (DeGeval)	Evaluation Standards (DeGeval-Standards) (adapted from the Program Evaluation Standards (Joint Committee) and the SEVAL-Standards)	<a href="http://www.degeval.de">www.degeval.de</a> (see Beywl 2000 and Beywl/Taut 2000:367)
French Evaluation Society	Evaluation standards in preparation	<a href="http://www.sfe.asso.fr">www.sfe.asso.fr</a> (see Perret/Barbier 2000)
Italian Evaluation Society	Guiding Principles	<a href="http://www.valutazioneitaliana.it">www.valutazioneitaliana.it</a>

### 1.2.2. The SEVAL Standards: Structure and Content

The SEVAL Standards are based on the premise that an evaluation should at once be useful, feasible, proper, and accurate so as to fulfill the demands placed on it: good evaluations should therefore demonstrate all these characteristics. To make these category characteristics more tangible, the SEVAL Standards are subdivided into 27 individual Standards that fall into one of the four larger categories (Widmer/Landert/Bachmann 2000):

**N Utility**  
*The utility standards are to guarantee that an evaluation is oriented to the information needs of the intended evaluation users.*

#### **N1 Stakeholder Identification**

Those participating in, and those persons affected by, an evaluation are identified in order that their interests and needs can be taken into account.

#### **N2 Clarifying Evaluation Objectives**

All persons who take part in an evaluation are concerned that the objectives of the evaluation are made clear to all stakeholders.

#### **N3 Credibility**

Those who conduct evaluations are both competent and trustworthy; this helps ensure that the evaluation results are accorded a maximum amount of acceptance and credibility.

<sup>6</sup> For further comparisons between quality guidelines and standards that are employed internationally, see Widmer/Beywl 2000.



#### **N4 Scope and Selection of Information**

The scope and selection of the information collected permits pertinent questions to be asked of the evaluandum while simultaneously taking the interests and needs of the party contracting the evaluation, as well as other stakeholders, into account.

#### **N5 Transparency in Assessment**

The underlying reasoning and point of view used to arrive at the interpretation of the results are described in such a manner that the bases for the value judgments are clear.

#### **N6 Comprehensiveness and Clarity of the Report**

Evaluation reports describe the evaluandum, including its context, objectives, the questions asked, the procedure(s) used, and the results of the evaluation. This ensures that the pertinent information is made available and can be readily understood.

#### **N7 Timeliness in Reporting**

Significant interim results and (the) final report(s) are made available to the intended users in such a fashion that they can be used in a timely manner.

#### **N8 Evaluation Impact**

The planning, execution, and presentation of an evaluation encourages stakeholders both to follow the process of evaluation and utilize the evaluation.

### **D Feasibility**

*The feasibility standards ensure that an evaluation is conducted in a realistic, considered, diplomatic and cost effective manner.*

#### **D1 Practical Procedures**

Evaluation procedures are designed in such a manner that the necessary information is collected without disrupting the evaluandum or the evaluation.

#### **D2 Political Viability and Support**

Evaluations take the various positions of different interests into consideration in their planning and execution in order to win their cooperation and discourage attempts by one or another group to limit evaluation activities or distort or misuse the results.

#### **D3 Cost Efficiency**

Evaluations produce information of a value justifying the financial resources expended.

### **K Propriety**

*The propriety standards ensure that an evaluation is conducted in a legal and ethical manner and that the welfare of stakeholders is given appropriate attention.*

#### **K1 Formal Agreements**

The obligations of the parties to an evaluation (specifying what, how, by whom and when what is to be done) are set forth in written agreements that obligate the parties to fulfill all the conditions of the agreements, or if not, to renegotiate these agreements.

#### **K2 Protecting Individual Rights**

Evaluations are planned and executed in such a manner as to protect and respect human rights and welfare.

#### **K3 Human Interaction**

Evaluations are laid out in such a manner that contacts between participants are characterized by mutual respect.

#### **K4 Complete and Fair Assessment**

In their assessment and depiction of strengths and weaknesses in the evaluandum, evaluations are both complete and fair, such that strengths can be built on and problem areas addressed.

#### **K5 Disclosure of Findings**

The contracting parties to an evaluation ensure that the evaluation results are made available to the affected persons as well as to all those who have a legitimate claim to receive these results.

#### **K6 Declaring Conflicts of Interest**

Conflicts of interest are addressed openly and honestly in a manner that they will least compromise evaluation processes and conclusions.



**G Accuracy**

*The accuracy standards ensure that an evaluation produces and disseminates valid and usable information.*

**G1 Documentation of the Evaluandum**

The evaluandum is clearly and precisely described and documented so that it can be unambiguously identified.

**G2 Analysis of Context**

The influences of the context on the evaluandum are identified.

**G3 Description of Objectives and Procedures**

The objectives, questions, and procedures used in and by the evaluation are described and documented with sufficient precision that they are identifiable and can be judged.

**G4 Trustworthy Information Sources**

The sources of information used in an evaluation are sufficiently precisely described that the appropriateness of the information can be assessed.

**G5 Valid and Reliable Information**

The procedures for collecting information are so chosen or developed, and then employed, that the validity and reliability of the interpretation arrived at is ensured for the given purpose.

**G6 Systematic Information Checks**

The information that is collected, analyzed, and presented in an evaluation is systematically checked for errors.

**G7 Qualitative and Quantitative Analysis**

Qualitative and quantitative information in an evaluation are systematically and appropriately analyzed so that the questions posed in the evaluation are actually answered.

**G8 Substantiated Conclusions**

The conclusions reached by an evaluation are explicitly substantiated in a manner that stakeholders can follow and that permits them to assess the conclusions.

**G9 Objective Reporting**

Reporting is free from distortion through personal feelings or preferences for one or another party to the evaluation; evaluation reports fairly reproduce the results.

**G10 Metaevaluation**

The evaluation itself will be evaluated on the basis of existing (or other important) standards so that the execution can be done accordingly and so that stakeholders can in the end assess the strengths and weaknesses of an evaluation.

Beyond this list, the SEVAL Standards also include a general introduction (with information about the objectives, scope, and addressees of the Standards), further explication of each individual Standard, and various supporting materials (functional overview, creation process, etc).<sup>7</sup>

As can readily be seen, the number of individual Standards in each category varies; feasibility, for example, has only three while accuracy has ten individual Standards. This should not be misunderstood as implying a *weighting* among the four categories. Rather, there is a deliberate wish to avoid weighting either categories or individual Standards, because the significance of any specific Standard can only be determined in each individual case. A generalized weighting that made claims to be valid across cases thus would be inappropriate.

The question of *weighting* is particularly important because the SEVAL Standards contain demands that are, at least in part, at odds with one another. In practice, in fact, evaluation is often confronted with the question which Standard should enjoy prominence of place. But this lack of internal consistency in the SEVAL Standards should not by any means be seen as weakness, since it reflects the (previously described) tension-ridden and conflictual arena in which evaluations move.

<sup>7</sup> The complete SEVAL Standards can be found, in German and French, on the SEVAL homepage (<http://www.seval.ch/>); Italian and English versions are in preparation.



Put differently, the SEVAL Standards are maximum demands - not minimal standards of what is an absolute must but rather a statement of what a good quality evaluation should ideally try to achieve. In practice, it will also rarely ever be possible to completely fulfill every one of the 27 individual Standards. Nevertheless, all participants, and not just the evaluator, should try to take the SEVAL Standards into account as much as possible. With the SEVAL Standards, evaluation has an instrument available to it that describes, in a precise and sufficiently differentiated manner, what the demands of quality are.

## 2. Evaluation with sustainability

Some weeks ago, when I told a few colleagues of mine that I would be participating in a conference on the Evaluation of Sustainability, they asked me: "Sustainability of what?" I couldn't give them an answer then, but perhaps this conference will help me to do so.

Sustainability is not an object of evaluation. Sustainability is a quality, a capacity – an evaluation criterion. So it makes no sense to evaluate sustainability itself. But it does make sense to evaluate something, an object, a program or a public policy, using sustainability as an evaluation criterion.

Sustainability is a non-trivial concept, which – in contradiction with what the title of this conference suggests – is not so easy to deal with empirically than most other evaluation criteria. The concept of sustainability is multidimensional. Indeed, we face at least three dimensions of assessment: the economic, social and ecological dimensions. Some scholars add a fourth dimension, namely the quality of the political process (Thierstein 1998, Widmer/Schenkel/Hirschi 2000). It is thus unavoidable to use a multicriteria approach for assessment.

But this is not the only problem we are confronted with when trying to evaluate the quality of sustainability. The core problem lies elsewhere. The concept of sustainability has other characteristics that make it difficult to use as a measure to assess a policy, a program, a project or any other object. In particular, the following three aspects make the evaluator's life difficult:

1. the long-term perspective of the concept of sustainability,
2. uncertainty as a crucial component of the concept, as well as
3. the fact that the concept of sustainability is holistic, thus refusing to restrict itself to one specific aspect.

The problem with the *long-term perspective* of the concept lies in the difficulties connected with investigating future situations empirically. How do you measure the future development or the future state of affairs? It is certainly possible to develop expectations about the way in which current behavior will impact on the future. But it is impossible to avoid making assumptions when trying to forecast the future, assumptions that are not empirically testable and that may be wrong.

The second problem - *uncertainty* - bothers the evaluator even more than the first one does. People familiar with risk assessment or with a basic understanding of random theory will follow me on this one. The main question is the following: how do you assess the significance for the future of knowledge about risks elaborated in the present? Estimating the incidence of a specific risk is a hard task in itself. Assessing its relevance for the present is difficult. Assessing its relevance for the future is almost impossible.

Third, the *holistic* nature of sustainability, referred to in German as "Ganzheitlichkeit", forces the evaluator to give an overall, comprehensive judgment. A sound empirical investigation, however, needs definitions to focus on. Investigating the whole world is not possible in one evaluation.

Many discussions on the subject of sustainability systematically ignore these three aspects. This is the case especially when people discuss indicators (or the measurement) of sustainability. Even in the field of evaluation this is normally the case. In my view, this reduces the complexity of the concept of sustainability too much. I am tempted to call this practice a corruption of the underlying assumptions of sustainability.

I conclude this discussion about the evaluation *with* (and not *of*) sustainability with the statement that to take sustainability seriously means to avoid using it as an overarching evaluation criterion. Would that mean that we have to stop using sustainability as an evaluation criterion in general?



### 3. What conclusions can we draw from all of this?

If you have a positive image of both evaluation and sustainability – and as attendees of this meeting, I suppose you do! –, how do you manage the dilemmas outlined? Let me make some propositions on how to approach the problem:

- The first proposition would be that we should be more honest and acknowledge the restrictions we are confronted with when using the concept of sustainability in evaluation. This would include the more sensitive use of the term sustainability, even though the concept it is now quite trendy.
- As I already mentioned, we need an object to evaluate. Sustainability is not an object. Therefore, let's not try to evaluate sustainability. Let's evaluate specific objects with sustainability as an evaluation criterion.

In this context, I would suggest having a look at the SEVAL Standards. They might help solve some of the problems.

- Evaluation Standards teach us something about the multicriteria problem I mentioned. The Standards formulate 27 criteria to judge evaluation quality. The Standards reflect the different demands that the evaluation is confronted with. They sometimes conflict; they sometimes are not strictly separated from one other; some are highly interdependent. The situation is quite similar for the multifaceted concept of sustainability. But in the SEVAL standards, and this is my point, no weighting of the criteria exists. In the case of sustainability, this would mean that we should avoid calculating aggregated values, consisting of values of indicators of the different dimensions of sustainability. We should restrict ourselves to a more modest level of judgment, separated for the different dimensions of sustainability.
- A second proposition could thus be to use the SEVAL Standards as a tool to assess the quality of evaluations in the context of sustainability. With the Standards, we are able to distinguish the better from the worse and we are in a well-founded position to judge the quality of the results produced by the different studies. This would be a contribution to sustainability in itself.

Evaluation Standards cannot make miracles happen and – I must be modest on this point too – they can't solve all the existing problems. But in my view, this is also a benefit. If evaluation standards gave us all the answers to all of our problems, then there would be no need to continue the EASY-ECO workshops.



## References

- Altschuld, James W. (1999): The Certification of Evaluators. In: *American Journal of Evaluation* 20(3): 481-93.
- American Evaluation Association, Task Force on Guiding Principles for Evaluators (1995): Guiding Principles for Evaluators. *New Directions for Program Evaluation* 66: 19-26.
- Beywl, Wolfgang (1988): Zur Weiterentwicklung der Evaluationsmethodologie. Frankfurt: Lang.
- Beywl, Wolfgang (2000): Standards for Evaluation: On the Way to Guiding Principles in German Evaluation. *Evaluation Center Occasional Papers Series*, Vol. 17. Kalamazoo, Western Michigan University: 60-5.
- Beywl, Wolfgang/Taut, Sandy (2000): Standards: Aktuelle Strategie zur Qualitätsentwicklung in der Evaluation. *Vierteljahreshefte zur Wirtschaftsforschung* 3/2000: 358-70.
- Beywl, Wolfgang/Widmer, Thomas (2000): Die 'Standards' im Vergleich mit weiteren Regelwerken zur Qualität fachlicher Leistungserstellung. In: Joint Committee on Standards for Educational Evaluation (Hrsg.): *Handbuch der Evaluationsstandards*. Opladen: Leske+Budrich: 259-95.
- Cook, Thomas D./Gruder, C.L. (1978): Metaevaluation Research. *Evaluation Quarterly*, 2: 5-51.
- Dreyfus, Hubert L./Dreyfus, Stuart E. (1986): *Mind over Machine*. New York: Free Press.
- Joint Committee on Standards for Educational Evaluation (1981): *Standards for Evaluations of Educational Programs, Projects, and Materials*. New York: McGraw-Hill.
- Joint Committee on Standards for Educational Evaluation (1994): *The Program Evaluation Standards*. Newbury Park: Sage.
- Joint Committee on Standards for Educational Evaluation (Hrsg.) (1999): *Handbuch der Evaluationsstandards*. Opladen: Leske + Budrich.
- Joint Committee on Standards for Educational Evaluation (Hrsg.) (2000): *Handbuch der Evaluationsstandards*. 2. Auflage. Opladen: Leske + Budrich.
- Perret, Bernard/Barbier, Jean-Claude (2000): Ethical Guidelines, Process and Product Quality Standards, What For?. Paper presented at the EES Conference, Lausanne, October, 12<sup>th</sup> – 14<sup>th</sup> 2000.
- Scriven, Michael (1969): An Introduction to Meta-Evaluation. *Educational Product Report* 2(5): 36-8.
- Scriven, Michael (1991): *Evaluation Thesaurus*. Fourth Edition. Newbury Park: Sage.
- Stufflebeam, Daniel (1974): *Meta-Evaluation*. Evaluation Center Occasional Paper Series, Vol. 3. Kalamazoo, Western Michigan University.
- Stufflebeam, Daniel L. (2000): The Methodology of Metaevaluation. In: Stufflebeam, Daniel L./Madaus, George F./Kellaghan, Thomas (eds.): *Evaluation models*. Boston: Kluwer: 457-71.
- Stufflebeam, Daniel L. (2001): *Evaluation Models*. *New Directions for Evaluation*, 89.
- Thierstein, Alain (1998): *Raumordnung und nachhaltige Entwicklung: Handlungsansätze für eine nachhaltige Raumentwicklung in der Schweiz..* Hrsg. vom Bundesamt für Raumplanung. Bern: EDMZ.
- Widmer, Thomas (1996a): *Meta-Evaluation: Kriterien zur Bewertung von Evaluationen*. Bern: Haupt.
- Widmer, Thomas (1996b-h): *Fallstudien zur Meta-Evaluation*. 7 vols. Reihe 'Schlussberichte' des NFP 27. Bern: Schweizerischer Nationalfonds.
- Widmer, Thomas (2000): Evaluating Evaluations: Does the Swiss Practice Live up to The Program Evaluation Standards? *Evaluation Center Occasional Papers Series*, Vol. 17. Kalamazoo, Western Michigan University: 67-80.
- Widmer, Thomas/Beywl, Wolfgang (2000): Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder. In: Joint Committee on Standards for Educational Evaluation (Hrsg.): *Handbuch der Evaluationsstandards*. Opladen: Leske+Budrich: 243-57.
- Widmer, Thomas/Landert, Charles/Bachmann, Nicole (2000): *Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards)*. Bern / Genève: SEVAL.
- Widmer, Thomas/Rothmayr, Christine/Serdült, Uwe (1996): *Kurz und gut? Qualität und Effizienz von Kurzevaluationen*. Zürich: Rüegger.
- Widmer, Thomas/Schenkel, Walter/Hirschi, Christian (2000): Akzeptanz einer nachhaltigen Verkehrspolitik im politischen Prozess. Deutschland, Niederlande und Schweiz im Vergleich. Nationales Forschungsprogramm 'Verkehr und Umwelt' (NFP 41), Bericht D13. Bern: BBL/EDMZ.
- Wilensky, Harold L. (1964): The Professionalization of Everyone. *American Journal of Sociology* 70(2): 137-58.



**Astrid Kuffner**

## **EVALUATION OF SUSTAINABILITY- SUSTAINABLE EVALUATION? LESSONS LEARNED FROM A SURVEY OF EVALUATION REPORTS**

### **Task and Objectives**

#### **Project description**

This analysis marked the first step of a comprehensive research project on the evaluation of sustainability in the Research Focus “Managing Sustainability” at the Vienna University of Economics and Business Administration. It was initiated by the German Evaluation Society (DeGEval) with its working group on environmental evaluation. Its main objective was to perform a survey allowing conclusions on the current state of evaluation in this sector in the German- speaking region.

### **Methodology**

#### **Research**

The author went through a list of selected evaluators of the DeGEval<sup>8</sup>, who were most helpful by providing their own materials as well as new leads through the “snowball system”. The author offered the contacted evaluators a broad definition of evaluation and sustainability, which naturally resulted in a variety of studies provided.

After a period of collection from winter 1999 to spring 2000 the sample of reports amounted to 43. The collection provided the first data base for the evaluation market (customers as well as suppliers) in this special field of application for Germany, Switzerland and Austria. The approach might have resulted in a potential sampling bias, but was the most apt given the green-field nature of sustainability-evaluation.

#### **Analysis**

To analyse the reports the author developed a system for classification which consisted of two elements. First she defined common traits of the reports by reading and re-reading them several times (qualitative element). Then the author developed a grid of appraisal factors based on descriptive statistics, sustainability criteria, a checklist for sustainability projects<sup>9</sup> and selected Program Evaluation Standards<sup>10</sup> (quantitative element). Out of scope for the analysis were reports which did not even treat one column of the triad of sustainability, Structural Funds Evaluations<sup>11</sup> and examinations of methods without an example or use in practice.

It should be noted, that the two elements of the analysis were not meant to exclude some of the reports in the first place, but to iteratively define types of evaluation reports. Both the fields of evaluation and sustainability still lack precise definitions- as do the reports. Hence, the author took an explorative approach to the problem as outlined above.

---

<sup>8</sup> Special thanks to Wolfgang Meyer of the DeGEval (working group for environmental evaluation).

<sup>9</sup> The checklist for sustainability projects of the so called “NachhaltigkeitsTATENBank”, served as a basis (a project of the Austrian Federal Ministry of Environment in cooperation with the Interdisciplinary Institute for Environmental Economics).

<sup>10</sup> The Program Evaluation Standards of the Joint Committee on Standards for Educational Evaluation, compare SANDERS, James R. (Editor): Handbuch der Evaluationsstandards- Die Standards des Joint Committee on Standards for Educational Evaluation, Leske + Budrich, Opladen 1999

<sup>11</sup> because they represent a type of evaluation which is already regulated by the European Union in interval, methodology and form and would have disturbed the explorative character of the analysis.



Given a strict time budget and methodological constraints, the analysis focused on measuring levels of analysis rather than the intensity of the analysis- for the given areas of in-depth appraisal, and the sustainability criteria.

As an example for these “levels of evaluation” might serve the degree of in- depth analysis, which was categorised the following way:

- description: description of measures taken for documentation purposes
- efficiency evaluation: follows the subjective valuation of participants in a project- comparison of target and actual state
- outcome and/or effectiveness evaluation: tries to establish cause-effect chains, measures outcomes quantitatively or qualitatively, against a given reference value
- success factors: tries to open the “black box” of impacts even further by defining prerequisites which helped or inhibited the success of a project

## Results

### Descriptive statistics

43 reports were collected and analysed in this survey. The results show trends in this specialised field, and allow statements on the market situation for environmental and sustainability evaluations.<sup>12</sup>

A general increase of environmental and sustainability evaluation in the German speaking region from 1994 until 1999 can be deduced.

The major part of the evaluations are concomitant (65%), on the regional level (54%) and performed externally (86%). Politics are preferably evaluated in Switzerland, otherwise programmes and projects dominate as evaluanda (both 42%). 30% of evaluations are formative, 30% summative and 40% are both. The type of intervention dominantly evaluated are subsidies (33%).

Almost 50% of the evaluation data were collected and analysed with quantitative methodology (46,5%) (also compare table 1).

**table 1: Descriptive Statistics (n=43), marked with \* multiple choice possible**

n= 43	percentage	evaluanda	percentage
<b>timing</b>			
ex-ante	7%	policies & regulations	16%
concomitant	65%	program	42%
ex-post	28%	project	42%
<b>supplier</b>		<b>level of analysis*</b>	
internal	14%	description	12%
external	86%	valuation (efficiency)	93%
		cause effect (effectiveness)	40%
		success factors	30%
<b>country</b>		<b>methodology used</b>	
Germany	49%	qualitative	21%
Switzerland	14%	quantitative	46%
Austria	37%	both	33%
<b>regional level</b>		additional ecologic*	7,%
national	25%	additional technical*	2%
regional	54%	<b>use</b>	<b>percentage</b>
local	21%	formative	30%
		summative	30%
		both	40%

<sup>12</sup> For full details of the results also view: KUFFNER, Astrid: Evaluation von Nachhaltigkeitsaspekten- nachhaltige Evaluation?- Eine vergleichende Analyse von deutschsprachigen Evaluationsberichten, die sich mit nachhaltiger Entwicklung beschäftigen, Universität Wien, Wien 2000





## Sustainability check

In general a long-term perspective is not a criterion for performing a good evaluation – an evaluation is normally done to give support choosing between alternatives (before or during a project) or looking at it afterwards. The projection of future effects is a well-known methodological problem and in addition normally not in scope of the ordering party. Only one report out of 43 treated possible effects after the intended end of the program.

We can safely say that about one third of the reports treated all aspects of the sustainability triad, although only one gave them approximately equal weight.

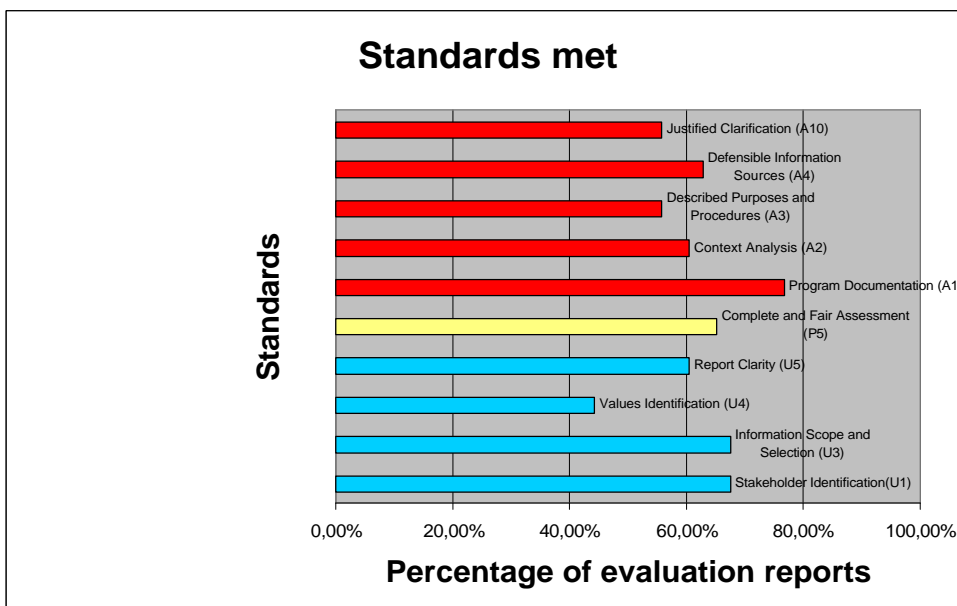
If one defines participation as the questioning of participants in a program/project this criterion is fulfilled by 65% of the reports due to the questionnaires used. If one considers developing the design and/or discussing the results with those, who are stakeholders of the evaluandum, the rate converges towards 0%.

## Quality Standards

The Program Evaluation Standards have originally been developed to evaluate evaluations (metaevaluation). This was not the purpose of the study at hand. Nonetheless the Standards have been the only accredited way to rate the quality of evaluations - at that time<sup>13</sup>. The Standards used for this analysis were selected as to apply to evaluation reports only. That is why the category “Feasibility Standards” has not been applied at all: one can not draw conclusions by reading the evaluation report only.

The author showed how the reports would have performed if accredited Standards were used (for the performance and the standards used also view graph 1).

graph 1: The percentage of reports which met the selected quality standards.



<sup>13</sup> Today, the DeGEval (10/2001) and the SEVAL (Swiss Evaluation Society, 12/2000) have developed standards for evaluation.



## Lessons learned

### Remarks

The following collection of experiences reflects the view of the author. It contains good ones as well as bad ones. The list is certainly far from complete, but may be of help to the Young Researchers who will be working in a similar field after EASY- ECO I.

### Trash and Flash Experiences

- The collection of reports is surely a good way to get in touch with evaluators in this field. You can define hot spots in the market- be it in the field of application (e.g. environmental management and energy sector), the region, or the person of the evaluator. Collecting itself has an multiplier effect: the more persons you contact the more persons contact you. Spread the rumor and make use of it.
- It is common sense that there are different uses of evaluations e.g. affirmation, legitimacy support or severance for programs, establishing a new consulting branch, etc. The question is not, whether evaluations are used, but what for. When assessing an evaluation it is very important to keep the quality of the content, the line of arguments and the analytic insight in mind, not personal judgement. The accredited Program Evaluation Standards were the perfect means for this purpose.
- It will be very easy to define a “who is who” of the evaluation market. The evaluators with reputation, who evaluated big projects with a lot of subsidies behind them will be easy to find. The challenge lies in finding evaluators, who do good work, but are not well-known - yet. Keep your eyes open for institutions who promote small projects, local governments, and University Institutes. The motivation for choosing one evaluator over another are variegated and there are many evaluations in the field of subsidies.
- Although the public sector is probably the richest source of evaluations, it is not the only one. When looking for evaluations, one sometimes has to behave like a truffle swine, going where it apparently leads astray. Follow up every organisation that gives subsidies e.g. government, NGO, foundations, societies, associations, unions, banks, institutes, or clubs. The tighter the link to firms/economy and the more experimental the approach (test runs) the higher the potential. But still, highly interesting evaluations may be done internally and often be confidential.
- The importance of an evaluation is not automatically linked to its quality, but to the importance of the evaluandum. The perspectives of the researcher and the consumer could sometimes not differ more. Keeping in mind that every evaluation is used for a certain purpose and trying to understand the perspective of the ordering party might help.
- Although the objects of evaluation differ, the methods do not. This makes it easier to classify evaluations and compare them. Depending on the field of application, additional methods from the core discipline of the evaluandum are used. After several evaluations with multiple choice questionnaires you may appreciate the contrasts provided by an interdisciplinary approach.
- Precise definitions are not always given- not for evaluation and not for the evaluandum. A mix of methods may be of help in this case. Reading evaluation reports several times helps you to develop “Gestaltwahrnehmung”- a sense for important common traits and properties- commonness itself. Once you own this special sense you will find it easier to label the evaluations and put them in order.



## **EVALUATION**

**Jon Lovett, United Kingdom**

Ecological Sustainability: A Policy Oxymoron?

**William Lafferty, Norway**

Political evaluation: Premises, Approaches and Methods

**Wolfgang Meyer, Germany**

Sociological evaluation – Methods and Limits

**Clive Spash, United Kingdom**

Valuing Society and Environment: Economic Methods and Limits

**Jürgen Freimann, Germany**

Evaluating the Impacts of Corporate Environmental Management Systems. A Comparison between EMAS and ISO 14.001



**Jon C. Lovett**

**Ecological Sustainability: A Policy Oxymoron?**

Centre For Ecology, Law & Policy & Conservation International

**Ecological Sustainability:  
A Policy Oxymoron?**

**Jon C Lovett**  
**Centre for Ecology, Law and Policy**  
**University of York**  
**England**

1

Centre For Ecology, Law & Policy & Conservation International

**“ Oxymoron: a figure of speech in which apparently contradictory terms appear in conjunction ”**

(Oxford English Dictionary)

2

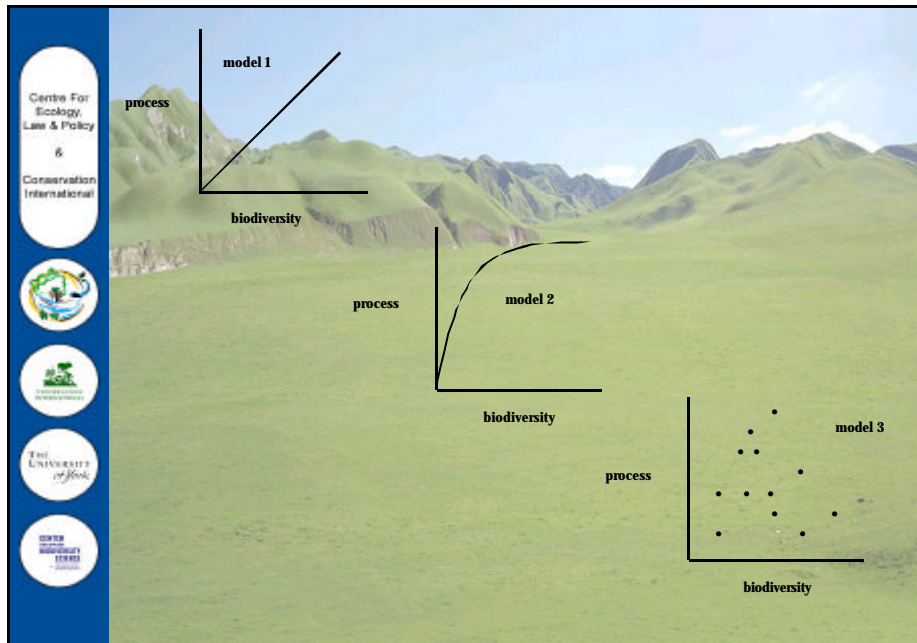


- **Ecology**
- **Sustainability**
- **Evaluation**

3

- 1) Does the productivity of ecosystems depend on their Biodiversity?**
- 2) Does ecosystem stability depend on biodiversity?**
- 3) Does the long term sustainability of ecosystem function depend on biodiversity?**

4



Centre For Ecology, Law & Policy & Conservation International

**Tilman & Downing. Nature (1994) 367: 363 – 365**

**Hypotheses:**

- More diverse communities are more resistant
- More diverse communities are more resilient

**Reasoning:**

- Species differ in their traits

**Conclusions:**

- Species poor plots were more greatly harmed by drought (less resistant)
- Species poor plots took longer to return to pre-drought conditions (less resilient)
- Diversity and ecosystem stability are positively correlated
- Functional diversity should be a saturating function of species richness because, in species rich ecosystems, additional species are more likely to be similar to existing species
- Progressive loss of species should have progressively greater impacts on ecosystem stability

6



Centre For Ecology, Law & Policy & Conservation International

**Tilman & Downing. Nature (1996) 379: 718 – 720**

**Hypotheses:**

- More diverse communities are more productive
- Nutrient losses are lower in more diverse communities

**Reasoning:**

- Interspecific differences allow more diverse communities to more fully utilise limiting resources

**Conclusions:**

- Productivity increases with diversity
- Soil nitrogen utilisation increases with diversity
- Diversity and sustainability are positively correlated
- Loss of species threatens ecosystem functioning and diversity

7

Centre For Ecology, Law & Policy & Conservation International

**1) Productivity of ecosystems does depend on biodiversity**

**2) Ecosystem diversity does depend on biodiversity**

**3) Long term sustainability of ecosystem function does depend on biodiversity**

**Or does it . . . . . ?**

8



Centre For Ecology, Law & Policy & Conservation International

### A question of scale

- 25 hotspots occupying 1.44% of the Earth's land surface contain 43.8% of the world's plants
- Most plant species are rare

9

Centre For Ecology, Law & Policy & Conservation International

### The problem of dynamics

10





Centre For Ecology, Law & Policy & Conservation International

## Sustainability

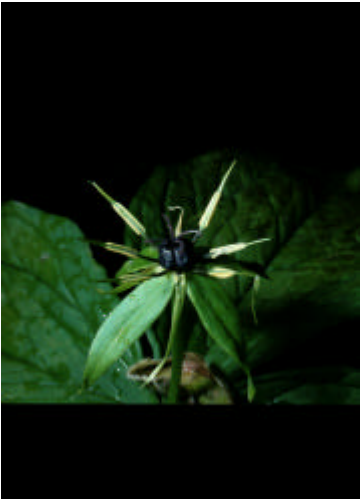


11

Centre For Ecology, Law & Policy & Conservation International

## Evaluation

**Count species diversity?**  
**What about the microbes!**  
**Measure ecosystem functions?**  
**What about rare species!**



12



Centre For Ecology, Law & Policy & Conservation International

## Ecology in policy

- **Species specific legislation**
- **Conserve full range of habitats**
- **Shift subsidies to provision of environmental goods**

13

Centre For Ecology, Law & Policy & Conservation International

## Conclusion

- **Empirical evidence for positive relationship between diversity, productivity, stability and sustainability**
- **Global diversity concentrated in a few hotspots**
- **Empirical relationship only holds at very local scale with small numbers of species**
- **Ecology is dynamic, responding to climate change**
- **Sustainability: a matter of opinion**
- **Evaluation: ecosystem function or public perception?**

14



**William M. Lafferty**

## **POLITICAL EVALUATION: PREMISES, APPROACHES AND METHODS**

### **Introduction**

The use of evaluation as approach and method in political science has very little prominence. A review of seven major “handbooks” of political science published between 1975 and 1999 reveals no significant reference to the approach as an independent method.<sup>14</sup> The only area where evaluation has played a noticeable methodological role is in the sub-field of “policy analysis”. The tradition here is, however, very strong, connected mainly to the study of policy implementation in the United States, Scandinavia and the Netherlands. A key methodological text in this area is Vedung (1997), and there have recently appeared two major collections of the most significant contributions to the field across a broad spectrum of topics and approaches (Rist 1995 and Nagel 2002).<sup>15</sup> The approach is also strongly promulgated by The Policy Studies Organization in the United States (subtitled “The International Association for Decision Makers”), and is often featured in major journals on policy analysis.

The common logic underlying nearly all of these studies is that policy implementation is a central feature of public administration, and that it is the responsibility of policy analysts to improve our understanding of both how implementation works, and the degree to which programmes and policies actually achieve their declared goals. But whereas there is considerable disagreement and competition on the first point (among different “schools” of explanatory policy analysis), discussions of different evaluative approaches and methods are relatively seldom. In general one can say that modern political science has been very self-conscious about distancing itself from its normative roots in classic theory, and this has contributed to a certain “nervousness” around issues of values and objectivity. For many positivists, behaviourists, logical-empiricists and rational-choice analysts in political science, “evaluation” sounds too much like “values” and applied science; directions which the mainstream of “academic” political science would prefer to keep at arm’s length. It is no accident, I believe, that the vast majority of evaluative studies in the area of environment and development are carried out by disciplines *other than* political science.

Be that as it may, as the references cited in the present paper indicate, there is more than enough of relevant “discourse” within the sub-field of policy analysis to nourish the development of a systematic and well-disciplined “school” of political evaluation. The key to such a development is not to avoid either values or normative political theory, but to combine the two with established standards of empirical enquiry. This is particularly important in the area of sustainable development, where conflicts on values and principles are rife, but is also a vital aspect of democratic performance in general. The present author has, for the past several decades, been working on research issues in this area, and can illustrate the approach by first looking briefly at what has come to be known as the “normative-empirical approach” to democratic performance, and then turning more specifically to the evaluation of strategies for sustainable development.

### **Normative-empirical analysis of democratic performance**

It is central concern of democratic governance that democracies should be *both* legitimate and effective. Most arguments for democracy are addressed to the first characteristic, claiming that democracies are *per se* the most legitimate form of authority. Many (hardly all) are also willing to argue that, in the long run at least, democracies are also the most efficient form of governance. These arguments usually build on one or another version of the “mobilization-of-competence” argument, namely that complex societies require highly integrated systems of personal and institutional interaction, and that it is only by including as many actors as possible in the decision-making process that one can maximize the effectiveness of the overall performance.

---

<sup>14</sup> The works consulted were *The Handbook of Political Science* (Greenstein and Polsby, 1975 and subsequent years); *The International Handbook of Political Science* (Andrews, 1982); *The Dictionary of Political Analysis* (Plano, Riggs and Robin, 1982); *Bridges to Knowledge in Political Science: A Handbook for Research* (Kalvelage, Meland and Segal, 1984); *Political Science Research: A Handbook* (1996); *A New Handbook of Political Science* (Goodin and Klingemann, 1996); and *The Blackwell Dictionary of Political Science* (Baley, 1999).

<sup>15</sup> Additional recent, and highly relevant, contributions are: Mayne and Zapico-Goñi (1997), Riper and Toulemonde (1997), and Bemelmans-Videc et al. (1998).



Combinations of both arguments are currently prevalent in the debate over the legitimacy and effectiveness of the European Union.

Evaluation is an approach that rather uniquely cuts across both types of argument. On the one hand, evaluation is necessary for supporting democratic legitimacy since it is vital that the “people” see that programmes related to electoral majorities are being carried through. Democracy must be *perceived* to be implementing the goals, programmes and policies that politicians have promised. Such perception can, however, be either highly symbolic or highly instrumental. Democracy *may look like* it is working to follow up political commitments – but the *actual degree of instrumental output and change* may be another story all together. Evaluation is thus also necessary to verify the instrumental effectiveness of the implementation process.

Further, both tasks – guaranteeing legitimacy and documenting instrumental effectiveness – require *external objective* evaluation. “Form follows function”, and it goes without saying that *political neutrality* is necessary to monitor the follow-up of political commitments; and *methodological objectivity and discipline* are necessary to document effectiveness. Some journalists *may* be neutral enough to perform the first task (most are not), but even the best and most neutral journalist will lack the professional ethos and methodological schooling of a scientifically trained evaluator.

While the second of the two tasks indicated (monitoring effectiveness) is a common declared goal among professional evaluators, the first task (guaranteeing legitimacy) is not a goal – or at least not a *declared* goal. Indeed the essence of conflict associated with most programmatic or institutional evaluations is the *degree of association or compliance with the contracting party*. Anyone who has either conducted or been subjected to a “conditional evaluation” (an evaluation where the result will directly effect the continuation of the activity under evaluation), knows that the degree of “sympathy” between the contractor and contractee is a key issue affecting the results of the evaluation. The notion that an agent would contract an evaluation which seriously undermines the legitimacy of the agent itself, is simply not part of the normal evaluation “game”. It is a part, however – and *must be* a part – of the evaluation of democratic societies and their political programmes. Hence the very basic initial premise that, if evaluations of democratic performance in general, and of national strategies for sustainable development in particular, are to be effectively carried out, the evaluating unit must have a mandate guaranteeing political neutrality. Whether or not the unit manages to live up to the mandate and manifest it in practice, is, of course, another question.

Given that such a mandate has been both issued and maintained, how can it be applied in practice to democratic performance? As an approach within the Department of Political Science at the University of Oslo, “academic freedom” enabled the development of a programme of “normative-empirical analysis”. During the period 1981-92 (roughly), numerous projects related to different aspects of democratic performance were carried out, all operating on a dual normative-empirical track. The basic steps of the approach were quite simple – though up to that time there were virtually no “paradigms” from the discipline to build on. Over the years the approach led to numerous publications in both English and Norwegian, including several graduate theses and doctoral degrees. The elemental steps of the approach are listed in Box 1



**Box 1: Basic steps of normative-empirical research**

1. Identification of a specific practical discourse where questions of democratic norms were at issue: these could be questions of equality, freedom, participation, rights, or whatever, and the discourse could be related to any aspect of political performance: national or local democracy, institutions, workplace democracy, women's rights, etc.
2. Connection of the specific issue in question to one or more academic discourses related to the problematic.
3. "Translation" of the practical-discourse problem into a normative-theoretical discourse problem: clarifying the implications of the problem within a normative-theoretical context
4. Formulation of empirical criteria, drawn from the field of the practical discourse, by which the normative problem could be addressed and clarified
5. Determination of the relevant empirical methodology necessary to an objective analysis of the normative problematic
6. Execution of the empirical analysis, with conclusions for *both* the practical discourse and the normative-theoretical discourse.

**Several different types of study were carried out during this period:<sup>16</sup>**

Parliament as a discursive forum: An analysis of the practice of the Norwegian parliament with respect to norms of democratic discourse. Focusing on the debate over the 4-year long-term plan in Norway – where expectations as to meaningful dialogue and constructive discourse should be stronger than for normal parliamentary inter-party exchanges – the study first clarified the issue vis à vis discourse theory, and then applied the norms to an empirical analysis of the debates. The results revealed a clear bias inherent in the rules for debate to the advantage of the largest parties, indicating that normal procedures for allocating parliamentary seats according to electoral pluralities can detrimentally influence the quality of discourse on long-term, overarching issues.

Norway as a "Distant Democracy": Taking its point of departure in a major empirical study of political participation in Norway (Martinussen 1973), the goal of this analysis was to demonstrate that the negative conclusions of the study (portraying Norway as a "Distant Democracy" characterized by low levels of political involvement and alienation), were not warranted on either empirical or normative grounds. By initially explicating and reworking the vague normative premises of the study, and then applying many of the same indicators and analytical methods to similar data sets, the study clearly showed that the situation for participation and democracy was much less negative than originally indicated (Lafferty 1981). The debate raised by the study led to numerous other studies and debates focused on the normative criteria for participation in a democracy, with significant consequences for both normative theory and everyday expectations as to what kind of participation should/can lead to what kind of results.<sup>17</sup>

Democracy and the media: The general theme here was a major trend during the 1980's in Norway to "democratize the media". Taking its point of departure in high-profile political proposals and debates as to the "democratic nature" of the mass media – primarily radio and the press at this juncture – the study did a thorough analysis of the public documents in relation to relative norms for democratic media. Criteria were developed for different aspects of the democratization problem, and numerous analyses of different types of media were then carried out. Once again a major finding was that norms had been vaguely expressed and often incorrectly related to different proposals for change.

<sup>16</sup> All of the studies carried out are in Norwegian. An overview of the relevant theses and dissertations is available at the website of the Department of Political Science: <http://www.statsvitenskap.uio.no/fag/hoved/sammen/>. Some of the more relevant works are Hovde (1982), Langeland (1985), Christensen (1987), Mageli (1987), Tørres (1987), Auråker (1988), Raaum (1988), Skogerbø (1988), Hov (1989), Strandhagen (1990), and Kjelland-Mødre (1991).

<sup>17</sup> The issues have been debated primarily in Norwegian. Interested readers of a Scandinavian language can consult Berg 1983 and Hagtvet 1980. There is also an interesting graduate thesis devoted to the debate between Martinussen and Lafferty on democracy in Norway (Strømsnes 1993).



The empirical analyses also showed (again) that different, and often competing, democratic norms could lead to contradictory performance and conflicts among valued norms (Skogerbø 1988 and 1996).

Democracy and learning at the workplace: Finally, we can mention a whole series of studies where the theme was workplace democratization, and the problem was to untangle a whole spectrum of broadly and vaguely expressed values and expectations. A major challenge for this subset of studies was to differentiate between the specification of democratic norms for a political community *per se*, and the specification of democratic norms for both workplaces in general, and specific types of workplace in particular. It was here that the notion of “form follows function” emerged as a key guideline, indicating that specific types of functional activities – activities which by functional definition require a division of labour and responsibility – require specific modes of democratic procedures and values. In short, the “democratization” of an activity is an adaptive process whereby the essential nature of the activity in question must be given due and balanced consideration in relation to democratic norms.<sup>18</sup>

A major result of all these studies was to develop a specific awareness and identity as a separate research approach to practical problems where democratic norms were directly at issue. We were able to address political issues which were often very contentious; to place those issues within relevant systematic discourses of normative political theory; and to assess existing states of affairs against the normative criteria derived. All this was done, however, within a normal university setting, where the task was principally conducting basic research and schooling political scientists. In the next phase of the work, the profile of the approach was specifically recognized in a programmatic context, and connected directly to the evaluation of national democratic performance on achieving sustainable development.

### **Strategic research for sustainable development: Institutional prerequisites**

In 1992 the current author was asked to head an existing programme devoted to the development of alternative macro scenarios for the socioeconomic development of the Norwegian society. The “Project for an Alternative Future” (PAF) had been initiated in 1982 by a small group of NGO-leaders and concerned academics, with the intention of developing alternative paths of development which could have political influence. The project was given direct support by the Norwegian parliament, and for a number of years functioned as a more-or-less ad-hoc think tank for the Committee on Education and Research within the parliament. The project was evaluated by an independent academic committee in 1988, and it was decided to transfer responsibility for the activity to the then Norwegian Research Council for Science and the Humanities (NAVF). A new five-year period was inaugurated in 1990, but the constellation of diverse actors and interests had difficulty finding an appropriate form of applied research. When new leadership was introduced in 1992, it was gradually decided that the major thrust of the activity should be altered from one of developing alternative holistic scenarios, to one of assessing Norway’s progress on issues of sustainable development. The change was inspired largely by the broad promulgation of the findings and recommendations of the World Commission on Environment and Development in its concluding report *Our Common Future* (WCED 1987). The change of leadership took place, moreover, at the exact time of the follow-up conference to the WCED, the United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro in June of 1992.

In 1993 the research community in Norway underwent a major transformation when five separate councils were merged in one Norwegian Research Council (NRC). In addition – and most significantly for the present “story” – a totally new division was created within the Council, the Division on Environment and Development (DED). A major task of the new division was to integrate research on environment and development, both within the division itself but also, as far as possible, within the Research Council as a whole. It was also indicative of the times that one of the principle new objectives adopted for the NRC was to promote research that contributed to “the creation of value-added capital within a framework of sustainable development” (“*verdiskapning innenfor en bærekraftig utvikling*”).<sup>19</sup>

---

<sup>18</sup> See the list of graduate theses in note 2. More general treatments are presented in Lafferty 1984b and 1989, and Lafferty and Raaum 1992.

<sup>19</sup> The interesting conjunction of “value-added capital” and “sustainable development” was also typical of the times, and particularly in line with similar signals from the European Union, particularly in the Maastricht Treaty.



It was within this context that the Project for an Alternative Future was terminated in 1995, and replaced by a new “Programme for Research and Documentation for a Sustainable Society “ (ProSus). The leadership and most of the staff were the same, but the remit from the Research Council was significantly altered. ProSus was defined as an “applied strategic research program” designed “to produce new knowledge and conduct informational activities in support of a better realization of national goals for sustainable development” (ProSus 2000: 1). The activity of the programme was to be conducted along three lines (Box 2).

**Box 2: Principle activities of ProSus as defined by the Research Council of Norway**

- Documentation and evaluation of Norway’s follow-up of the Rio accords and the guidelines of the United Nations Commission for Sustainable Development. The programme is to focus on the political, social and economic goals of the UNCED process, and should submit regular reports on Norway’s progress with respect to the declared goals and values.
- Goal-oriented strategic research on the barriers which stand in the way for a more rational and effective realisation of sustainable development. The activity here should be conducted in cooperation with other research institutions, both nationally and internationally, and in close dialogue with voluntary organisations and representatives for business and labour.
- Information and dissemination of alternative strategies of governance, steering instruments and normative future perspectives for more sustainable societies, locally, nationally and globally. The activity here to be coordinated through networks with other research and information efforts in this area.

Of most direct relevance for the issue of political evaluation, is the fact that ProSus receives the major part of its funding from the Ministry of Education and Research, but the allocation is channelled through the Research Council of Norway, guaranteeing that there is no political connection between the Ministry and the Programme. Furthermore, ProSus was, as of 2000, designated a “strategic university programme”, placing its evaluation and research activities within the administrative domain of the Centre for Development and the Environment (SUM) at the University of Oslo. SUM is an “all-university” research and teaching unit, established directly under the University Senate. Its remit also prescribes research and information on sustainable development, but is both more basic-research oriented and more focused on development and environment issues in developing countries. The two unites are, therefore, highly complementary, with ProSus focusing more on the pursuit of sustainable development within Norway, and SUM concentrating on sustainable development in an international, North-South context.

These administrative details are necessary to understand the nature of the evaluative strategic research being conducted at ProSus. By monitoring and evaluation of Norway’s initiatives to achieve sustainable development, ProSus fulfils an important task for democratic legitimation, at the same time that an empirical foundation is laid for focused research on the barriers to, and new possibilities for, more effective implementation. Such a research operation is a vital adjunct of the strategic intentions of both UNCED and, more recently, the European Union. The political bodies of the international and European communities have made sustainable development an overarching goal for economic and developmental activities in general, and it is absolutely essential that scientific bodies be accorded a role in evaluating and improving implementation efforts.

**“Grounding” evaluation: What is “sustainable development”?**

The first issue that must be resolved in any evaluation is a common understanding of the nature of what is to be evaluated. A high degree of consensus on this issue is a necessary, but not sufficient, condition for a “successful” evaluation. This is true for professional programme and policy evaluation; and it is *particularly* true for independent political evaluation. Clarification of the goals and mandate for change under evaluation is a crucial condition for converting analytic results into better practice.



In developing evaluative criteria for ProSus, we have relied on a relatively simple three-point programme for “grounding” analyses of initiatives for sustainable development.

**Firstly**, we consciously opt out of the business of trying to tell anyone what sustainable development “really is”. ProSus does not have a substantive programmatic position on sustainable development. Our strategy is to declare that sustainable development is what the governments of the world have committed their countries to within the UNCED process. All of the accords adopted at Rio, particularly the *Rio Principles* and *Agenda 21*, have sustainable development as their goal, and virtually all members of the United Nations have committed themselves to pursuing the goal. Our solution to the definitional problem is, therefore, conventionally political. We believe that this solution provides by far the closest approximation to a “consensus” on what sustainable development is. It is an *instrumental* and *political* solution, which, in our opinion, means that it provides a much better point of departure for change than any of the myriad of alternative definitions, positions and programmes that one can find in the voluminous literature on sustainable development.<sup>20</sup>

**Secondly**, we point out that the Rio documents make no extensive effort to define sustainable development. This is because the Rio Summit was designed and constructed to specifically follow up, specify and facilitate the understanding of sustainable development as put forth in the WCED report, *Our Common Future*. Insofar as UNCED commits the nations of the world to work for sustainable development, it is the sustainable development of *Our Common Future*. Does this not mean then that we are simply “passing the buck” on the definitional problem to *Our Common Future*, the meaning of which is notoriously contended? We think not – for two reasons: (1) we believe that a textual analysis of *Our Common Future* reveals that the type of development envisaged as “sustainable” is distinct enough on its own terms to contrast sharply with major characteristics of currently competing paths of *non-sustainable* development<sup>21</sup>; (2) we believe that the Rio accords (particularly *Agenda 21*), and the follow-up work of UNCSD, provide more than enough specifics for change (objectives, guidelines, actions, indicators) to move the agenda on sustainable development a *considerable* distance towards change.

**Thirdly**, we have developed three distinct types of “evaluative criteria” to provide meaningful, and politically relevant, nuances to our evaluations: (1) *External criteria*: the documents and guidelines provided by the WCED, UNCED, UNCSD – and more recently UNEP, the OECD, and the European Environmental Agency. These sources provide the most general level of standards for assessing Norway’s follow-up to the UNCED-accords; (2) *Internal criteria*: the strategies, goals, programmes, action plans, policies etc., adopted by government as implementation of the UNCED programme; (3) *Comparative criteria*: the activities, initiatives, programmes, policies, etc., undertaken by other governments in their pursuit of the Rio goals.

This position on sustainable development is explicitly “operative”. It is designed to build as much objectivity into the evaluation process as possible; an objectivity anchored in the commitments of democratic governments, and applicable within a very intuitive understanding of the relationship between research and politics. Given that governments mean what they say when they sign international or regional agreements to pursue sustainable development, social research (in this case) has both the right and the obligation to monitor the implementation process. The position is, therefore, not designed to either defend or “preach” a specific understanding of sustainable development, but to accept, clarify and evaluate the semantic and operational implications of the political discourse in question.

---

<sup>20</sup> See Lafferty 1996 and 1999 for the complete argument here.

<sup>21</sup> The argument for this position is presented in a collection of studies devoted solely to *Our Common Future* (Lafferty and Langhelle 1999).





### **Operational guidelines: The need for “rules of thumb”**

While the “grounded definition” and standards outlined in the previous section serve as basic orientations for political evaluation, it is also the case that one encounters numerous situations in the evaluation process where the criteria of either *Agenda 21*, the UNCSO guidelines, or the proscriptions of the EU on sustainable development, are open to interpretation, weighting, ranking, etc. This is a most normal situation when pursuing highly normative goals; goals moreover which clearly must be realized in a multitude of different types of society, with markedly different resource bases, economies, social structures, levels of development, etc.

Even the most specific sections of *Agenda 21*, or the most refined of UNCSO indicators for sustainable development, will have to be *relativized* to the individual community. Form must be adapted to function, but form must also operate within contextual constraints. Mediating between the definition and standards outlined above and the specifics of any given society requires “rules of thumb” for effective communication and consistent research results. Demands for “cumulative knowledge” are no less relevant for evaluative research than for more “basic” modes of nomothetic science.

“Rules of thumb” are simple prescriptions for the basic application of any craft or discipline. They are what carpenters, plumbers and electricians learn as apprentices, and what social scientists learn (ideally) as graduate students. In applying evaluative standards for sustainable development, we can identify three types of “rules of thumb” which aid in the standardization of the research effort: *differentiation criteria*, *operational paradigms*, and *performance benchmarks*.

### **Differentiation criteria**

One of the key operational tasks in the evaluation of sustainable development is to know it when you see it. There are a great number of phenomena going on in any society, at any level of development, which may, or may not, be judged as “sustainable development”. Shakespeare made the basic point by pointing out that “Beauty is in the eye of the beholder”; but beauty is not a strategic programme for change, and Shakespeare was not a social scientist. We need help.

Two logical “devices” can provide assistance. Though neither is so intuitively clear as to guarantee consensus, they are at least consistent with the grounded position here outlined. The first has to do with differentiating sustainable development phenomena from other similar phenomena, and the second has to do with differentiating among the so-called “environmental, economic and social pillars” of sustainable development. Both aspects are illustrated by the simple list provided in Box 3. The list provides an overview of “basic components” of sustainable development, with the term “gradients” also used in parentheses to highlight an underlying logic of stages. The list thus reflects a *selection of identifiable characteristics*, formulated as a *sequence of prioritized issues and standards*.

As a “rule of thumb” for evaluation, the list says: “These are the essential elements of sustainable development (within the UNCED context), and they are to be given prominence in the order listed”. The key implications of this in an operational context are as follows:



### **Box 3: Basic components (“gradients”) of sustainable development**

#### The environmental component:

*Consisting of three major aspects (phases) of normative environmental politics:*

- nature conservation
- environmental protection
- ecological balance

#### The economic component:

*Consisting of the key elements of the Brundtland/UNCED goal of a “qualitatively” different mode of (“sustainable”) economic development (“de-coupling”):*

- eco-efficiency (reducing the impact on natural resources for producing goods and services)
- sustainable consumption (consumer responsibility for eco-efficiency)

#### The social (equity) component:

*Consisting of four dimensions of equity for adjusting the satisfaction of “basic needs” to the sustainable functioning of natural life-support systems:*

- national equity for current generations
- national equity between current and future generations
- global equity for current generations
- global equity between current and future generations

(1) Sustainable development incorporates, but is different from all three of the “environmental components”. When “looking for” sustainable development to evaluate, it is important to differentiate it from “pre-sustainable-development” environmental concerns. The conservation of nature, protection of the environment from pollution, and maintenance of ecological balance are all vital aspects of sustainable development, but they were in place and being pursued prior to the political commitment to sustainable development, and their achievement is not sufficient to qualify for high marks on SD implementation.<sup>22</sup>

(2) Having “said” this, it is nonetheless the case that the most fundamental dimension of sustainable development is its environmental or “ecological” component. This is what anchors the concept logically, making it different from other normative concepts, ideologies and programmes devoted to socio-economic welfare, justice, political reform, etc. In short, the ecological component is a necessary, and “lexicologically” prioritized, dimension of sustainable development – but it is not sufficient to qualify as an independent standard for implementation.

(3) The “economic” component is the second-most important aspect of the concept. It reflects the vital, and highly differentiating, message of *Our Common Future*: namely that any level of economic growth must be fundamentally assessed as to its impact on the sustainable functioning of life-support systems. This involves, if one will, the “materialist bias” of the WCED position. It is the basic relationship between natural life-support systems and the dispositions of the “means of production” which warrants primary attention in a sustainable development context. Moreover, this relationship must be conceptually and analytically “de-coupled”. Where it can be demonstrated that modes of production driven by the under-satisfaction of basic needs are causing harm to sustainable life-support systems, these must be rectified. Likewise, where it can be demonstrated that modes of production driven by excessive satisfaction of basic needs, these too must be rectified. In general, conditions of poverty characterize the first type of imbalance, and conditions of over-consumption characterize the second type.

<sup>22</sup> The classic statement on this aspect from *Our Common Future* is as follows: “The environment does not exist as a sphere separate from human actions, ambitions and needs, and attempts to defend it in isolation from human concerns have given the very word ‘environment’ a connotation of naivety in some political circles. The word ‘development’ has also been narrowed by some into a very limited focus, along the lines of ‘what poor nations should do to become richer’, ... But the ‘environment’ is where we all live; and ‘development’ is what we all do in attempting to improve our lot within that abode. The two are inseparable.” (WCED 1987: XI)



(4) The “social dimension” of sustainable development (once again, in the UNCED context), is a generalized reference to the ethical stipulation of *Our Common Future*, to the effect that issues of decoupling should be assessed and adjusted according to four dimensions of equity applied to the interdependency between modes of production, the level of satisfaction of basic needs, and impacts on natural life-support systems. These are: (a) national equity for current generations, (b) national equity between current and future generations, (c) global equity for current generations, and (d) global equity between current and future generations. The “social dimension” is viewed in this light as a normative perspective for the assessment and regulation of the relationship between economic activity and natural life-support systems according to fair standards for the satisfaction of basic needs.

(5) Though it is clearly difficult to reach a consensus on benchmarks for both “sustainable life-support systems” and “basic needs”, the conceptual and operational difficulties encountered here are considerably over-dimensioned by both public and academic discourse. Reasonable indicators for critical thresholds of life-support systems – air, water, nutrients, sinks, and non-renewable vital resources – are available, as are reasonable conceptual standards and indicators for basic needs. In an evaluation context, the emphasis must be placed on “reasonable” within the context in question, and not on “reasonable beyond doubt”. The community of evaluation practitioners should, however, be willing to devote considerable more attention and resources to the question of consensual standards.

Once again, it is important to emphasize that these criteria are not meant to impose *general* standards on the understanding and evaluation of sustainable development (probably an impossible task), but to assist in the evaluation of sustainable development as conceptualized and adopted as a political programme within the UNCED process.

### Operational paradigms

Another device which has proved particularly useful in the evaluative work done at ProSus, is the notion of an “operational paradigm”. This is particularly important when the task is one of comparative assessment across different political constituencies, where there is an acute need for common points of reference. By way of illustration we can look at two types of criteria which have been employed with considerable success in the SUSCOM project, an EU-funded “concerted action” to determine the status and implications of the implementation of “Local Agenda 21” (LA21: Chapter 28 of *Agenda 21*) in Europe.<sup>23</sup> Before the project could proceed to an empirical monitoring of progress on LA21, we had to establish common criteria across the 12 research partners as to what would be considered “a Local Agenda 21”. This was accomplished in two steps.

The initial step was on a normative-descriptive level: Once again the question had to be answered – “How would we know ‘an LA21’ if we saw one”? The problem was particularly acute for an assessment of Chapter 28 of the *Agenda*, since the guidelines for objectives and activities were (of necessity) extremely open. Given the enormous diversity of communities throughout the world that would have to work with the Chapter, it was clear that any attempt to provide substance benchmarks for achievement would be counter-productive. We developed, therefore, a relatively simple logic: (A) Chapter 28 was an explicit appeal to “local authorities” to engage in a “dialogue” with their citizens and stakeholders, so as (implicitly) (B) to promote sustainable development. After considerable internal “dialogue” on our own, the SUSCOM team arrived at the set of criteria shown in Box 4.

---

<sup>23</sup> The SUSCOM project was conducted between 1997 and 1999, covering LA21 activities in 12 European countries. The project is comprehensively reported in Lafferty and Eckerberg, 1998, Lafferty 1999 and Lafferty 2001.



**Box 4: Basic criteria for identifying Local Agenda 21**

- A more conscious attempt to relate environmental *effects* to underlying economic and political *pressures* (which in turn derive from political decisions, non-decisions and markets)
- A more active effort to relate local issues, decisions and dispositions to *global impacts*, both environmentally and with respect to global solidarity and justice.
- A more focused policy for achieving *cross-sectoral integration* of environment-and-development concerns, values and goals in planning, decision-making and policy implementation.
- Greater efforts to increase *community involvement*, i.e. to bring both average citizens and major stakeholder groups, particularly business and labour unions, into the planning and implementation process with respect to environment-and-development issues.
- A commitment to define and work with local problems within: (a) a broader ecological and regional framework; and (b) a greatly expanded time frame (i.e. over three or more generations)
- A specific identification with Agenda 21 and the UNCED process

SUSCOM Project (Lafferty 2001: 4-5)

Once these criteria were agreed – and brought to bear in our joint assessments of the original monitoring efforts – the question arose as to whether it was possible to further clarify the reporting effort by outlining a more dynamic and systematic “model” – an “ideal-type” of “best case” which would reflect the most ambitious aspirations of the Rio action plan. While the criteria were highly useful in established a set of individual characteristics, we soon found that the criteria could be fulfilled individually and piecemeally, without other criteria being manifest at all.

On this second task, it was highly fortunate that one of the most significant documents related to the entire LA21 effort – the Ålborg Charter – had in fact outlined just such a process. Though hardly promulgated at all across Europe (it was included in one of the Charter’s subsections), the list of integrated criteria here provided exactly what was needed for “evaluative imaging”. As shown here in Box 5, the paradigm from the Ålborg Charter outlines in much greater detail the logical steps which both could and *should* accompany a commitment to LA21. Since there were numerous signatories to the Charter in all 12 of the SUSCOM countries, this meant that we also here could bolster the validity and legitimacy of our assessments in and through the open democratic commitments to the LA21 goals. The devices in question thus provided key operational criteria on the basis of widely sanctioned and internally consistent programme standards.

**Box 5: The Ålborg Charter “paradigm” for Local Agenda 21**

- Information-gathering and consciousness-raising
- Interpretation and relativization of Agenda 21 to local conditions and problems
- Development of priorities and local action plans with both general and sector-specific targets
- Determination of appropriate steering instruments, including the procurement of voluntary agreements among sector-relevant social actors (“stakeholders”, “major groups”, “target groups”)
- Goal-specific procedures for implementation
- Monitoring and evaluation (of both the enactment process and its effects)
- Revision of goals, plans and initiatives

The Ålborg Charter, Part III



### Performance benchmarks

Finally – and very briefly – there is a clear need in all forms of political evaluation for developing some kind of performance benchmarks. Differentiation criteria are necessary to make the first “cut” among competing relevant phenomena; and operational paradigms can significantly improve comparative monitoring; but neither device can help us to determine whether the glass is “half full” or “half empty”. Even accepting that we are operating on a very general and complex level of evaluation (systemic political performance), we still need some form of basic benchmark for strengthening the practical implications of the evaluative results.

Turning once again to the SUSCOM project, we can conclude the overview of devices by considering the categories presented in Table 1. In contrast to the first two devices, which were largely developed prior to the actual monitoring and reporting of the implementation efforts, this final device summarizes some of the more significant findings of the comparative evaluation. The table gives expression, on the one hand, to an acknowledgement by the project of a need for differentiating between different levels or types of LA21 performance (the “Four P’s”): which in turn led to a broad categorization of the results in terms of 5 different “modes” of LA21 implementation. Taken together the results constitute a new set of “normative-empirical” benchmarks for LA21 achievement – with clear indications as to what will be the strengths and weaknesses of each mode. By employing the term “mode”, instead of “model”, we clearly indicate how and why the four dominant extant “modes” deviate from the Ålborg paradigm. The benchmarks in question thus reflect specific types of “outcome” from LA21 initiatives, expressed as deviations from the ideal type. The evaluative approach to LA21 is thus provided with a new knowledge base and point of departure for further empirical assessments, as well as new insights for revisions the action plan for sustainable development in this area.

**Table 1: Stylized modes of LA21 implementation in Europe**

Mode of implementation	<u>The Four P’s</u>			
	Emphasis is place on:			
	<u>Process:</u>	<u>Plan:</u>	<u>Policy:</u>	<u>Product:</u>
	New methods of mobilisation and cooperative governance for achieving SD	Adopting a strategic plan for SD with targets and indicators	Adoption of single-issue SD policies and programs	Achievement of confirmed SD-goals and targets
Ålborg	Yes/strong	Yes/strong	Yes/strong	Probable/strong
Paternal	No	Possible	Possible	Possible/Weak
Integrated/Single-issue	Possible/partial	Partial/strong	Partial/strong	Probable/narrow
External/Forum	Partial/strong	Possible/weak	No	Unlikely
External/Fragmented	Possible/partial	No	No	Possible/narrow

Source: Lafferty 2001: 292



### Concluding perspectives: Strengths and weaknesses?

The major strength of the type of political evaluation here described is that it is not political. Indeed, insofar as it *is* political, its quality and impact are weakened. The evaluation of political processes requires, in short, a studiously *apolitical* institutionalization and execution if it is to have broad-ranged and long-lasting impact. This is particularly true with respect to at least four crucial “mediators” of social change: the mass media, science, public administration and business. No meaningful progress towards sustainable development will be achieved without the positive engagement of each – and all four are particularly wary of politicized research. Insofar as evaluation of political implementation is perceived as politically biased in any recognizable party-political sense, one can guarantee that the effects of the results will be proportionate to the influence of perceived party beneficiaries. Not only that, but one will also contribute to a direct politicization of the programme, policy or issue in question, thereby enhancing the conflictual nature of the issue, resulting in yet higher thresholds for change. The negative prospects of such a development are particularly acute with sustainable development, since the issue has thus far enjoyed a very high moral status. Such an admonition does *not* mean, however, that evaluation of initiatives for sustainable development must be non-controversial. To the contrary, the more controversial the better, since change is nourished by controversy – but it must be controversy related to deviations from clearly documented political commitments, and clearly delineated efforts to achieve change.

To the degree that this sounds like an overly naive and relatively hopeless prescription for evaluative research in this area, we stand face to face with the potential limits of evaluation as a medium for change. The slightest concession towards the institutionalization of “deep structure” politics in and through evaluation practice, however “balanced” in relation to other similarly “disguised” practices, the more quickly professional evaluation will transcend, and ultimately sever, its contacts with academic policy analysis. The ability to maintain a proper disciplinary balance – and even more crucially, the ability to finance such a balance over the long run – is the outstanding challenges of the persuasion. They are, however, only marginally different from similar challenges in the natural, technical and economic sciences. It’s all part of developing a truly effective science for change: a science particularly suited to the democratic pursuit of sustainable development.

### References

- Berg, T. (1983). *Deltagerdemokratiet* (“Participatory Democracy”). Oslo: The University Press.
- Bemelmans-Videc, M.-L.; R. C. Rist and E. Vedung (eds) (1998). *Carrots, Sticks and Sermons: Policy Instruments and Their Evaluation*. New Brunswick: Transaction Publishers.
- Hagtvet, B. (1980). *Den Vanskelige Lydigheten* (“The Difficult Obedience”). Oslo: Pax Forlag.
- Hagtvet, B. and W. M. Lafferty (eds) (1984). *Demokrati og Demokratisering* (“Democracy and Democratization”). Oslo: Aschehoug Forlag.
- Lafferty, W. M. (1981). *Participation and Democracy in Norway: The “Distant Democracy” Revisited*. Oslo: University Press.
- Lafferty, W. M. (1983a). “Deltagelse og demokrati: Momenter i en uendelig dialog” (“Participation and Democracy: Moments in an Unending Dialogue”), in Trond Bergh (ed.) *Deltagerdemokratiet* Oslo: University Press.
- Lafferty, W. M. (1983b). “Political Participation in the Social-Democratic State: A Normative-Empirical Framework for the Analysis of Decision-Making Involvement in Norway”. *Scandinavian Political Studies* 6: 281-308.
- Lafferty, W. M. (1984a). “Decision-Making Involvement in Norway: The Nature and Scope of Citizen Access in a Social Democratic Polity”. *European Journal of Political Research* 12: 43-58.
- Lafferty, W. M. (1984b). “Workplace Democratization in Norway: Current Status and Future Prospects with Special Emphasis on the Role of the Public Sector”. *Acta Sociologica* 27: 123-138.
- Lafferty, W. M. (1985a). “Decision-Making Involvement and Socioeconomic Resources in Norway: A Normative-Empirical Analysis”. *Scandinavian Political Studies* 8: 253-282.
- Lafferty, W. M. (1985b). “Stortinget som demokratisk forum: Interessesetninger eller dialog?” (“The Storting as a Democratic Forum: Conflict of Interests or Dialogue?”). *Norsk Statsvitenskapelig Tidsskrift* 1: 25-34.



- Lafferty, W. M. (1987). *Democracy and Citizenship: A Conceptual Framework for Comparative Normative-Empirical Analysis*. Report from the project on "Varieties of Democratic Experience", Department of Political Science, University of Oslo.
- Lafferty, W. M. (1989). "Work as a Source of Political Learning Among Wage-Laborers and Lower-Level Employees", In Roberta Sigel (ed.) *Political Learning in Adulthood: A Sourcebook of Theory and Research*. Chicago: University of Chicago Press.
- Lafferty, W. M. (1996). "The Politics of Sustainable Development: Problems and Prospects after Rio". *Environmental Politics* 5: 185-208.
- Lafferty, W. M. (ed) (1999a). *Implementing LA21 in Europe: New Initiatives for Sustainable Communities*. Oslo: ProSus and the European Commission.
- Lafferty, W. M. (1999b). "Sustainable Development as a Research Problem and Theme for Political Science". *International Political Science Review*, 20: 123-128.
- Lafferty, W. M. (ed) (2001a). *Sustainable Communities in Europe*. London: Earthscan.
- Lafferty, W. M. (2001b). "Local Agenda 21: The Pursuit of Sustainable Development in Sub-National Domains", In *Sustainability Assessment at the Local Level*, Dimitri Devuyst (ed), New York: Columbia University Press.
- Lafferty, W. M. and K. Eckerberg (eds) (1998). *From the Earth Summit to Local Agenda 21: Working Towards Sustainable Development*. London: Earthscan.
- Lafferty, W. M.; and O. Langhelle, P. Mugaas and M. Holmboe Ruge (eds) (1997). *Rio +5: Norges oppfølging av FN-konferansen om miljø og utvikling*. Oslo: Tano-Aschehoug Forlag.
- Lafferty, W. M. and O. Langhelle (1999). *Towards Sustainable Development* London: Macmillan Publishers.
- Lafferty, W. M. and J. Meadowcroft (eds) (1996). *Democracy and the Environment: Problems and Prospects* Cheltenham: Edward Elgar
- Lafferty, W. M. and J. Meadowcroft (2000). *Implementing Sustainable Development: Strategies and Initiatives in High-Consumption Societies*. Oxford: Oxford University Press.
- Lafferty, W. M. and N. Raaum (1992). *Makt og Politikk med en Annen Røst ("Power and Politics in Another Voice")*. Research Report No. 02/92, Department of Political Science, University of Oslo.
- Lafferty, W. M. and E. Rosenstein (eds) (1993). *The International Handbook of Participation in Organizations*. (Vol. III: "The Challenge of New Technology and Macro-Political Change"). Oxford University Press.
- Martinussen, W. (1973). *Fjerndemokratiet ("The Distant Democracy")*. Oslo: Gyldendal.
- Mayne, J. and E. Zapico-Goñi (eds) (1997). *Monitoring Performance in the Public Sector*. New Brunswick: Transaction Publishers.
- Nagel, S. S. (2002). *Handbook of Public Policy Evaluation*. London: Sage Publications.
- ProSus (2000). *Strategi- og virksomhetsplan for ProSus, 2000-2005 ("Strategy and Action Plan for ProSus, 2000-2005")*. Oslo: Centre for Development and the Environment, University of Oslo.
- Rieper, O. and J. Toulemonde (eds) (1997). *Politics and Practices of Intergovernmental Evaluation*. New Brunswick: Transaction Publishers.
- Rist, R. C. (1995). *Policy Evaluation. The International Library of Comparative Public Policy Vol. 3*. Cambridge: University Press.
- Skogerbø, E. (1988). *Demokratiperspektiv på lokalfjernsyn: En studie av Jevnaker lokal-TV ("Democratic perspectives on local television: A Study of Jevnaker Municipality")*. Report No. 2. Oslo: Department of Media and Communication, University of Oslo. (Master's thesis in Political Science)
- Skogerbø, E. (1996). *Privatising the Public Interest: Conflicts and Compromises in Norwegian Media Politics 1980-1993*. Report no. 20. Department of Media and Communication, University of Oslo. (Doctoral dissertation in Political Science)
- Strømsnes, K. (1993). *Politisk Deltakelse I Norge: Variasjoner over et Tema ("Political Participation in Norway: Variations on a Theme")*. Master's thesis in Political Science, Institute for Comparative Politics, University of Bergen.
- Vedung, E. (1997). *Public Policy and Program Evaluation*. New Brunswick: Transaction Publishers.



**Wolfgang Meyer**

## **SOCIOLOGICAL EVALUATION – METHODS AND LIMITS**

### **Abstract**

Realization of sustainable development is impossible without participation of the majority of people and organizations. There is a rising interest in evaluation research by stakeholders like national state agencies or international political organizations because of limited progress achieved by their programs. But scientific methods and general orientation of evaluation research are not only useful for the judgement of huge policy programs. They can even help low-budget organizations to improve their performances and results. This article emphasizes on three closely related aspects: the development of target systems, measurement problems and the construction of adequate indicators, and finally the embeddedness of these tools within more complex and durable management systems. Presentation includes a short balance of practical use, an outlook on new challenges related with sustainable development, and some arguments why impact monitoring and evaluation systems using the equipment of evaluation research could probably be a solution to some existing problems. Concluding, some future tasks for the development of appropriate sustainability management systems including applied methods of evaluation research will be addressed to different actors.

### **1. Introduction<sup>24</sup>**

For the world-wide discussion on sustainable development, the year 2002 is a very important year. In autumn, the World Summit at Johannesburg, South Africa, will balance the progress since the influential World Summit in Rio de Janeiro 1992 and, hopefully, will leave us with new guidelines, international commitments and political decisions to continue the process. Surely, many people and organizations will criticize the former development as non-satisfactory facing the tremendous global problems which are still waiting for proper solutions (among others see for example South African NGO Caucus 2001: 16f.). In addition it is easily foreseeable that after the meeting in Johannesburg some people and organizations will be dissatisfied again. But irrespective whether the results of Johannesburg will be seen as an important step forward or as a great disappointment, it is quite sure, that the majority of people on earth does not even take notice of this conference. And if they do so, they probably look at the Summit as a big palaver and a stage for second-class politicians, trying to get on one's soapbox from time to time. The World Summit – no doubt – will be a meeting for world elites and even in the so-called developed countries at least the majority of people will not be very interested and sometimes even annoyed with this meeting.

Obviously, there is a great variety of reasons for this lack of interest. Among others, many people are of the opinion that this discussion and its topics does not really bother them and their private life. The solution of global problems is merely seen as the task of nation states and international organizations while the contribution of a single person, enterprise or private association is attributed as to be too small to change anything at all. Comparable to environmental problems, sustainable global development seems to be a too complex task for average people to handle. "Think global – act local" needs someone to coordinate collective action and to tell us what to do. And there seems to be no reason why we should spend some of our spare time to develop ideas for this abstract concept of sustainable development and to leave our comfortable position of a "free rider" in world development.

However, there have been some changes since Rio – not only in political action but also in people's minds. Business organizations realized that environmental protection is not mainly a threat but a big chance to open up new markets and increase their own image at the consumers level. Economic associations try to avoid regulative states policy by self-obligations and by developing their own strategies for environmental protection. Non-governmental organizations with environmental targets like Greenpeace learned about the importance of economic and social factors for their own success and started to co-operate with industrial partners instead of

---

<sup>24</sup> The author wants to thank Angelika Nentwig and Nicolà Reade for helpful comments and language improvement. They are, of course not, responsible for any of the remaining mistakes.





seeing them solely as their enemies. Nation states and their bureaucracies introduce indirect measures to set incentives trying to expand the scope of action for different social groups and to improve their inclusion in political decision processes ("round table policy"). They launch support programs to increase the offer for informing and consulting several different target groups about environmental topics and questions of sustainability. Transnational commitments homogenize more and more and lead national developments towards the targets of sustainable development (e.g. the local agenda 21 process for communes). And especially in Austria, Germany, and Japan a remarkable group of enterprises implemented environmental management systems and sometimes spent a lot of money for running them.

To sum it up: although the majority of people is still not interested, there are a lot of promising initiatives to disseminate the idea of sustainability and to reach several different groups within societies. The diffusion of sustainable development is yet not an automatic social process which would run without any initiatives of transnational organizations or nation states. Because these interventions are very difficult (and sometimes frustrating) tasks, the initiators more often want to know something about the impacts of their used measures. This is one reason for rising interest in evaluation research at transnational organizations and nation states in Europe.<sup>25</sup>

By using an extensive definition of evaluation, everybody evaluates in one or the other way its own action and the achieved impact. The task of evaluation *research* is to systematize observations, to improve measurement and to objectify judgements by using scientific methods. Rossi et al (1999: 4) for example define program evaluation as "the use of social research procedures to systematically investigate the effectiveness of social intervention programs". Nowadays, there are a lot of more or less similar definitions and a great variety of evaluation concepts for different policy fields, application areas and various types of problems.<sup>26</sup>

Scientific evaluation differs from other scientific studies as well as from auditing. While scientific research is usually oriented to expand human knowledge without any practical interests, evaluations are undertaken for more practical reasons and shelve their contribution to scientific progress. "Audit is primarily concerned with verifying the legality and regularity of the implementation of resources (inputs) in a programme. Evaluation, on the other hand, is necessarily more analytical." (Nagarajan and Vanheukelen 1997: 13).

While in most cases state organizations and other stakeholders are clients of evaluation researchers, scientific evaluation is also useful for every organization including small firms and non-profit organizations, too. One major effort of evaluators is to develop useful concepts for self-evaluation and continuous monitoring within all kinds of organizations. Of course, expensive external evaluations are only needed for huge programs or projects. But scientific evaluation methods are obviously also useful for small projects and working units. Moreover, to realize the ambitious goals of sustainable development, a common use of scientific evaluation methods is indispensably needed. For the diffusion of evaluation methods for sustainability, the development of appropriate toolboxes for different social groups is required. Although some fundamentals already exist, this remains as one of the most important future tasks for evaluation research. This is the main message of this paper.

This article wants to introduce some important aspects which should be mentioned when running any kind of project (business, private or public projects) or intervention. Three different but closely related topics will be presented. First, requirements and problems of developing appropriate target systems will be discussed. The need for hierarchically formulated and systematically connected goals will be shown and some basic rules of formulation will be presented. Moreover, the reasons for resistance against these widely accepted rules and the importance of target systems for sustainable development will be addressed.

Goal-attainment and impact control are most important aspects for further development of target systems. This is strongly connected with measurement problems and the use of appropriate indicators. In a second step, a systematic overview on different kinds of indicators and some rules for indicator construction will be presented.

---

<sup>25</sup> For the development of evaluation research in Europe see Leeuw (2000). Referring to environmental evaluation, evaluation activities are still very poor and the leading actors have only just begun to discuss the possibilities and the need for evaluation research. For the Austrian and German example see Meyer and Martinuzzi (2000). Even in the United States, where evaluation research has a quite longer tradition and the research infrastructure is much better developed, environmental evaluation is only a matter of peripheral importance (see Rich 1998).

<sup>26</sup> Some examples are: Breakwell and Millward (1995), Chelimsky and Shadish (1997), Chen (1990), Conseil scientifique de l'évaluation (1996), Mohr (1995), Nagarajan and Vanheukelen (1997), Patton (1997), Rossi et al. (1999), Scriven (1991), Stockmann (1997), Worthen et al. (1997). For the use of evaluation methods for environmental policy programs see Knaap and Kim (1998).



Again the specific problems of formulating sustainability indicators especially for low-budget organizations will be kept in focus.

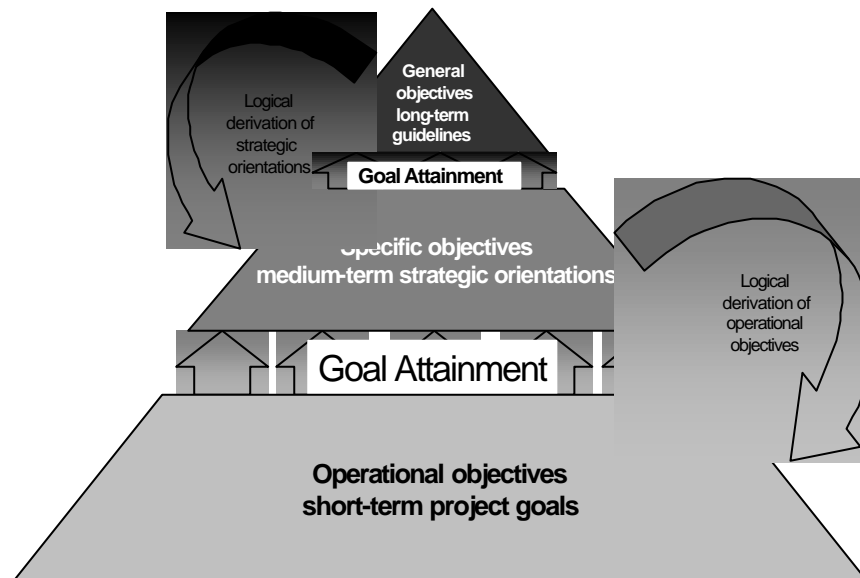
For managing sustainable development by using appropriate target and indicator systems, the implementation of a corresponding sustainability management system is indispensable. Restrictions set by a lack of resources within the representative organization have to be mentioned when developing such a sustainability management system. Existing TQM and environmental management systems have been proved to be too costly for low-budget organizations. In chapter three, some outlines for implementing adequate monitoring and evaluation systems with respect to further development for sustainability management will be presented.

Although this article tries to give some useful information for practical work, it will of course not be able to solve the problems related with the development of appropriate sustainability management systems. In conclusion, some future tasks derived from the presented ideas will be defined and addressed as responsibility to different groups.

## 2. Developing Target Systems for Sustainable Development

The first important step to sustainable development is designing a guiding target system which have to be continuously adapted to social changes. This target system should lead decision makers step by step to more and more complex goals. Objectives should be hierarchically ordered (Figure 1): on the first level, there should be short-term objectives which could be obtained in a well-defined time period. The achievement of these short-term objectives should add to the attainment of medium-term strategic orientations. Again several of these strategic orientations should sum up to the realization of a long-term guideline.

Figure 1 Hierarchical Target System



Source: Meyer (2000), slightly modified

To give you an example: if your long-term guideline is the realization of technically optimized reductions of energy consumption within your enterprise (community, private or public household etc.), it might be possible to do several pilot projects to find out the best reduction strategies for different situations, subgroups, departments or working areas. Your medium-term strategic orientation for these projects might be to decrease the energy consumption within a well-defined social surrounding by using a limited number of testable measures. In most cases this bundle of used measures to achieve this strategic orientation might be too inhomogeneous to observe them as a whole. One of these used measures could for example be the change of thermostats in one room and the related objective is to reduce the energy consumption in this room. Again the achievement of several objectives should lead to the fulfillment of the strategic orientation like pieces in a jigsaw puzzle.



Such a target system form is not new nor too complicated to realize it in any kind of social system (beginning with your private household and ending with global organizations). Although all experts agree to the effectiveness of the presented target system forms and a lot of scientific and even popular literature has been published on this topic<sup>27</sup>, reality is far away from realizing such an easy and understandable planning instrument. In most cases goals are primarily politically formulated, which means, that they will be reached anyway. The reason again is very easy to understand: if I'm trying to jump ten yards long, I will surely fail to reach my goal – but if I want to jump ten inches, I will always win (and everybody likes to win). Therefore, in nearly every case you will hear that persons or organizations have achieved their goals. But unfortunately, in many cases this is not a big success.

Another strategy to avoid the risks to fail when running a project is to use vague formulation of goals. For example, if I define my goal to do a very long jump and don't bother about the length or accept every length as "a very long jump", there will be no problems to achieve my goal. Probably, you consider these examples for extremes, but even in huge organizations with a fair amount of money invested in projects this is the reality. For example, in an evaluation study in commission of the federal environmental agency in Germany we analyzed more than 30 projects of almost the same amount of federal associations.<sup>28</sup> Besides some slight exceptions almost none of these organizations used a well-developed target system to run environmental projects or even had an idea about their medium- or long-termed orientations on environmental topics. In consequence, nobody trusts reports of success, although everybody likes winners (and avoids personal contact with losers). As the main driving force of sustainable development, continuous learning processes are therefore very difficult to establish.

On the other hand, if there's no possibility to fail, there is no chance to learn from faults and to reach some future progress as well. To design a realistic and guiding target system we need the courage to formulate risky goals and to accept that we will not reach everything we wanted to reach. We should not conceal uncomfortable results but take the chance to learn about the reasons why we failed. In many cases it is even much easier to learn from failures than from success. Only then we will get the possibility to avoid the same mistakes in different situations and to improve our future performances.

However, the first step will be to reform our practice of defining goals. Every goal definition should have five key elements (for details see BUWAL 1997: Chapter 4.2):

- object (who should be reached?)
- content (what should be reached?)
- extent (how much should be reached?)
- area (where should it be reached?)
- time period (when should it be reached?)

Likewise, a useful goal of this article could be formulated as following: Within six months ten journal readers in all European nations should read this article and as a result change their practice to define goals. Whether this ambitious goal will be reached or not could be measured by suitable indicators and the resulting information could be used for adjustments or re-formulation of goals for a better adaptation to reality if needed. The most important question, not only for this example, will be why some goals are achieved by using these means and why others are failing in doing so. To analyze these causal connections is the main task of evaluation research.

---

<sup>27</sup> Target systems are for example discussed in nearly every book on project management. There are several textbooks and introductions in various languages from scientists of different disciplines and countries. Some examples are: Andersen et al. (2001), BUWAL (1997), Diethelm (2000), Gareis (1990), Lientz and Rea (2001), Lock (2001), Meredith and Mantel (2000).

<sup>28</sup> These federal associations include some of the most important environmental NGO's e.g. NABU or BUND, powerful economic alliances like the chambers of commerce, huge public organizations like the Deutsche Landkreistag or the Hochschulrektorenkonferenz and professional organizations like the Deutsche Architektenkammer. For further information on the evaluation study take a look at the English summary in our research report (Meyer, Jacoby and Stockmann 2000). See also the comparable results of another evaluation study in Stockmann et al. (2001).



The continuing process of adjustment and adaptation of target system is the heart of managing sustainable development. Sustainability means durable effect control and rational reaction to collected information. Sustainability is a continuing exchange between target system and observed results. Having the long-term guidelines in mind, short-term success or defeat loses their importance. They are only steps towards the future goals or reasons to change courses – sometimes only to make a detour. On the other hand, short-term success might be the starting point for reformulating goals and objectives – mostly to increase demands.

One important aspect of target systems is to mark the priority of goals. Sustainability usually means preference of long-term to short-term goals. But of course the importance of the contribution to subordinated aims varies within various objectives. Therefore, a differentiation between objectives should be used. One possibility is to distinguish goals which have to be achieved (“Must-goals”), should possibly be achieved (“Should-goals”) or can additionally be achieved (“Can-goals”).

And finally sustainability management is some kind of Sisyphean task – it will never end. As soon as long-term guidelines have been achieved, new guidelines have to be formulated. Developing target systems, adjust or adapt goals to reality and permanent control of measurement effects are the continuing burden of sustainable development.

As mentioned before, we are quite far away from this ideal of rational sustainable management. Especially for environmental tasks non-governmental organizations, at least in Germany, are strongly oriented at the topics set by stately sponsorship and their actions are in general not the result of an internal, systematic development of target systems. And even the states behavior to support environmental projects is far from being continuous – there are several breaks not only stimulated by political changes but also influenced by public interest and fashions. Of course, there have been some achievements gained but actually one can not speak of a well-developed social infrastructure for sustainability management. In most cases not even an initial stage has been achieved.

Furthermore, one has to recognize that implementing environmental targets and on-going systematic development of these targets is only one element of sustainability. Surely, there are some promising outlines to integrate environmental and economical targets and – very seldom – to integrate environmental and social targets. But as far as can be surveyed, uniting all three targets has not yet been successfully realized anywhere.

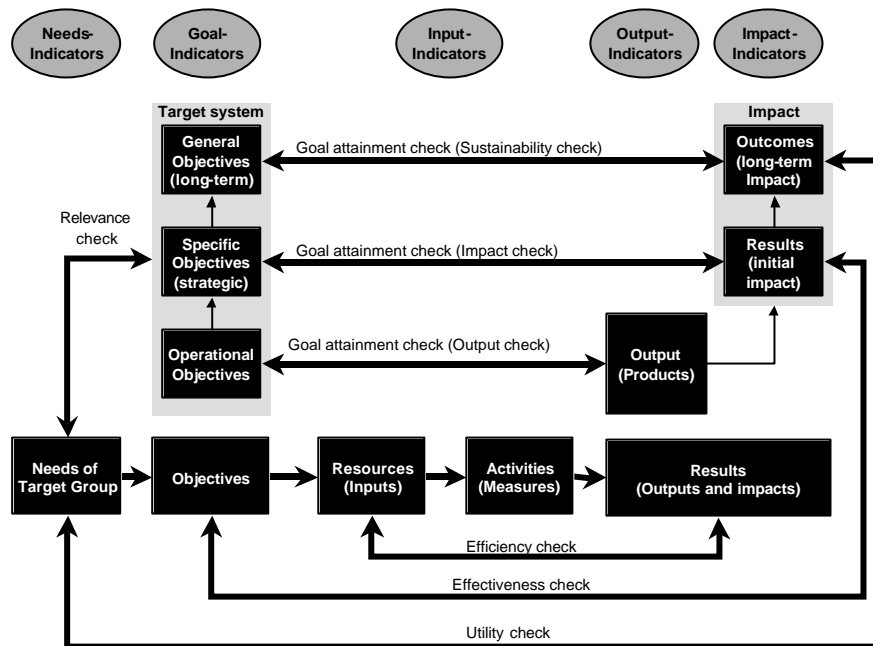
Especially for small and medium sized firms or poorly financed non-profit organizations the requests of sustainability management are not easy to handle. Those organizations lack of resources (personal, money and time) and normally they do not have the needed scientific know-how to implement ambitious target systems. Standardized and practicable model systems which could be used as a portfolio still do not exist. There's a visible need for developing negotiable target systems for small units which could be spreaded as examples for best practice. In order to realize this, barriers of professional jealousy within business sector will have to be overthrown – a problem which frequently stops diffusion of innovations between enterprises.

### **3. Developing Indicators for Sustainable Development**

For continuous control of the effects of implemented measures the development of adequate indicators is needed. There are seven different possibilities to check effects with corresponding indicators constructed for its measurement (Figure 2).



Figure 2 Indicator systems and effect control



Source: Nagarajan & Vanheukelen (1997); modified and supplemented Meyer (2000)

- efficiency (input-output-analysis – the relation of used means and realized output)
- effectiveness (comparison of aimed and achieved results)
- utility (comparison between needs of the target group and sustainable changes)
- relevance (comparison between needs of the target group and strategic orientations)
- goal attainment
  - a) output (realization of planned measures)
  - b) immediate outcome (realization of strategic results)
  - c) sustainable outcome (realization of durable changes)

In general, the goal of these checks is to optimize the correlations within the causal chain from the needs of the target group to the long-term impacts. As far as possible, the formulated target system should be a perfect expression of target group needs. And for to achieve proper results, precisely those resources needed for necessary activities should be used. Any weak link within this chain of causes and effects threatens success. Because of the great variety of influences, in reality these links could not easily be observed. Indicators are simplifications of this more complex reality (Nagarajan and Verheukelen 1997: 19). They should help us to understand interrelations, to identify problems and to find their causes. Therefore some general requests for adequate indicators could be formulated:

- quality of measurement (indicators should be valid and reliable instruments which steadily produce information on the topics they are formulated for)
- usability (indicators should be easy to handle and understandable)
- applicability (indicators should have a clear interpretation and an unambiguous differentiation between variable results)
- punctuality (indicators should give us information right in time before any decisions have to be made)

Because of the great variety of possible indicators, the question who should construct indicators is not easily answered. Referring to the topics of sustainability, there are huge differences in competencies and the possibility to get support from the consulting market. No doubt, businessmen normally have a lot of experience by using economic indicators to control market development, income and expenditure. And there are professional consultants offering their help if there are any problems.



In Germany (and presumably in a number of other states, too), there are also environmental consulting and auditing enterprises with competencies in developing and implementing environmental indicator systems. Furthermore, economic associations like the chamber of commerce offer their help in environmental questions, too. But yet there are no organizations with enough experience in developing sustainable indicator systems, especially for the use of low-budget organizations. Even for the construction of social indicators not much support is to be found on the consulting market.<sup>29</sup>

But fortunately, in many cases professional (and expensive) support for developing indicators is not needed. Even simple indicators can give us some insight into impacts of used measures. For several problems, standardized indicators are available and can easily be transferred to personal requirements. In addition, it is possible to measure things quite accurate things which seems to be immeasurable by using “proxy-indicators” which are highly correlated with the “true” concept – this is for example the way psychologists develop their indicator systems to measure constructs as “happiness”. Sometimes indicator construction is more a question of fantasy than of specialized know-how.

On the other hand, because of measurement limits the construction of a “perfect” indicator is impossible. And there is still another insuperable dilemma between scientific demand and practical possibilities to realize this demand. Even experts with many years of experience in constructing indicators can’t find perfect solutions for these problems. Also well-known and broadly used indicators like the gross domestic product could be discussed and there are in fact good reasons to criticize them. Finally, the most important thing when constructing an indicator is to recognize the problems and limits of expressiveness which will give us the chance to improve measurement by adding other indicators avoiding these limitations.

The construction of indicators and appropriate indicator systems is therefore primarily a question of basic knowledge about the use and the property of indicators. Experts can help to reduce generating costs and to improve the quality of indicator systems, because they know a lot about practical measurement and the requests for constructing indicators. Without mentioning measurement, the development of indicators might lead to unsolvable problems when collecting the necessary data.

Especially in low-budget organizations, the staff has usually has no knowledge and experience in social scientific measurement and the problems related with data collection. On the other hand, external experts first have to learn about the opportunities to measure within the organization. These conditions are normally best known by the staff. Therefore, the optimum constellation is bringing methodological experts and organization staff together for the development of indicator systems.<sup>30</sup>

By developing indicator systems for sustainable development some special requests have to be taken into account. First of all, three different dimensions (ecological, economical and social) should be measured simultaneously and be related in an adequate way. Secondly, long-term impacts and changes should be viewed and properly reported. And thirdly, needs and interests of a great variety of social groups should be represented and they should also be involved into decisions concerning indicator construction. Besides the unsolved practical problems in doing so, up to now there are no indicator systems and measurement solutions which seem to be able to fulfill both the special requests of sustainable development and the above presented requirements for indicator systems. Such an indicator system could not be constructed by scientists alone, there must be an intensive co-operation between scientific experts and practical working people in different kinds of organizations. Those indicator systems have to be developed for practical use with a strong orientation towards the needs and possibilities in the respective area of application. Up to now, there are only rare initiatives in doing so.

---

<sup>29</sup> In Germany, for macro-level analysis and scientific research the social indicator department at the Centre for Survey Research and Methodology (ZUMA) in Mannheim offers its help. This group is also involved in international co-operations for the development of sustainability indicators. Because the units of these indicators are nation states and therefore national statistic systems or regular conducted representative surveys are used for data sources, these indicators can not easily be transformed to micro-level analysis or to the application of various organizations.

<sup>30</sup> This is for example the way we try to implement monitoring systems in development projects all over the world (see for an actual example in Mexico: Jacoby 2001).



#### 4. Developing Management Systems for Sustainable Development

For sustainable development, target and indicator systems have to be included in a systematic management system. One important element of this management system is some kind of regular effect controlling. Today, several well-developed forms of management systems have been implemented and standardized with international norms (e.g. ISO-norms), like for instance Total Quality Management (TQM) or Environmental Management systems. Of course, these systems include effect controlling as one important element.<sup>31</sup> And as continuously working and durable implemented systems, they have some important aspects of sustainability in mind.

Nevertheless, there are still a couple of limitations. Firstly, none of those systems really integrates appropriate the three dimensions ecology, economy and social which are in the focus of sustainability debate. Secondly, most of the existing standardized management systems concentrate their observation tools aim on output or product control just for doing some input-output analysis. Some more developed management systems are heading for process analysis instead of target-performance comparison but again usually only production processes, sometimes also marketing processes are taken into account. Open observation systems, understanding interventions to implement innovations primarily as a social process with complex, unforeseeable results and impacts, are yet only rarely used. And finally the most important factor for low-budget organizations is: implementation and continuation of quality and/or environmental management systems are very expensive. Moreover, these systems seem to be more obliged to external criteria and standards than to the own possibilities and needs. Although the benefits of these systems are largely uncontradicted (environmentally as well as economically) the majority of small and medium sized firms and nearly all non-profit organizations are still skeptical to the durable implementation of TQM or environmental management systems.

Impact monitoring and evaluation systems (M&E-systems) have a totally different starting point. They are not bound for some external standards and they are not aiming on some kind of certification. The goal of M&E-systems is simply to improve knowledge about impacts for decision makers. They do not want to be “perfect” in measurement with respect to the state of the art in social sciences, but they want to present valid and reliable information as far as this can be realized within the existing social surrounding. Optimizing measurement is not a goal for itself but an important basis for rational decisions to achieve aspired impacts and to improve positive results.

Four main functions of monitoring and evaluation can be distinguished (Figure 3):

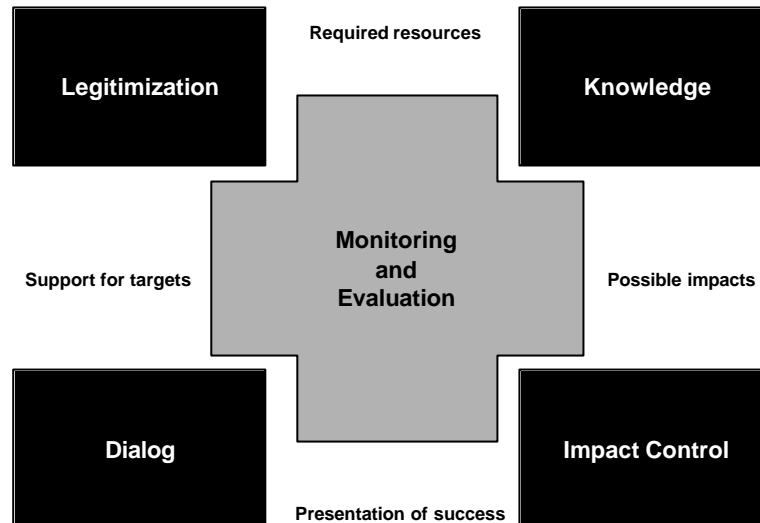
---

<sup>31</sup> See for example DIN (1998: EN ISO 9001-2000-12, Chapter 8) for quality management.



Figure 3 Functions of Monitoring and Evaluation

Source: Meyer (2000)



- Knowledge (collected data should help to recognize possible impacts and its required resources)
- Impact Control (collected data should document impacts and serve as realistic and reliable source for external presentations)
- Dialog (collected data should verify successes to significant others and therefore promote external support for the targets)
- Legitimization (collected data should justify the former and planned use of required resources to supporters)

M&E-systems should not be recognized as competitors or as simple extension of TQM and environmental management systems. As far as they are only developed for continuous measurement, they can not overtake the extensive tasks of management systems. On the other hand, they are also not simply a modified form of effect controlling because of the strict orientation on the social processes behind it instead of putting the focus primarily on output effects as controlling systems do. As an open procedure M&E-systems will always change depending on perceived new impacts or changing observation interests of all involved social groups. Because they are formulated as social observation systems, they can be used universally for any kind of impact research irrespective of the issue. Effect controlling in environmental management systems, for example, are strictly associated with environmental topics while TQM is strongly oriented on customers needs and interests. TQM therefore entirely ignores other important groups and their requests (e.g. politicians, competitors, stakeholders, the staff and its families) as far as they are not concerned with product quality.

The main characteristics of M&E-systems are:

- Orientation on social processes related with implementation of innovations
- Extensive view on interests of participated groups and all kinds of impacts attributed to interventions by people who had been involved
- Participatory development of indicator systems including for example organization staff and external experts in methods of empirical social research
- Regular measurement of impacts with an appropriate and continuously improved indicator system (monitoring)
- Periodical analysis of causes and effects related with observed impacts (evaluation)
- Results of M&E-systems are intended to be important information for decision processes for improving impacts
- M&E-systems are concerned in the development of target systems which defines the focused areas but they are not limited on measurement of pure goal attainment questions





M&E-systems should be seen as useful and pragmatically formulated tools for decision processes which can be easily implemented and improved with respect to an organizations potential. They fulfill some of the basic requirements for implementation of sustainable development:

- Integration of ecological, economical and social goals: M&E-systems are open to any kind of interest and therefore suitable for bringing the main targets of sustainable development together
- Participation: without respect to all kind of information from and the different views of various groups involved in the intervention process, M&E-systems can not work
- Sustainability: M&E-systems are designed for durable use, should be steadily improved and have to be regularly adapted to changing social surroundings

M&E-systems are only one element of sustainability management which has to include important aspects of quality and environmental management as well. To design such a complex management system is a task yet not be realized. However, the implementation of appropriate M&E-systems can be an important first step. Especially for low-budget organizations, which consider TQM or environmental management system for being too costly and difficult to implement, M&E-systems are probably an adequate solution.

## 5. Conclusions

Sustainability management requires at least a M&E-system for measuring impacts achieved by any kind of intervention. Important tasks for running such a system are the continuous improvement of target and indicator systems. Hierarchically formulated target systems assure successive and complementary attainment of goals and objectives. Appropriate indicator systems for different observation problems try to deliver information needed for decisions to improve performance and results. Hence, sustainable management has to put its focus on social processes to implement innovations and their adequate measurement.

The actual reality is still far away from some kind of effective sustainable management. Even the well-known and highly respected utility of hierarchically formulated target systems does not lead to a widespread use. Although some organizations use outstanding controlling systems (sometimes already for environmental concerns), impact monitoring is hardly to be found in any organization at all. The development of social and, even more, of sustainability indicators nowadays is entirely the task of academics. For the use of low-budget organizations, no appropriate indicator tableaus exist. But some extraordinarily examples show, that even in very small non-profit organizations useful information can be collected by simple, self-developed indicators. Nevertheless, in Germany, the majority of associations, especially when they are big and economically better situated, does not bother about impact measurement. To what extent these results can be generalized for other organizations or for other countries is still an open question.

Target and indicator systems have to be included in impact monitoring and evaluation systems, guaranteeing continuous supply of information and its adequate measurement just in time when they are needed for decisions. To develop the scientific basis of such a monitoring and evaluation system is the target of evaluation research. Nowadays, evaluations are almost exclusively used for judgments about the impact of public projects and programs, having nation states and public organizations as stakeholders. In Europe, infrastructure for evaluation research is still in construction and yet poorly developed. For systematic improvement of sustainable development public investment in this area of applied scientific research has to be raised and basic research institutes (like for instance in Germany the Max-Planck- or the Fraunhofer institutes for a broad spectrum of other scientific topics) have to be implemented.

Finally, M&E-systems have to be embedded in sustainability management systems which should also include elements of existing quality and environmental management systems. Beside of some modifications needed to adapt existing management systems to the targets of sustainable development, it is most important to increase the usability especially for low-budget organizations. Nowadays, the majority of such organizations considers management systems for too expensive and too difficult to handle for their own purposes. For a widespread use of sustainability management, appropriate M&E-toolboxes for different kinds of organizations must be developed in co-operation between staff and external experts. To finance such kind of projects should be one important target for sponsorship. Sustainable development can not be realized without the participation of low-budget organizations such as non-profit organizations or small and medium sized firms. To help them by implementing appropriate impact monitoring and evaluation systems is the first step for developing professional sustainability management.



## List of References

- Andersen, E. S., K. V. Grude, and T. Haug (2001) 'Goal directed project management: effective techniques and strategies' (London, UK: Kogan Page, 2<sup>nd</sup> ed.).
- Breakwell, G. M., and L. Millward (1995) 'Basic evaluation methods. Analysing performance, practice and procedure' (Leicester, UK: British Psychological Society).
- BUWAL Fachstellen Naturschutz der Kantone Aargau und Zürich, Abteilung Naturschutz (eds.) (1997) 'Projekte erfolgreich abwickeln. Arbeitshilfe für den Natur- und Landschaftschutz' [Successfully running projects. A working aid for the protection of nature and the countryside] (Buchs, CH: RegioPrint).
- Chelmsky, E., and W. R. Shadish (eds.) (1997) 'Evaluation for the 21<sup>st</sup> Century : A Handbook' (Thousand Oaks, USA: Sage).
- Chen, H.-T. (1990) 'Theory-Driven Evaluations' (Newbury Park, USA : Sage).
- Conseil scientifique de l'évaluation (1996) 'Petit guide de l'évaluation des politiques publiques' (March. Paris, F: CSE).
- Diethelm, G. (2000) 'Projektmanagement. Bd. 1: Grundlagen' (Herne/Berlin, D: Verlag Neue Wirtschafts-Briefe).
- DIN Deutsches Institut für Normung, K. Graebig (eds.) (1998), Qualitätsmanagement, Statistik, Umweltmanagement. Anwendungshilfen und Normensammlungen, Berlin/Wien/Zürich, D/A/CH: Beuth).
- Gareis, R. (ed.) (1990) 'Handbook of Management by Projects' (Wien, A: Manz).
- Jacoby, K. P. (2001) 'Implementación de un Sistema de Monitoreo y Evaluación (M&E). Informe final sobre la implementación de un Sistema de Monitoreo y Evaluación en un proyecto de la cooperación técnica mexicano-alemana (CANACINTRA / HOPE / GTZ: Formación Profesional Orientada al Medio Ambiente)' [Implementation of a Monitoring and Evaluation System (M&E). Final Report on the Implementation of a Monitoring and Evaluation System in a Mexican-German development co-operation project (CANACINTRA / HOPE / GTZ: Environmental oriented Vocational Training)] (Saarbrücken, Mexico D.F.: Camera Nacional de la Industria de la Transformacion; also available in German).
- Leeuw, F. L. (2000) 'Evaluation in Europe', in: R. Stockmann (ed.), Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder [Evaluation Research. Basics and selected research fields] (Opladen, D: Leske+Budrich): 57-76.
- Lientz, B. P., and K. P. Rea (eds.) (2001) 'Project management for the 21<sup>st</sup> century' (San Diego/London, USA/GB: Academic, 3rd ed.).
- Lock, D. (2001) 'The essentials of project management' (Aldeshot, USA: Gower, 2<sup>nd</sup> ed.).
- Knaap, G. J., and T. J. Kim (eds.) (1997) 'Environmental Program Evaluation. A Primer' (Urbana/Chicago, USA: University of Illinois Press).
- Meredith, J. R., and S. J. Mantel Jr. (2000) 'Project Management – A Managerial Approach' (New York, USA: Wiley, 4<sup>th</sup> ed.).
- Meyer, W (2000) 'Umweltberatungsprojekte – Ein Ratgeber' [Environmental Consulting Projects – a Guide] (Bremen/Saarbrücken, D: Bundesverband für Umweltberatung).
- Meyer, W., K.-P. Jacoby, and R. Stockmann (2000) 'Evaluation der Umweltberatungsprojekte des Bundesumweltministeriums und des Umweltbundesamtes. Nachhaltige Wirkungen der Förderung von Bundesverbänden' [Evaluation of Environmental Consulting Projects sponsored by the Federal Environment Ministry and the Federal Environmental Agency. Sustainable Results of supporting federal associations], Research-Report with an english summary (Saarbrücken, D: Saarland University; to be published within the UBA-Texte series).
- Meyer, W., and A. Martinuzzi (2000) 'Evaluationen im Umweltbereich. Ein Beitrag zum nachhaltigen Wirtschaften?' [Environmental Evaluation. A contribution to sustainable economy?], Vierteljahreshefte für Wirtschaftsforschung 3/2000: 449-463.
- Mohr, L. B. (1995) 'Impact analysis for programme evaluation' (Thousand Oaks: Sage, 2<sup>nd</sup> ed.).



- Nagarajan, N., and M. Vanheukelen (1997) 'Evaluating EU Expenditure Programmes: A Guide. Ex post and Intermediate Evaluation' (Luxemburg, LUX: European Commission XIX/02 – Budgetary overview and evaluation, EU).
- Patton, M. Q. (1997) 'Utilization-focused Evaluation' (Thousand Oaks: Sage, 3<sup>rd</sup> ed.).
- Rich, R. F. (1998) 'Program Evaluation and Environmental Policy: The State of the Art', in: G. J. Knaap and T. J. Kim (eds.) 'Environmental Program Evaluation. A Primer' (Urbana/Chicago, USA: University of Illinois Press): 23-44.
- Rossi, P. H., H. E. Freeman, and M. W. Lipsey (1999) 'Evaluation. A Systematic Approach' (Thousand Oaks u.a., USA: Sage, 6<sup>th</sup> ed.).
- Scriven, M. (1991) 'Evaluation Thesaurus' (Newbury Park, USA: Sage, 4<sup>th</sup> ed.).
- South African NGO Caucus on the World Summit for Sustainable Development (2001) 'Towards the World Summit on Sustainable Development, South Africa, 2002', World Summit Papers of the Heinrich Böll Foundation No. 1 (Berlin, D: Heinrich Böll Foundation).
- Stockmann, R. (1997) 'The Sustainability of Development Cooperation' (Baden-Baden, D: Nomos).
- Stockmann, R., W. Meyer, H. Gaus, U. Kohlmann, and J. Urbahn (2001) 'Nachhaltige Umweltberatung. Evaluation eines Förderprogramms der Deutschen Bundesstiftung Umwelt' [Sustainable Environmental Consulting. Evaluation of a support program by the German Federal Environmental Foundation] (Opladen, D: Leske + Budrich.)
- Worthen, B., J. R. Sanders, and J. S. Fitzpatrick (1997) 'Program Evaluation: Alternative approaches and practical guidelines' (White Plains, USA: Longman).



## Clive Spash

### VALUING SOCIAL CONSEQUENCES OF ENVIRONMENTAL CHANGE: ECONOMIC METHODS AND THEIR LIMITS

#### 1. Introduction

This paper provides a survey of methods of valuation in economics as an approach to assessing social welfare resulting from environmental change. In doing so I will:

- (a) outline a range of values associated with environmental change;
- (b) set out the methods of cost-benefit analysis (CBA) used in empirical work;
- (c) explain the main problems associated with those methods; and
- (d) recommend the most appropriate method(s) for empirical work to evaluate the monetary benefits and costs.

Section 2 provides an overview of the environmental values and divides these into market, non-market and non-economic values. Section 3 summarises the principal methods for generating estimates of the non-market values and major problems in their application. The techniques discussed are the travel cost method (TCM) in Section 3.1, the production function approach (PFA) in Section 3.2, hedonic pricing (HP) in Section 3.3, and stated preference techniques, mainly the contingent valuation method (CVM), in Section 3.4. The final section summarises the extent to which these valuation methods can be used in the European context to inform the decision making processes. Emphasis throughout is placed upon awareness of the limitations of these methods, which therefore demand prudent use in any policy context.

#### 2. The range of relevant environmental values

Economic assessment, associated with evaluating externalities, can differ from impact assessment under a natural science or engineering approach where physical impacts are central. That is, under the economic approach, emphasis is placed upon the physical impacts only to the extent that they follow a path to specific targets which affect human welfare. Environmental change is then linked to human welfare via characterising the environment as goods and services (which may be broadly defined). Thus, changes in the provision of environmental goods and services form the focus of attention. This means economic value categories relate to what are regarded as relevant welfare generating aspects of the environment in terms of the environmental goods and services provided rather than the source of physical changes, e.g. loss of ecosystems or biodiversity. Another difference between economic values and physical changes is that the expectation of an impact and its psychological effect can be important in terms of economic welfare, even if there is no physical change, as, for example, in the case of nuclear power stations. Economic valuation techniques appeal to human preferences and as a result are influenced by whatever impacts upon those preferences and ignores whatever those preferences fail to take into account.

##### 2.1 Potential impacts on existing markets

Goods and services sold directly to consumers or to firms as inputs to production may be impacted by certain environmental changes. For example, air pollution may impact agricultural crops raising production costs, shifting supply functions and raising prices in the market place. In such cases a market exists where goods and services are traded so that standard economic models of supply and demand in that market can be employed to estimate the economic impacts. Producers may find the costs of supply rise due to pollution and may reflect the cost of additional treatment of inputs or outputs in accordance with governmental or international standards, or the cost of paying others to change their practices. Where environmental quality is reduced, industry reassurances may be insufficient to prevent consumer substitution away from the supplier, thereby reducing demand. In order to capture these effects a market model both consumer and producer welfare surpluses to be estimated.



Pollution from production processes may also reduce the option for use of products which are currently regarded as non-commercial. For example, recreational and tourist opportunities may be temporarily or permanently lost. The impact of lost demand for recreation and tourism would be felt in the local economy if no alternative (perfect) substitute site were available in the area for the activities involved, and a regional substitution effect might occur. The site may have unique characteristics with no opportunities for substitution while others have close market substitutes. Thus, non-commercial harvesting of food can be related to the markets for the commercial substitutes to estimate the value of the product loss. However, in general, recreational activities involve an experience related to the quality of the environment which depends not only upon site characteristics and the availability of substitutes, but also the value gained from participation in the activity. Thus, recreational and tourist activities can involve values beside those costs associated with related markets or identifiable physical aspects of ecosystems. For example, gathering flowers, wild berries or mushrooms may be related to market substitutes but the experience of harvesting is of more central concern to the welfare gained. This means people may incur travel costs in excess of the value of the harvest gathered, as measured by the equivalent commercial product costs, because the experience is valued in itself. If markets are directly related to an activity (e.g. travel costs, entry fees, equipment costs), the opportunity cost can be approximated but not the welfare gained (i.e. consumer surplus). The welfare in terms of surplus would require using a technique which would derive a surrogate demand function for the site's non-market values (e.g. CVM, or TCM as discussed below).

## **2.2 Non-market values**

The activities outlined above tend to result in either direct markets being set up to buy and sell services or goods related to the environment, or in impacts upon uses which are directly related to market activities. However, non-market aspects and non-marketable values are associated with the impacts of environmental changes. These values include a range of goods and services including aesthetics, cultural heritage, health, peace and quiet and ecosystems functions & biodiversity.

Several attempts have been made by economists to value aesthetics (see Graves 1991). The aesthetic appeal of a given environment or site involves the subjective perception of what is beautiful or stimulates the emotions. The relationships between a person and their physical environment can form an important part of their identity affecting their feelings of personal and/or social worth, self-efficacy, distinctiveness or uniqueness (see e.g. Twigger-Ross and Uzzell 1996). Hence, the concept of 'landscape' and the particular characteristics of different landscapes (such as rugged hills, a gently rolling lowland valley, or a river plain; or birdsong and 'natural' sounds associated with different landscapes) can form an important aspect of people's identification with, or reaction to, a place and their perception of any changes to the status quo. Such aesthetic/landscape appeal could be partially reflected as a characteristic of the local housing market. Thus, sites of great natural beauty may be associated with specialised markets (e.g. retirement homes, tourism and holiday sites). Where such aesthetic qualities are regarded as rare or unique the lack of substitutes will mean aesthetic enjoyment of the location is reflected by the premium placed upon living in the area (which might then be assessed via an HP approach). Studies which attempted to assess aesthetic values by direct questioning of the public (e.g. CVM) have tended to estimate an individual's preference for a specific aspect of aesthetic quality, such as improved visibility due to better air quality or the introduction of specific tree species). Following this reasoning, aesthetic quality would then need to be identified with specific site characteristics by visitors or local residents. Such reductionism may be difficult or impossible to achieve (e.g. relating aesthetics to the type and quantity of individual landscape features). Questions can also be raised as to the relevance of a current aesthetic preference. That is, some landscapes may be regarded as worthy of preservation for the values they encapsulate and therefore should be removed from the vagaries of consumer choice, fashions and fads.

Cultural, historical and archaeological sites can be disturbed or destroyed by land use practices. Cultural values may be associated with geological features, ecosystems (e.g. an ancient woodland) as well as buildings or ruins. The site specific nature of these values can, in theory, allow estimation of their economic benefits (using for example, HP, TCM or CVM), although in practice defining and measuring what is meant by the concepts of culture and history would prove difficult. The importance of a site in public perception can vary greatly from that of an expert, say an archaeologist or historian. Local or regional perception of a site's cultural and historical features may also diverge strongly from national or international opinion (e.g. World Heritage Sites versus local economic interests).



Mortality and morbidity valued in monetary terms has proven highly controversial. In terms of preference based measures of economic welfare loss, the risk perception of individuals is all important, rather than the judgement of scientists, engineers or other experts. As long as the aim is to learn about individual's preferences and willingness to pay or accept then knowledge is required of the individual's view of the trade-off, including subjective risk assessment (see Freeman 1993, Chapter 8). Thus the fact that health and safety hazards have been largely eliminated through tighter regulation of industries may be irrelevant in terms of the value of the impact that the public perceives. Of course, public preferences based on mistaken beliefs or poor information may be deemed an inappropriate point of reference for policy.

Noise from human activities can prove a substantive impact on well-being. The reaction may be to take ameliorative steps such as installing double glazing in houses and these protection costs could then be used to estimate the damages resulting from noise pollution. However unless the protection is a perfect substitute for the benefits foregone this approach will be an underestimate of the loss. For example, double glazing is only effective as a sound barrier inside a house with the doors and windows closed. A fuller range of economic loss may be reflected in the value of houses and land in the vicinity which could be depressed by the presence of say a busy road so that the externality is capitalised. (In this case HP models might be employed to estimate the damage.)

The functions which ecosystems perform are many and varied. Only a minority of these fall within the framework where they can be bought and sold on markets subject to private ownership. Amongst the most important functions performed by ecosystems are maintenance of climatic stability and nutrient cycles. Biodiversity of ecosystems, genes and species is seen as an important aspect of natural capital and a key to sustainable development. However valuation of biodiversity is complicated by a poorly informed general public and the extent to which people reject market valuation in this area (Spash and Hanley 1995). Direct questioning of the public (e.g. via CVM) may be able to provide some aspects of species value or even ecosystem diversity but is unable to address many of the concerns raised by the need to maintain ecosystems functions and protect biodiversity. Where an identifiable output or service can be related to an ecosystem function and this output or service is connected to a market product the economic value of changes in ecosystem functions may be assessed (i.e. using PFA). Impacts can be complex, highly uncertain or unknown, such as the loss of a site specific species which has never been classified.

### **2.3 Intrinsic value in Nature**

This category of values is by definition outside of the economic calculus to evaluate. The category is mentioned here in order to qualify the discussion over the extent to which economic techniques can achieve a comprehensive valuation of the benefits of the environment. Intrinsic values are related to non-consequentialist and therefore non-utilitarian aspects of the environment. For example, a species may be valued as a food source and because it is beautiful and because of its potential to benefit science, but it may also be valued outside and separately from all these uses or aspects of its nature which create good consequences for humans.

Individuals who conform to economic assumptions regard the world from a perspective where all values can be traded. The ultimate criterion of morality lies in some value (e.g. welfare, utility, happiness) that results from acts. Such theories see only instrumental value in such acts, but intrinsic value in the consequences of these acts. In contrast, deontological ethical theories attribute intrinsic value to features of the act, themselves. This could be apparent as an expression of the rights of animals to welfare or the rights of humans to life. Thus, non-compensatory choices arise and Freeman (1986) has suggested that lexicographic preferences may be taken as a belief in such rights. When preferences are lexicographic, the individual cannot be compensated for the loss of a quantity of one good by increases in the quantity of one or more other goods, no matter how small the former or how large the latter. However this approach reduces the difference between payment offered and compensation demanded to an anomaly within utilitarianism rather than a fundamental difference in philosophical outlook. The refusal to trade becomes particularly relevant when disruption of the environment affects such things as human health, animal welfare and ecosystems functioning and structure. In such cases intrinsic values in non-human animals, plants or ecosystems are recognised by individuals as a serious constraint on economic trade-offs. Studies show a significant proportion of respondents to valuation surveys on biodiversity and wildlife can hold rights based beliefs as their motivation for environmental concern (Stevens *et al.* 1991, Spash 1998a, 1998b, 2000).



### 3. CBA Methods for social appraisal of environmental change

This section turns to a discussion of the specific methods for monetary valuation of the environment which have been mentioned above. In doing so, the main problems associated with applying these methods to different aspects of the environment are raised. This shows how the idea that environmental damages and improvements might be reflected in monetary terms confronts both theoretical limitations and practical problems.

#### 3.1 Travel cost method

Travel cost method (TCM) is the oldest of the non-market valuation techniques predominantly used in outdoor recreation modelling. The basic method is to place a value on non-market environmental goods by using the costs of consumption behaviour in related markets. For example, to evaluate recreational fishing, a TCM survey would typically gather information on travel costs, access / fish license fees, on-site expenses, and capital expenditure on fishing equipment. Varying such costs and predicting fishing activity changes can then be used to derive surrogate demand functions for fishing at a specific location. TCM is normally applied to site-specific locations and the cost of travel tends to be road based, via cars. The basic problems are: whether to use zonal or individual visits, how to treat visits to other sites, deciding the treatment of costs, and statistical problems. On theoretical grounds neither zonal nor individual visits ranks above the other. Unfortunately consumers' surplus estimates for a given site or class of sites have been shown to vary substantially with the choice of measure. A visit to a specific site may be only part of the purpose for an individual's journey. This is problematic because the full travel costs cannot be attributed to the site in question. Mendelsohn *et al.* (1992) regard most TCM studies as having either ignored multiple destination trips or arbitrarily allocating trip costs across visited sites. Calculating the cost of distance travelled involves setting a price per mile. This requires either using petrol costs only, as an estimate of marginal cost, or allowing for all the costs of motoring by including an allowance for sunk costs such as depreciation and insurance. Individuals, in maximising utility, are assumed to compare the marginal utility with the marginal costs of consumption to achieve an economically efficient outcome. This process implies marginal costs should be used, since including all costs will result in a measure of price per mile using average costs. The choice will influence the consumers' surplus figures. Similarly, time is expended both in travelling to a site and whilst enjoying the site itself. As a scarce commodity, time clearly has an implicit (or shadow) price. If individuals are giving up working time, in order to visit a site, the wage rate is the correct opportunity cost; if a site visit occurs while on holiday leave, the opportunity cost will be measured with reference to the value, at the margin, of other recreation activities foregone. Ideally, a separate value would be calculated for each individual to reflect their set of leisure activities and valuations. In practice, such data is too difficult and expensive to collect. Debate on the appropriate approach continues.

The general conclusion would seem to be that TCM researchers are forced to assign their own subjective estimation of visit costs. Randall (1994) has argued that visit costs are inherently subjective, but give an ordinal measure if the cost increases with distance travelled. Thus, the traditional TCM yields an ordinal measure of welfare. However TCM cannot then serve as an independent technique for estimating recreation benefits. Randall (1994) has suggested the subjective treatment of costs means that TCM must be calibrated using information generated from fundamentally different methods so that TCM is no longer an independent tool. This might follow the example of Cameron (1992) who combined TCM with CVM, an approach successfully followed by Kling (1997).

As part of the process of calculating welfare changes via TCM a regression analysis will be undertaken to predict visits, e.g. visits per capita are a function of travel costs. Loomis (1995) has argued that regional economic effects associated with the improvement of a recreation site will be underestimated unless all aspects of the decisions to undertake recreation are included. He identifies four recreation aspects of choices: participation in a given recreation activity, the site(s) to visit, the frequency of trips to a given site, and the length of stay. He goes on to explain recent advances in statistical techniques for modelling each recreational choice, and illustrates (using deer hunting) how linking two of the four recreation choices yields more complete estimates of the change in number of trips, income and employment.

Another more general problem is the extent to which aggregating preferences reflects the type and range of values of concern. For example, the presence and size of human settlements in areas bordering a site or national park may play a decisive role in determining attributed monetary values. The closer an ecosystem is to large human settlements the more there are likely to be frequent visitors and hence a larger aggregate



monetary value may be calculated as being associated with the site. Those coming from further away to a remote site will have a higher willingness to pay which can counter this impact on total site value, but a green space in the city may easily prove to have a higher monetary value on the basis of low cost but frequent visits. In the extreme a wilderness area which restricted all access would be regarded as having no value under the TCM. Thus, contrary to a criterion of environmental prioritisation based upon the pristine or virgin status or biodiversity of an ecosystem, an altered and ecologically degraded site can appear socially preferable and of more value under the TCM. Using TCM in Europe to indicate which ecosystems should be protected as rare habitats may therefore lead to the loss of ecosystems in remote regions, regardless of their ecological quality or significance.

The extent to which TCM might help assess the environmental externalities can be summarised as follows. Recreation and tourism cover several activities such as hiking, painting, photography and bird watching. TCM was developed with the site specific recreational visitor in mind and is therefore well suited to assessing these values, but the qualifications mentioned above (e.g. non-site benefits) must be taken into account. The extent to which TCM can address other aspects of environmental value is often limited. Aesthetic changes (such as 'wilderness' aspects) are only likely to be discernible in TCMs where the site has special features creating a recreational demand, and even then they may be difficult to define or measure. Cultural and historical values would only be part of TCM estimate of demand where they are associated with site specific features, and require site visits for enjoyment or evaluation. Peace and quiet, as part of a site's recreational experience, might be included under TCM, but the impact on the wider community of noise pollution (e.g. from farming machinery or aggregates extraction) would be excluded. Health and safety concerns associated with certain rural land uses would fall beyond the scope of TCM to evaluate. In general the site characteristics valued by TCM are only those recognised by visitors as important. That is, the values are implicit in the preferences of the visitors. This means, if visitors fail to recognise the importance or even existence of a characteristic of a site (e.g. biodiversity) then this characteristic will be absent from the valuation via TCM. In particular, genetic diversity and ecosystem functions are unlikely to form part of site values obtained under TCM.

### **3.2 Production function approach**

The production function approach (PFA) generally uses scientific knowledge on cause-effect or dose-response relationships, i.e. the relationship between environmental quality variables and the output level of a marketed commodity. The PFA has been popular in studying air pollution impacts on agricultural crops (Spash 2001), but has also been applied elsewhere, for example pollution impacts on fisheries, e.g. Kahn (1991) and Silvander and Drake (1991). The PFA requires a quantifiable definition of the environmental change of concern, linking this change to a receptor response function, and then applying the results to an economic model for a related market good. Thus, applications have been determined by the availability of existing scientific information on dose-response functions.

Physical characterisation of environmental quality change requires the analysis of biological processes, technical possibilities, their interactions with producer decisions and the effect of resulting production changes on consumer and producer welfare. Biological or production response data provide a link between an environmental variable (e.g. water or air quality) and the performance parameters of an ecosystem. The response relationship may be quantified directly from biological experimentation, indirectly from observed producer output and behavioural data (secondary data) or from some combination of data sources. Procedures based upon producer data (e.g. production or cost functions) are preferable from the viewpoint of economic analysis (Adams 1983) and can avoid the need for explicit cause-effect functions. However data and statistical difficulties have restricted their applicability. Scientifically derived cause-effect functions have been most commonly applied in economic assessments. Cause-effect models offer a means of measuring the economic costs of several important environmental quality changes. However controversies over the appropriate way in which to model responses mean that widely varying estimates of economic damages can emerge. In addition, the model must be linked to data on the physical environment (e.g. water or air quality) in order to make accurate impact predictions. Such data needs to be locally or regionally disaggregated and is normally unavailable.

The choice of the economic model brings its own set of problems. Three main categories of economic approach can be defined: traditional models, optimisation models and econometric models. The traditional approach takes given market prices and multiplies these by losses in output. Thus the latter two approaches are to be preferred as they avoid this crude approximation of economic impacts in an oversimplified market structure. The main requirement for easy assessment of consumer and producer welfare is the existence of well established economic models. In their absence the analyst must construct a model and test its validity.





In the case of the environmental externalities from environmental changes the main areas in which a PFA could operate would be in looking at impacts on market products, materials damage and health effects related to labour markets. The PFA is unsuitable for estimating the benefits from wildlife conservation, or recreation and tourism unless there is an associated market good or service with which to link dose-response functions. For example, if wildlife conservation increased or reduced production costs (or increased or reduced marketed output), then a link could be feasible with the impacts on existing product supply, cost and/or profit functions. Cultural and historical values which might be lost due to land use practices could only be assessed via physical impacts on materials (e.g. the rate at which chemical applications might cause erosion) which would relate only to maintenance costs rather than the socio-economic aspects of culture.

Ecosystems functions are commonly ignored by consumers and producers and therefore difficult to value in economic terms. The PFA is the only method which seems appropriate here because of its basis in scientific knowledge which can then be linked into economic processes, although the values it will be able to assess would still be limited. A more general problem is the lack of suitable scientific information.

### **3.3 Hedonic pricing**

In this section reference is made to house prices as an example of how hedonic pricing (HP) could be applied to rural land use change, although similar models have been applied in labour markets where environmental risks are internalised in the wage rate. Aspects of the environment can be valued via their impact on the price of housing by being one of the characteristics which contribute to the definition of the commodity which a particular house and location represent. Thus, proximity to clean water and air, recreational opportunities, peace and quiet can all be expected to be factors increasing prices of housing in certain markets. Conversely, the existence of pollution or environmental nuisance may reduce the price of housing in the vicinity. The problems with implementing the HP approach include choosing the variables to include and the functional form.

The analyst must decide which factors to include as explanatory variables in the HP equation and the demand curve. Excluding a variable which has a significant effect on house prices, and which is correlated with some or all of the other variables in the model, will influence the estimation of coefficients. This leads to biased estimates for these coefficients and for the implicit prices (see Atkinson and Crocker 1992). Several of the independent variables included in the HP equation may be closely correlated with each other. This is the problem of multi-collinearity. The HP equation is non-linear so as to allow derivation of a demand equation. However there is no universally preferred functional form. Criteria which might be used to select the functional form include: restricting the number of parameters, selecting parameters justified by economic theory, choosing the form which economises on computing time, finding a form which provides a good explanation of the observed data and gives correct predictions, e.g. house prices falling with increased pollution (see Garrod and Allinson 1991). Housing markets are often segmented on grounds such as ethnic composition, rental versus owner-occupied, and price bracket. This can bias coefficients in the price function because segmentation implies that demand parameters vary across sectors. The HP analyst must then estimate separate price equations for each segment of the market. In a study of Boston, Michaels and Smith (1990) asked estate agents to segment the housing market and so identified four distinct sub-markets. Separate HP equations were then estimated for each segment in order to value the disutility of living close to hazardous waste sites.

HP assumes that current levels of environmental quality are the main influence on house prices, but they can also be influenced by expected changes in environmental quality. For example, the prospect of strict dust and noise regulations relating to quarries can keep prices higher in zones near such sites than in the absence of such expectations. The implicit price would fail to measure the valuation of current noise and air quality levels alone. Similarly, the expectation that a quarry will close in the near future, and the site will be restored, can raise prices where they had been depressed by the quarrying activities.

The HP approach is related to regional or site-specific characteristic but can be used to gain aggregate estimates of demand on the basis of site data. Where recreation and tourism are recognised components of the housing market HP could be used to value changes in their quality, subject to the qualifications above. Aesthetics, or aspects thereof, have been subject to benefit estimation using HP, although CVM (discussed below) is seen to have several advantages. The extent to which the aesthetic concept is seen to be captured will depend upon how it is defined, e.g. beauty versus water quality. Providing a disaggregation of aesthetic values seems unlikely due to the difficulty of finding and agreeing upon any measure of aesthetic variation. On



a more practical level, Graves (1991) points out that data is severely limited, not only due to a lack of measurement but also, because many important aesthetic features are located away from well developed markets. Cultural and historical values could also, in theory, be derived from HP where the characteristics associated with these values are identifiable, e.g. a particular geographical location or building. However the extent to which such values are capitalised into land or house prices is questionable; for example, house location can be unrelated to the enjoyment of cultural and historical sites. In general HP would be useless due to the inability to provide data measuring quantity or quality variations of these concepts by location. Health and safety, noise and air pollution may be assessed via HP. For example, Hughes and Sirmans (1992) have used a standard HP model to show a substantial negative effect of traffic externalities on single-family house prices. Similarly, house prices have been shown to be influenced by the positive externalities from urban forests which include benefits derived from pleasant landscape, clean air, peace and quiet and screening, as well as recreation (Tyrvaainen 1997). However HP studies do not generally disaggregate these estimates but rather tend to assess the total (positive or negative) externality.

The theoretical assumptions underlying HP mean it will give inaccurate estimates of environmental externalities if buyers lack perfect information about relevant environmental quality variables, buyers are unable to attain their utility maximising position, or the housing market is in disequilibrium. Furthermore, HP requires weak complementarity which means only those environmental externalities of rural land uses that have an impact on the property market will be measured. Finding suitable variables to measure environmental quality attributes can be problematic.

### **3.4 Stated preference methods**

Stated preference methods directly survey individuals to obtain their preferences rather than analysing their actual behaviour as revealed in the market place. This has led to some criticism from those economists who prefer to use secondary data collected by government agencies but which is related to actual behaviour. Research collecting primary data on intended behaviour is common in other social sciences and the conditions for convergence with actual behaviour have been studied extensively (see Fishbein and Ajzen 1975). In terms of the CVM there have also been tests comparing actual with stated willingness to pay. The apparent advantage attributed to methods such as TCM and HP is that they relate to actual behaviour, but as has been explained above the links being drawn between observational data and the underlying motives are often weak or purely speculative. This suggests the need for economists to pay far more attention to motives for behaviour, whether intended or actual.

CVM is the principal stated preference method although both conjoint analysis and choice experiments have received some recent attention. In contrast to the other three methods reviewed above, CVM has received considerable and increasing attention in the literature with academic journal articles on the subject in excess of 1000 studies world-wide. The main advantage attracting this attention is the ability of CVM to estimate what are termed option, 'existence' and bequest values in addition to direct use values. The combination of these indirect or passive use values can be large compared to the direct use values associated with non-market goods to which the other methods are solely restricted.

There are several stages to conducting a CVM study: survey design, pre-testing, carrying out the main survey, estimating willingness to pay (WTP) and/or willingness to accept (WTA), bid curve analysis, data aggregation, and final assessment. In particular application of the technique requires careful survey design, awareness of potential biases and a decision on whether to use WTP or WTA. The design of a CVM study includes the amount, type and way information is presented to individuals, the order in which it is presented and the question format. There is a wide body of evidence to suggest that survey design can affect responses. Survey design requires framing a realistic decision concerning the environment where the monetary question to be asked is accepted as a possible state of the world in which individual respondents might find themselves. Important decision by the analyst include a reason for the payment and how, including how often, funds will be raised (the bid vehicle). The technique for bid elicitation may be an open-ended question, a dichotomous choice, or a bidding game. Also, at this stage information on physical changes will be summarised and the method of their description chosen (e.g. text, graphics, maps).

Due to the sensitivity of responses to the information supplied the pre-testing of the survey has become of increasing importance. This can be conducted via a small sample test run of the survey or a focus group. The pre-test will enable the identification of problems with regard to the framing of the decision problem as well as divergence between encoding and decoding of information. The conduct of the main survey can use several variations. The in-house interview is now most favoured, although the expense of this approach often means



surveys are completed in the street, by telephone interviewing or mail. The sample is often weighted in terms of the local or regional population which is seen as politically more important to the decision and likely to have strong direct economic connections to the outcome.

WTP and WTA are the two welfare measures available for a CVM survey. Willig (1976) showed that the two measures would be close if the ratio of consumer's surplus to income (expenditure) was sufficiently small, and if the income elasticity of demand for the good in question was sufficiently low. Where these conditions failed to hold, precise limits on the difference between the two measures could be calculated. Whilst Bockstael and McConnell (1980) criticised the applicability of Willig's findings to environmental benefits, Randall and Stoll (1980) extended Willig's theorem (which was derived for price changes) to the quantity changes more commonly encountered in environmental valuation. Despite this stated WTP has still been found to be significantly lower than stated WTA (e.g. Rowe *et al.* 1980; Hammack and Brown 1974), and lower than the Willig approximation would suggest. In addition, experimental work by Knetsch and Sinden (1984) and Gregory (1986) has also found that WTA exceeds WTP. On practical grounds, the status quo reference position is preferable in terms of the property rights structure. If an alternative is imposed by the blanket imposition of WTP formats in all CVM surveys, the result can be to create an unrealistic trade-off, hypothetical market bias and protest bids. Thus, rather than follow a generic prescription to always use WTP formats as a conservative estimate of values (e.g. guidelines suggested by NOAA 1994), the property rights prevalent in a given situation should be used as guidance. This reinforces the theoretical argument for using WTA to measure a loss and WTP for a gain (Knetsch 1994).

Typically, median bids are less than mean bids so both are reported. At this stage the treatment of 'protest bids' becomes problematic and these are often omitted from the mean calculation. Protest bids are zero bids given for reasons other than a zero value being placed on the resource in question. For example, a respondent may refuse any amount of compensation for loss of an environmental asset which they regard as unique or a species which they feel should be protected at all costs. Respondents may refuse to state a WTP/WTA amount because they reject the survey as an institutional approach to the problem, or because they have an ethical objection to the trade-off being requested, e.g. a lexicographic preference. Analysis of the bid curve is used to test construct validity, i.e. that the socio-economic variables have the expected signs, and the regression is statistically significant. Other relationships can also be investigated at this stage. In general bid curve analysis has tended to be of academic, as opposed to policy, interest despite the relevance for judging whether the exercise has produced the expected results in accordance with economic theory.

The method of aggregating data, both across time and space, requires deciding on the relevant population, the method of aggregating from the sample bid, and the time period or discounting procedure for aggregation. These are major concerns in CBA and have serious impacts on any resulting monetary values. The sensitivity of the results to variations in such factors should be tested and presented as a central aspect of the findings, however this is rarely the case. Sensitivity analysis in CBA is generally ignored or extremely limited.

Final reflection upon the CVM study can include convergent validity and success of repeatability where there exist other similar studies. The overall success of the exercise will also become apparent as the results are being analysed, e.g. a high number of protest bids. There are several specific problems which are recognised as possible causes of bias, some of which have been mentioned: strategic bias, design bias (choice of bid vehicle, prompting a bid). More problematic are the impact of information, as this is by necessity restricted but can have serious influence upon the resulting bids, and the problem of embedding as raised by Kahneman and Knetsch (1992). These two issues are discussed next.

In a hypothetical market, respondents combine information provided to them regarding the good to be valued, and how the market will work, with information they already hold on that good. Their responses may be influenced by either hypothetical market or commodity-specific information given to them in the survey. This phenomenon implies that WTP/WTA values are endogenous to the valuation process. Samples *et al.* (1986) found bids to preserve different animal species varied significantly according to the information provided by researchers. Ajzen *et al.* (1996) concluded from experimental research that the nature of the information provided in CVM surveys can profoundly affect WTP estimates, and that subtle contextual cues can seriously bias these estimates under conditions where the good is of low personal relevance. Whitehead and Blomquist (1991) show both theoretically and empirically how information on environmental substitutes changes the value of related goods. In their research, telling respondents about alternative wetland sites significantly altered WTP to protect the Clear Creek wetland in Kentucky. However Randall (1986) has argued that CVM answers should vary under different information sets, otherwise the technique would be insensitive to significant changes in commodity framing. The divergence of opinion here relates to information provided which economists regard as important to the decision and should therefore have an impact but fails to do so, and information which



would be regarded as peripheral or irrelevant by economists but which does have an impact on stated behaviour. Indeed, the effects of information may be inappropriately labelled as bias, depending on the way in which WTP/WTA is changed. Information which improves the knowledge of an individual concerning the characteristics of a good can be regarded as informing a consumption decision. Information which alters the preferences is more problematic in the neo-classical framework and could be regarded as creating a bias. For example, Baron and Maxwell (1996) show that individuals' WTP can be biased by information on the cost of provision of public goods and suggest eliminating information from which costs could be inferred, so that respondents can focus more easily on benefits alone. While such redesign may avoid some types of bias, a more general issue, which remains, is how far individual preferences can be regarded as exogenous to the valuation process and especially so when goods are unfamiliar and/or never traded in a market.

This problem arises when the component parts of an individual's valuation are evaluated separately and when summed found to exceed the valuation placed upon the whole. Evidence that such behaviour exists was provided by Seip and Strand (1990). CVM studies have found part-whole bias, also termed embedding, and this has been attributed by some to valuation of the moral satisfaction from contributing to a worthy cause ('warm glow') rather than the good itself (Kahneman and Knetsch 1992). The counter reaction has been that CVM surveys finding embedding are flawed in some way which creates the part-whole bias, and that this can be corrected by careful survey design (Carson and Mitchell 1993, 1995; Hanemann 1994). However Bateman *et al.* (1997) have provided experimental evidence for the existence of part-whole bias for private goods outside of the CVM context. They therefore suggest the problem lies with economic preference theory rather than the CVM approach.

CVM can in theory, although to varying degrees, address most of the non-market value categories related to environmental change, with the exception of biodiversity and ecosystems functions. Biodiversity itself has rarely been valued in CVM, instead individual species (rather than their diversity) have been a focus of research. Aesthetics have been subject to benefit estimation by HP and CVM, with Graves (1991 p. 225) regarding CVM as having several advantages, i.e. distinguishing the aesthetic dimension of a policy change; generating data rather than relying on remote proxies; imposing fewer behavioural assumptions; and yielding plausible results, particularly in applications to visibility. However defining aesthetic qualities and conveying them to individuals so they become familiar with the concepts in a commodity framework can be difficult or impossible. Also, the description of an aesthetic quality may itself introduce distortions.

One practical issue in terms of applying CVM concerns the disaggregation of benefit categories. Whilst the CVM survey can, in theory and as mentioned, be applied to a range of value categories any one survey will be limited in scope and is normally restricted to an aggregate assessment of total benefits, as with the other methods. The most common disaggregation is by direct use and passive use (i.e. two categories) and in some cases by use, option, existence and bequest values (i.e. four categories). This kind of disaggregation could be carried out for classes of an externality by one of three main methods. First, obtain a total bid and then ask respondents to split the total, e.g. state the percentage attributable to each category. Second, ask for an evaluation of each externality either by the same individual or, if a large enough sample is available, by different individuals (i.e. sub-samples for each externality). Both these approaches may be difficult for individuals to comprehend. Third, develop alternative scenarios so that aspects of the externality are removed (e.g. noise/no noise) for different sub-samples. The problem here is to develop realistic scenarios responsible for creating the change in the externality. The issue of disaggregating benefits by externality and the preceding discussion highlights the continuing experimental nature of the CVM approach.

In areas where an environmental change can be easily described and understood in terms of a choice based upon individual preferences, and the market trade-off implied is accepted as appropriate to the decision then CVM seems applicable. However where aesthetics, cultural & historical values and ecosystems functions & biodiversity are concerned these conditions often seem unlikely to hold. Where only certain aspects of such environmental concepts are easily explained and/or captured in commodity terms the resulting monetary valuation will be a poor reflection of the environmental values they encapsulate.



#### 4. Conclusions

Table 1 provides a summary of the methods discussed so that the economic valuation techniques suitable for use in the context of several categories of environmental change can be easily identified. There is some danger in presenting such a simplified table and the qualifications and limitations of methods must be kept in mind. In addition there are general qualifications to the use of CBA. For example, income distribution is taken as given so that prices and monetary estimates will reflect relative purchasing power in society. Adjustments could be made to the results to test for the impact of changing income distribution, but, as with other sensitivity analysis, this is rarely done in practice.

HP can assess certain aspects of externalities after they have occurred and have been capitalised. The PFA is generally inapplicable due to a lack of scientific data. TCM is primarily concerned with recreation and tourism values at a site prior to any development. As can be seen CVM provides the most potential for comprehensive coverage of externalities and could be conducted before or after an environmental change. In terms of including option, existence and bequest values only CVM can attempt to do so. CVM also provides considerable flexibility in the types of non-market value which can be addressed in the survey. However there are several aspects of implementing CVM which restrict the extent to which it will be able to assess environmental values. In practical terms, the cost of and time needed for conducting a CVM survey can be relatively high and have increased due to the extent to which various design features are now regarded as required practice. However, CVM remains an experimental technique which has been accelerated into public policy use by legal action in the USA over natural resource damages. Perhaps the greatest contribution the technique is now making is in terms of forcing economists to reconsider the content and meaning of both observed and intended human behaviour with regard to a plurality of environmental values.

The extent to which societal well-being is related to the monetary estimates obtained by the methods outlined is highly context dependent. Thus, where there are large scale changes across space and time the micro-economic assumptions underlying welfare measures are unlikely to hold. Similarly, where there are diverse cultures or differences in agreement over the role of markets there will be rejection of the value basis of the measures. Economists have tended to overlook such issues but in applied policy arenas this merely means economic prescription become unrealistic and academically abstract. If an ethical concern or income distribution issue is key to the valuation of an environmental change then a more pluralistic approach is required than offered by standard monetary valuation methods. This means explicitly considering the role of monetary estimates in an openly discussed decision process rather than assuming CBA is merely one input into an undisclosed process of politics.

**Table 1: Current suitability of CBA methods for assessing the economic externalities of environmental change**

Non-Market Value Category	Method			
	TCM	PFA	HP	CVM
<i>Aesthetics</i>	?	-	?	?
<i>Cultural &amp; Historical Values</i>	?	-	-	?
<i>Health &amp; Safety</i>	-	+	?	+
<i>Ecosystems Functions &amp; Biodiversity</i>	-	?	-	?
<i>Peace &amp; Quiet</i>	?	-	+	+
<i>Recreation &amp; Tourism</i>	+	-	?	+
<i>Water &amp; Air Quality</i>	?	+	+	+

Key:

- Of little or no use
- ? Sometimes useful
- + Potential for application strongest



## References

- Adamowicz, W., J. Louviere and M. Williams (1994), 'Combining revealed and stated preference methods for valuing environmental amenities', *Journal of Environmental Economics and Management*, 26, 271—292.
- Adamowicz, W., P. Boxall, M. Williams and J. Louviere (1998), 'Stated preference approaches for measuring passive use values: choice experiments and contingent valuation', *American Journal of Agricultural Economics*, 80, 64—75.
- Adams, R. M. (1983), 'Issues in assessing the economic benefits of ambient ozone control: Some examples from agriculture', *Environment International*, 9, 539—548.
- Ajzen, I., T. C. Brown and L. H. Rosenthal (1996), 'Information bias in contingent valuation: effects of personal relevance, quality of information and motivational orientation', *Journal of Environmental Economics and Management*, 30 (1), 43—57.
- Atkinson, S. and T. Crocker (1992), 'Econometric health production functions: Relative bias from omitted variables and measurement error', *Journal of Environmental Economics and Management*, 22 (1), 12—24.
- Baron, J. and N. P. Maxwell (1996), 'Cost of public-goods affects willingness-to-pay for them', *Journal of Behavioural Decision Making*, 9 (3), 173—183.
- Bateman, I., A. Munro, B. Rhodes, C. Starmer and R. Sugden (1997), 'Does part-whole bias exist? An experimental investigation', *Economic Journal*, 107 (441), 322—332.
- Beaufoy, G. (1994), 'Impact of the CAP Reform on land use and rural amenities', in A. Dubgaard, I. Bateman and M. Merlo (eds), *Economic Valuation of Benefits from Countryside Stewardship*, Kiel: Wissenschaftsverlag Vauk, pp. 27—46.
- Bockstael, N. E. and K. McConnell (1980), 'Calculating equivalent and compensating variation for natural resource facilities', *Land Economics*, 56, 56—62.
- Bonnieux, F. and P. Le Goffe (1997), 'Valuing the benefits of landscape restoration: a case study of the Contentin in Lower-Normandy, France', *Journal of Environmental Management*, 50, 321—333.
- Bromley, D. and I. Hodge (1990), 'Private property and presumptive policy entitlements: reconsidering the premises of rural policy', *European Review of Agricultural Economics*, 17 (2), 197—214.
- Cameron, T. A. (1992), 'Combining contingent valuation and travel cost data for the valuation of nonmarket goods', *Land Economics*, 68 (3), 302—317.
- Carson, R. T. and R. C. Mitchell (1993), 'The issue of scoping in contingent valuation studies', *Journal of Agricultural Economics*, December, 1265—1267.
- Carson, R. T. and R. C. Mitchell (1995), 'Sequencing and nesting in contingent valuation surveys', *Journal of Environmental Economics and Management*, 28, 155—173.
- Clark, C. W. (1976), *Mathematical Bioeconomics*, New York: Wiley.
- Fishbein, M. and I. Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Reading, Massachusetts: Addison-Wesley.
- Freeman, A. M. (1986), 'The ethical basis of the economic view of the environment', in D. van der Veer and C. Pierce (eds) *People, Penguin and Plastic Trees: Basic Issues in Environmental Ethics*, Belmont, CA: Wadsworth Publishing Co., pp. 218—227.
- Freeman, A. M. (1993), *The Measurement of Environmental and Resource Values: Theory and Methods*, Washington, DC: Resources for the Future.
- Garrod, G. and P. Allinson (1991), 'The Choice of Functional Form for Hedonic House Price Functions', Discussion paper 23, Countryside Change Initiative, University of Newcastle-upon-Tyne.
- Graves, P. E. (1991), 'Aesthetics', in J. B. Braden and C. D. Kolstad, *Measuring the Demand for Environmental Quality*, Amsterdam: North-Holland, pp. 213—226.
- Gregory, R. (1986), 'Interpreting measures of economic loss', *Journal of Environmental Economics and Management*, 13, 325—337.
- Hammack, J. and G. Brown (1974), *Waterfowl and Wetlands: Towards Bioeconomic Analysis*, Baltimore: Johns Hopkins Press.



- Hanemann, W. M. (1994), 'Valuing the environment through contingent valuation', *Journal of Economic Perspectives*, 8 (4), 19—43.
- Hanley, N and C. L. Spash (1993), *Cost-Benefit Analysis and the Environment*, Aldershot: Edward Elgar.
- Hartwick, J. M. and N. D. Olewiler (1986), *The Economics of Natural Resource Use*, New York: Harper and Row.
- Hughes, W. T. and C. F. Sirmans (1992), 'Traffic externalities and single-family house prices', *Journal of Regional Science*, 32 (4), 487—500.
- Kahn, J. (1991), 'Atrazine pollution and Chesapeake fisheries', in N. Hanley (ed.) *Farming and the Countryside: An Economic Analysis of External Costs and Benefits*, Oxford: CAB International, pp. 137—158.
- Kahneman, D. and J. Knetsch (1992), 'The purchase of moral satisfaction', *Journal of Environmental Economics and Management*, 22 (1), 57—70.
- Kling, C. L. (1997), 'The gains from combining travel cost and contingent valuation data to value non-market goods', *Land Economics*, 73 (3), 428—439.
- Knetsch, J. (1994), 'Environmental valuation: some problems of wrong questions and misleading answers', *Environmental Values*, 3, 351—368.
- Knetsch, J. and J. Sinden (1984), 'Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity', *Quarterly Journal of Economics*, 94 (3), 507—521.
- Loomis, J. B. (1995), 'Models for determining environmental quality effects on recreational demand and regional economics', *Ecological Economics*, 12 (1), 55—65.
- Louviere, J. (1996), 'Relating stated preference measures and models to choices in real markets: calibration of CV responses', in D. J. Bjornstad and J. R. Kahn (eds) *The Contingent Valuation of Environmental Resources: Methodological Issues and Research Needs*, Cheltenham: Edward Elgar, pp. 167—188.
- Mackenzie, J. (1992), 'Evaluating recreation trip attributes and travel time via conjoint-analysis', *Journal of Leisure Research*, 24 (2), 171—184.
- Mendelsohn, R., J. Hof, G. Peterson and R. Johnson (1992), 'Measuring recreation values with multiple destination trips', *American Journal of Agricultural Economics*, 74 (4), 926—933.
- Michaels, R. and V. Smith (1990), 'Market segmentation and valuing amenities with hedonic models: the case of hazardous waste sites', *Journal of Urban Economics*, 28, 232—242.
- MLURI (1996), *Valuation of the Conservation Benefits of Environmentally Sensitive Areas*, Aberdeen: Macaulay Land Use Research Institute, 78pp.
- NOAA (1994), 'Natural resource damage assessment: proposed rules', *Federal Register*, 59 (5), 1062—1191.
- Pruckner, G. (1995), 'Agricultural landscape cultivation in Austria: an application of the CVM', *European Review of Agricultural Economics*, 22 (2): 173—190.
- Randall, A. (1986), 'The possibility of satisfactory benefit estimation with contingent markets', in R. Cummings, D. Brookshire and W. Schulze (eds) *Valuing Public Goods: An Assessment of the Contingent Valuation Method*, Totowa, NJ: Rowan and Allanheld, pp. 114—122.
- Randall, A. (1994), 'A difficulty with the Travel Cost Method', *Land Economics*, 70 (1), 88—96.
- Randall, A. and J. Stoll (1980), 'Consumers surplus in commodity space', *American Economic Review*, 70 (3), 449—455.
- Roe, B., K. J. Boyle and M. F. Teisl (1996), 'Using conjoint analysis to derive estimates of compensating variation', *Journal of Environmental Economics and Management*, 31 (2), 145—159.
- Rowe, R., R. d'Arge and D. Brookshire (1980), 'An experiment on the economic value of visibility', *Journal of Environmental Economics and Management*, 7, 1—19.
- Samples, K., J. Dixon and M. Gowen (1986), 'Information disclosure and endangered species valuation', *Land Economics* 62 (3), 306—312.



- Seip, K. and J. Strand (1990), Willingness to Pay for Environmental Goods in Norway: A contingent valuation study with real payment, Memorandum 12, Dept. of Economics, University of Oslo.
- Silvander, U. and L. Drake (1991), 'Nitrate pollution and fisheries protection in Sweden', in N. Hanley (ed.) *Farming and the Countryside: An economic analysis of external costs and benefits*, Oxford: CAB International, pp. 159—176.
- Spash, C. L. (1998a), *Environmental Values and Wetland Ecosystems: CVM, Ethics and Attitudes*. Cambridge, Cambridge Research for the Environment, Department of Land Economy, University of Cambridge, 111pp.
- Spash, C. L. (1998b), 'Investigating individual motives for environmental action: Lexicographic preferences, beliefs and attitudes', in *Ecological Sustainability and Integrity: Concepts and Approaches 13*, edited by J. Lemons, L. Westra and R. Goodland, pp. 46—62 Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Spash, C. L. (2000), 'Multiple value expression in contingent valuation: Economics and ethics', *Environmental Science & Technology* 34 no.8: 1433-1438.
- Spash, C. L. (2001) , in *Managing Pollution: Economic Valuation and Environmental Toxicology* , edited by C. L. Spash and S. McNally, pp. xii + 222 Cheltenham: Edward Elgar.
- Spash, C. L. and N. Hanley (1995), 'Preferences, information and biodiversity preservation', *Ecological Economics* 12, 191—208.
- Stevens, T. H., J. Echeverria, R. J. Glass, T. Hager and T. A. More (1991), 'Measuring the existence value of wildlife: what do CVM estimates really show?', *Land Economics*, 67 (4), 390—400.
- Twigger-Ross, C. L. and D. L. Uzzell (1996), 'Place and Identity Processes', *Journal of Environmental Psychology*, 16 (3), 205—220.
- Tyrvaainen, L. (1997), 'The amenity value of the urban forest: An application of the hedonic pricing method', *Landscape and Urban Planning*, 37 (3—4), 211—222
- Whitehead, J. and G. Blomquist (1991), 'Measuring contingent values for wetlands: Effects of information about related environmental goods', *Water Resources Research*, 27, 2523—2531.
- Willig, R. (1976), 'Consumers surplus without apology', *American Economic Review*, 66, 589—597.





**Jürgen Freimann/ Michael Walther**

## **EVALUATING THE IMPACTS OF CORPORATE ENVIRONMENTAL MANAGEMENT SYSTEMS. A COMPARISON BETWEEN EMAS AND ISO 14.001**

### **Abstract:**

The authors present and discuss the results of empirical research on the implementation of standardized environmental management systems in companies. As there are basically two competing standards – the European EMAS and the worldwide ISO 14.001 – the focus of the paper is on potential different and common features of the corporate implementations, experiences and impacts.

Although the methodical approaches of the empirical research projects differ to a great extent, their results are quite similar: The experiences of the companies with environmental management systems do not really differ with respect to the formal standards underlying but to the corporate culture and the strategic importance environmental orientation has for the company. There are some more or less important differences in the formal demands of the two standards (for example the duty to publish an environmental report within EMAS but not ISO 14.001), but these differences do not really lead to different corporate practices. The widespread opinion- especially in Austria and Germany - that EMAS leads to a 'star performance' in environmental management cannot be confirmed by empirical findings.

In addition to these results this paper points out the occurring problems of empirical research in the field of environmental management. Many of the given statements by the corporate actors cannot be verified by the researchers and therefore have to be handled carefully. This leads to the proposition that companies have to evaluate their grade of sustainability themselves and should publish the results to make it possible for the stakeholders to compare them. The main task for the scientific community then would be to develop the necessary indicators, benchmarks, standards etc. to help the companies.

### **Two competing EMS-standards**

Within the last years many companies all over the world installed standardized environmental management systems (EMS) in order to systematically manage the environmental effects of the companies' activities. There are basically two relevant systems, EMAS (Eco-management and Audit Scheme, the EU's standard, a state run system) and ISO 14.001 (the worldwide EMS-standard, privately run by the International Organization for Standardization). Several empirical studies tried to find out how the implementation of one of those systems influences environmental performance, costs and benefits of companies.

Since most of this research took place in German speaking countries, their results have not yet been published in English. In this paper the authors try to give a short overview on the results and a comparison between the EMAS and ISO 14.001 experiences. Both, EMAS and ISO 14.001, are formalised environmental management system standards that compete to gain favour of companies and organisations– at least in the European Union. Since EMAS was established by the EU and is furthermore provided with the force of law it is considered to be more significant. ISO 14.001 is just an industrial standard but as part of the 'ISO management standard family' it is very well known by potential users.

Participation is voluntary and both systems do not substitute national environmental law. The development especially of EMAS was motivated by the idea, that the pressure of competition would encourage a large number of companies to participate, even if their previous voluntary environmental care was rather small.

Both standards represent a new approach to corporate environmental care. They do not focus on results of actions in form of material objectives and limits, but on setting up organisational structures and managerial processes and subject these to continual control. To comply with these requirements corporate actors are no longer allowed to just obey the rules. They rather have to develop and make use of different abilities like formulating environmental objectives and implementing organisational structures and processes as well as measurement of achievements.



This engagement is rewarded with a certificate that shows the corporate environmental commitment to the public. In the case of EMAS the certificate is combined with the right to use the EMAS-logo in corporate advertising (but not in product advertising).

In the core of the EMS, the structure of the management system, ISO 14.001 and EMAS were quite similar at the time the studies were carried out. Both systems include specific organisational measures to avoid pollution and damages of the environment and intend to improve the environmental performance of a company. They force participating companies not only to implement particular measures (like environmental policy, aims and program) but also to carry out periodical internal checks and external audits by independent auditors.

Looking at both standards in detail there are some differences<sup>32</sup>. EMAS is above all a supranational instrument of environmental policy, which aims to stimulate the rather deadlocked and Europe-wide differently developed environmental legislation by turning away from the former 'command and control'-concept with its limited effects in comprehensive environmental measures.

While ISO 14.001 is primarily an internal management tool used by companies to improve their systematics, legal security and innovative ability, EMAS sets a second focus on the external stakeholders. Companies participating in EMAS have to publish an environmental statement to inform the public. ISO 14.001 does not include this duty.

The recent amendment of EMAS (EMAS II) defuses the rivalry between the two standards. ISO 14.001 is now the management system required by EMAS. A company, which has already installed an EMS based on ISO 14.001, has to accomplish some additional duties only (for example the proof of complete legal compliance (with national environmental laws) and the publication of the environmental statement) to get validated as an EMAS-participant as well.

Corporate self-control and the supervision by public and experts required by EMAS regulations are often considered to be a substitute for state control. Therefore, EMAS has frequently been misinterpreted as an instrument of deregulation of environmental laws. This is the main reason why several companies expected state support in return. In the case of ISO 14.001 as a private norm these wishes could not emerge, and consequently participation in this system does not require any public statement.

Though, both standards are open in content. The users themselves and not the standards set the goals they aim to achieve with the aid of the established EMS. Especially in Germany it has been argued, that EMAS is overall of a higher valence than the ISO 14.001 standard. However, this proposition can only be evaluated through the empirical study of practical experiences and not from the wording of the standards themselves.

Meanwhile numerous empirical studies have been published on EMS. The majority refers to the EU as well as – more often - to German or other national practices. They are mainly based on the implementation of the first version of EMAS, sometimes in combination with ISO 14.001. The results presented in this article are essentially based on our own studies<sup>33</sup>. We compared our results to a wide range of other empirical studies based on questionnaires and case studies.<sup>34</sup> The results of these studies are to a great extent consistent.

Switzerland is the only country in Western Europe, where, as a non-member of the EU, companies can only implement environmental management systems according to ISO 14.001. With most of the German studies concentrating on EMAS and the additional problem that a lot of eco-orientated companies in Germany participate in both systems, Switzerland is the country where we can gain pure experiences with ISO 14.001. A recently published study by Thomas Dyllick and Jost Hamschmidt<sup>35</sup> from St. Gallen/ Switzerland offers the opportunity to compare results and to clarify questions of the value of both standards. Another look at ISO 14.001 in practice in the form of collected worldwide case studies provides Ruth Hillary.<sup>36</sup>

---

<sup>32</sup> For a detailed comparison of all important aspects of both systems see BMU/UBA 2000, p. 28-34

<sup>33</sup> esp. UNI/ ASU 1997 and FBU 1995

<sup>34</sup> Mainly the following articles and reports: Schnauber/ Kiesgen/ Mangelmann 1994, Gerling Consulting-Gruppe 1995, Umweltakademie Fresenius/ IHK Dortmund 1996, Baum et al. 1996, Klemisch 1997, Wietschel/ Rentz 1997, Isaak/ Keck 1997, FEU 1998a/ b/ c, Hartmann 1998, Hoepfner et al. 1998, Jaeger et al. 1998, Seidel/ Weber 1998, BMU/UBA 2000.

<sup>35</sup> see Dyllick/ Hamschmidt 2000, for a short overview in English see Hamschmidt 2000

<sup>36</sup> see Hillary 2000



## The Spread of the Systems

It is not surprising that the ISO 14.001 standard attracted more participants, even though it came into force a year later than EMAS. After all it is valid worldwide. More significant than the absolute numbers are the relative changes. In the last two years the number of ISO 14.001-participants rose from 13.440 (December 1999) to 31.793 (October 2001), a growth ratio of about 135%. In the same time EMAS-participation increased only by 25%, from 3.155 to 3.995 with even some falling numbers within this time span (in the first half of 2000 participation dropped by 4%).<sup>37</sup>

Although these numbers show nearly eight times more ISO 14.001-participants than EMAS-participants a valid statement about the attractiveness of each standard is not possible, because of the different validity of the standards (world-wide vs. EU-only). Such a comparison is only feasible in the EU, where companies can choose between the two standards.

Analysing the data on participation in the EU the preference for ISO 14001 becomes evident. In every EU-country, except for Germany and Austria, ISO 14.001 is preferred to EMAS. More than 70%, about 2.650 sites, of the EMAS-participants are located in Germany (either sites of German companies or international companies situated in Germany). It is difficult to explain why so many German sites take part in EMAS. One reason can certainly be seen in the government-financed programs for participating companies that were offered especially in the early stages of the European EMAS-implementation-process. On the other hand, German companies are confronted with very strict environmental laws and therefore have to do significantly more in order to get validated than companies in other European countries. This is the same for Austria.

In relation to the different sizes of the EU member states we find the largest share of EMAS sites in Austria followed by Germany and the Scandinavian countries. Apart from Germany and Austria the ratio of ISO 14.001-participants compared to EMAS is 5/1 or even higher.

The Dyllick/ Hamschmidt study offers another interesting detail: They investigated the highly concentrated market for ISO 14.001-certifications in Switzerland. The three major certification-companies have an accumulated market share of 94%.<sup>38</sup> To our knowledge, comparable data of the corresponding market structures in Germany are not available, at least not publicly. It would be interesting to have more information on these conditions, because they are likely to influence the further development of validations and certifications respectively as well as the costs.

Another insight into instructive background information on the global spread of EMS-standards is provided by the BMU/UBA study. It deals with the acceptance of EMS in different countries and cultures and shows the great influence of different national eco-political constellations. Moreover, the importance of the varying cultural conditions of companies, law and economy becomes evident in this study.<sup>39</sup> Besides all superficial similarities of economic reality in different countries, which can mainly be put down to globalisation, the globalised economy and accordingly the environmental management remains rooted in national cultures. This explains at least partially the great differences in the propagation of EMS-standards.

In prospect the competition between EMAS und ISO 14.001 seems to be decided: the worldwide spread industrial standard has prevailed. It is questionable, whether the EMAS II amendment will be successful in winning back lost ground. Especially since EMAS II even converged closer to ISO 14.001 instead of increasing the demands and to position itself as the star-performance it already claims to be.

There is a wide range of possible explanations for this actual situation that this paper does not aim to discuss. Instead the focus lies on the impacts of each system. Only if EMAS really is the 'star performance' compared to ISO 14.001, the prevailing of ISO 14.001 is problematic from an ecological perspective. This question seems less relevant considering that the total number of eligible companies who implemented any EMS is less than 1%. On the other hand it can be stated for Germany that EMS affect over 50% of all employees, because the bigger companies participate to a greater extent than the small or medium sized companies.

---

<sup>37</sup> For up to date figures see [www.ecology.or.jp/isoworld/english/analy14k.htm](http://www.ecology.or.jp/isoworld/english/analy14k.htm)

<sup>38</sup> Dyllick/ Hamschmidt 2000 p. 30

<sup>39</sup> see BMU/UBA 2000, p. 103-109



### ***Experiences and impacts***

To begin with the punch line: The available studies do not prove that there are any significant differences in the effects of EMS depending on the different underlying standards.

Apart from details the observed differences of impacts between companies are not to be put down to the applied formal EMS-standard. They rather result from corporate political intentions which led to the implementation of the standard as well as from corporate cultural conditions the standards are confronted with and more or less fit in. This stringently results from the substantial findings of the surveys on hand.

### ***Ecological Impacts***

Environmental management systems get implemented to improve the corporate environmental care by detecting and removing ecological weak spots, either as a self-imposed aim or influenced by stakeholders. According to surveys on the motivation of companies this is stated to be the main reason for establishing EMS.<sup>40</sup> Therefore, the ecological impact is central to the question whether there might be a difference in valence of various EMS standards.

At first it is expected that the self-set eco-political aims in the context of EMS should significantly improve the ecological performance of companies with formalised EMS compared to those without. If EMAS really is of a higher value due to sanctions owing to the state-controlled validation, explicit orientation towards the public and the demand for documented ecological activities beyond compliance, EMAS companies should furthermore effect noticeable better results than companies certified only according to ISO 14001.

However, this conclusion cannot be drawn from any of the results of the empirical studies. Neither is it provable that companies with an implemented formalised standard show an ecological performance that is superior to those without any formalised EMS, nor do the studies point out that enterprises following EMAS aim for higher goals than those following ISO 14.001 or another EMS-standard.<sup>41</sup>

Early studies gave evidence that the established structures like formal responsibilities, working and procedural instructions, periodical audits and detailed environmental manuals - resulted in strict formalism and clear documentation rather than in adaptable and innovative organisational structures and processes.<sup>42</sup> This general statement was not only gained from empirical studies, but was also characterised as an inevitable symptom of EMAS.<sup>43</sup>

Conclusions drawn from subsequent empirical studies tend to confirm these early judgements. Environmental management systems are primarily expert systems. Participation of the employees is considered to be very important by most of the respondents but in practise this seems to be an exception. Participation of the works council is only realised in a small minority of companies. The established structures predominantly show only small differences to those requested by the legal regulations. In consequence EMAS does not improve the corporate environmental care to a higher extent than obligatory liabilities already did.

It is the formal structure of EMAS itself and not only the special German method of implementation that enforces this phenomenon.<sup>44</sup> EMAS implementation requires systematic checks of all relevant corporate activities and a complete documentation of all formal measures installed. Compliance to these requirements is to be proved to the accredited verifier and - by being forced to publish an environmental statement - even partly to the public. At the same time a process of continuous improvement of corporate environmental care is to be established. According to the findings of modern organisational theory this is only possible by loosening strict regulations and structures and enforcing organisational learning and development.

EMS in general as well as EMAS compared to ISO 14.001 have not yet proven their capability to lead companies to a better ecological performance, let alone to sustainability. With the use of new and therefore often cleaner technology, even companies with no or only small ecological ambitions realise cost savings by reducing the input of resources and the output of waste. Accordingly, cynics predicted that neither the

---

<sup>40</sup> see e.g. UNI/ ASU 1997, p. 48 and Seidel/ Weber 1998, p. 24

<sup>41</sup> see BMU/UBA 2000 p. 34-36

<sup>42</sup> see e.g. Freimann/ Schwaderlapp 1996

<sup>43</sup> see Freimann 1997

<sup>44</sup> see esp. Issak/ Keck 1997, Hartmann 1998 and FEU 1998a



ecological pioneers nor other well-managed companies, but those with the biggest (ecological) backlog would gain the most from participating.

In fact the Swiss study on ISO 14.001 confirms that the majority of participating companies (73%) had no experiences on environmental management before entering this standard.<sup>45</sup> Contrary to the popular self-assessment of participating companies as eco-pioneers, the standards seem to attract primarily those that are not above average on this field.<sup>46</sup>

Taking a close look at the environmental goals companies have to set by themselves according to both standards and the corresponding level of achievement one gets a first impression on the ecological impact of EMS. Here it becomes evident, that - at least in the beginning of the work with EMS- operational goals dominate, namely those that seem shortly accessible with a sufficient certainty.

The corporate environmental policy of companies participating in EMAS is dominated by technical measures. The focus is still on additive technologies (end of pipe). A change to technologies that avoid environmental pollution instead of treating it afterwards has yet to happen. At the moment investments in integrated environmental technologies are not common.

Nevertheless, the Swiss study confirms a trend that already emerged in the earlier German studies: organisational activities as well as the introduction of product-orientated ecology gain ground in the companies' programmes. Perspectively we can expect the step into a comprehensive ecological modernisation, especially in companies that already have more experiences in EMS by now.

These empirical findings contradict the criticism on EMAS, that it had a weak point in the field of products because EMAS I – in contrast to ISO 14.001 – did not require dealing with product-ecology.<sup>47</sup> The finding that the implementation of an EMS increases the sensibility of companies for ecological problems in regard to the use and disposal of their products applies to EMAS and ISO 14.001 accordingly.<sup>48</sup>

A main difference between the two systems is the request to evaluate the impact of all measures in terms of ecological effects in order to achieve effectiveness and not only efficiency. This is included in the EMAS-standard but not in ISO 14.001.<sup>49</sup> In practice companies lack the competence to do so and regional or national environmental goals are often missing. Therefore a valuation of the relevance of corporate ecological goals and measures does scarcely take place.<sup>50</sup>

Altogether the answer to the question on the ecological impact of standardized environmental management systems has apparently to be adjourned. It cannot be answered until there are long-term experiences with EMS as well as generally accepted indicators.

### ***Economic Impacts***

The economic impacts can be determined from the comparison of implementation costs and financial benefits. EMAS participation (as well as ISO 14001) is connected with considerable costs. Thereby we do not mean costs for technical measures of environmental care such as water treatment or waste gas cleaning. These costs should rather be called EMS consequential expenses. EMS-costs are costs that are directly connected with the implementation of the management system from the decision to participate up to the validation/certification.

Regarding the costs we find that in spite of printing and publishing the environmental statements, the validation by the state-accredited verifier and the registration participating in EMAS seems to lead to lower costs than participating in ISO 14.001. The Dyllick/ Hamschmidt study reports average costs of 287.000 Sfr (between 100.000 in the case of small and 535.000 Sfr for bigger companies), while the German EMAS-sites spend an average of 80.000 € (between 35.000 € and 130.000 €).<sup>51</sup> That is about half of the amount that Swiss companies spend on an EMS. This difference is much too high to explain it with the higher price-level in Switzerland. The allegedly simpler and less demanding system costs significantly more than the 'noble' EMAS?

---

<sup>45</sup> see Dyllick/ Hamschmidt 2000 p. 29f.

<sup>46</sup> Findings from German studies addressing this issue are not available to our knowledge.

<sup>47</sup> see e.g. Dyllick 1995 and Glatzner 1997

<sup>48</sup> see BMU/UBA 2000, p. 58-60

<sup>49</sup> see Stahlmann/ Clausen 2000

<sup>50</sup> see BMU/UBA 2000

<sup>51</sup> see UNI/ASU 1997



On the side of monetary benefits findings are quite similar. The Swiss study displays average monetary benefits (mostly cost savings) of 167.000 Sfr per year, the German study of about 50.000 €. The better part of these differences can therefore be explained by the different sizes of companies involved in the studies.

The payback period based on Dyllick/ Hamschmidt is 2.2 years; according to the UNI/ASU study it is 1.5 years. Many other German studies mainly report payback periods between 1.5 and 2.5 years. The available data regarding the different sizes of sites reveals a gap between smaller and medium/bigger sized companies. Especially for smaller companies the payback periods are much longer (up to 10 years).

At a closer look, the findings have to be put into perspective again. The empirical knowledge regarding EMS costs in that sense of the word is based on the statements of corporate managers in different questionnaires. Usually these statements are based on more-or-less precise data of corporate costing systems. These sources are of limited reliability. Most of the costing systems do not provide the possibility for exact cost allocation of environmentally relevant types of costs, especially of internal costs that do not directly lead to expenditures. The consulted corporate actors admitted that the statements on costs were mainly estimates. Additionally many managers who were asked in questionnaires do not exactly distinguish between EMS project costs and consequential expenses for environmental technical measures. This leads to considerable variations in the cost data given in empirical studies.

Furthermore, many companies have received financial aid from the government. Consequently, their statements about the level of costs do not remain under the planned costs quoted to the official sponsor. Besides these political circumstances there are micro-political interests to be considered.<sup>52</sup> The asked corporate actors combine their own internal goals with their statements. The environmental experts for example have to internally communicate their work as something that saves expenses rather than increases costs. These effects vary with the individual position and influence of the actors. It is not possible to take all these circumstances into account and translate them into financial parameters.

In ISO 14.001-companies as well as in EMAS-companies the internal costs amount to more than half of the total costs. Especially these costs are subject to problems in terms of accounting and consequentially to the validity. Therefore the findings of the ISO 14.001 study have to be carefully interpreted as well.

Taking a look only at the external validation/ certification, the environmental statement as demanded by EMAS and the registration, which are costs directly connected with expenses, the cost-findings get adjusted. EMAS-costs average out at about 18.000 €, while ISO 14.001 costs only 12.000 €. As expected, EMAS is significantly higher priced than ISO 14.001 on this field.

Even less valid than the estimated costs are the ascertained valuations of monetary benefits. Firstly, the accounting problem as mentioned above is even more precarious relating to the benefits. To impute cost savings by means of resource reduction directly to the implementation of an EMS is only possible in exceptional cases. Secondly, the above-mentioned problems of answers influenced by micro-politics occur here as well and even stronger, as economic benefits are the main internal argument to continue working with an EMS.

Concerning the effects on costs and benefits it has to be stated that the larger amount of intended benefits occur where monetary valuations are impossible to do. The implementation of an EMS has to be regarded as an investment. The costs have to be incurred immediately; the benefits accumulate over time and are often difficult to associate directly with the EMS. Finally, questionnaires are the least suited method to attain valid findings on this field. Therefore the quantitative data has to be interpreted with greatest care.<sup>53</sup>

Considering that the external benefits of validated/certificated EMS – although extremely difficult to evaluate - should be higher in the case of EMAS compared to ISO 14.001, because of the duty to publish an environmental statement, the necessary higher expenditure seems justified. To prefer ISO 14.001 only to save expenses is not advisable.

---

<sup>52</sup> see Freimann 1999b

<sup>53</sup> see Dyllick/ Hamschmidt 2000 p. 80



## Corporate benefits

It is not surprising that the stated corporate benefits too are nearly the same for EMAS and ISO 14.001 according to the different studies. Amazingly this counts even for the external corporate-political impacts. Due to the different constructions of the systems differences in this regard could have been expected.

In this field the impacts are much smaller than expected. Irrespective of the standard the EMS attract very little interest by the companies' stakeholders. Especially positive market effects were only slightly noticed by the surveyed managers. In contrast, they report on better contact and cooperation with authorities for environmental law, in the German EMAS-studies as well as in the Swiss ISO 14001-study. Managers predominantly mentioned overall image improvements as positive effects of EMS.

How can this be explained, considering the fact, that ISO 14.001 demands external information much less than EMAS, which sets an important focus on the public? Of course, even though it is not demanded, ISO 14.001 does not bar any user from the active communication about his environmental management and its ecological benefits. That is exactly what most of them do: 57% of the surveyed Swiss companies publish environmental statements or at least plan to do so. Inversely a noticeable part of the EMAS-companies do not perceive the chances of external communication, but regard it only as a bothering liability. Some companies even demand payments for their environmental statements. Public seems to be only slightly interested anyway. The authorities positively recognise the voluntary efforts for environmental care no matter whether based on the state-controlled EMAS or not.

In fact, positive corporate benefits through EMS evolve mainly inside the company organisation. The work with EMS leads to systematisation and documentation of competences and processes of operative environmental care. They affect sureness on the compliance with environmental laws. They help to get the idea of corporate environmental care – and with it the responsible persons – accepted by including corresponding rules into the company's basic principles and the therewith-documented support by the management. This happens firstly internally with the option to communicate it externally. Regarding the strategic dimension of corporate policy it has already been attested that EMS show little impact.<sup>54</sup>

## Methodical problems and perspectives

The insights in corporate environmental management and its substantial results as well as the experiences of companies with standardized EMS presented in this paper mainly come from empirical academic research projects and not from continuous internal or external evaluation of companies. This fact implicates some benefits but also several problems.

The most important benefit is that empirical academic research can afford to define specific goals and an adapted design for each study. It usually does not have to look after the costs unless at least one financier has declared his willingness to pay for the project. Those projects use to clearly define the goals and criteria, which shall be applied. Although they mainly focussed on corporate environmental care, they tried to evaluate the economic and social outcomes as well. Therefore one can say that the respective actual status of corporate sustainability was evaluated by the above-mentioned projects (see endnote 3). But the project-status of the evaluation leads to the fact that the produced results are valid only for the moment. All research projects that this paper informs about present snapshot-results of the social field they are dealing with.

On the other hand the surveys suffer from the problems as aforementioned: Most of the empirical research is based on oral or written interviews with corporate experts. The statements they give must be seen as subjective and often superficial judgements of more or less well informed persons, who mostly don't make much effort to answer the questionnaires. Even if they talk for example about costs or environment-technical figures such as quantities and qualities of waste, we do not know if they took a look into their technical or cost accounting documents before writing down the figures or if they just estimate. Even if we pay regard to all the sophisticated rules of empirical research we therefore cannot be sure that the results we present are valid and reliable.

Secondly no company can be forced to take part in questionings. They will take part only if they have any closer connection to the theme the questioning deals with. That is why usually only environmentally active or at least interested companies participate in projects about corporate environmental management topics.

---

<sup>54</sup> esp. Dyllick/ Hummel 1995, Dyllick 1999 und Freimann 1999a



The results of them do not really inform about the status in the general economy but only within its environmentally conscious sector.

Another disturbing factor is that the respondents are used to answer not only academic questionnaires but also questions of other public-agents such as journalists or environmental activists. They are professionals of corporate public relations, who are used to paint a positive picture of what they are asked for. Additionally they have to communicate their own social role within the company as a positive one, so that they will not give any information that could harm the company or themselves as environmental experts. This might lead to biases that cannot be eliminated from the results of empirical questionnaire-research.

But even if the respondents are willing to tell nothing but the truth and have the necessary knowledge, there remain some other problems. The usual financial and technical documentations of companies are not sophisticated enough that they show all those figures needed for comprehensive evaluation of corporate sustainability efforts. The normal cost accounting for example cannot exactly show the costs and benefits of special environmental measures, not only because some of the effects cannot be exactly evaluated in financial figures but also because the exact cost figures disappear in the general overhead costs of the company. A sophisticated environmental cost-and-benefit-accounting usually does not exist in 'normal' companies.

At least in EMAS-companies exists an environmental statement, which has to be published after the external validation-process by the accredited verifier. Most of the ISO-14.001-companies voluntarily publish environmental or sustainability reports as well. These documents could be used as more valid sources for the evaluation of corporate sustainability than questionings.

They really are subjects of empirical evaluation and rating at least in German speaking countries.<sup>55</sup> And, besides general information about the company, special information about its environmental policy, organisation, goals and measures, they give some 'hard' figures about the ecological and social impacts of the firm's activities and products. But the general guidelines for environmental reporting and the special directions within the EMAS regulation are so weak that most of the information given cannot be compared between different companies or branches. That is why most of the report-ratings do not evaluate the sustainability-status of the company but the features of the reports themselves such as completeness, intelligibility and credibility. Unless there are exact and binding rules for corporate environmental reporting, the existing reports do not present a valid base for corporate sustainability evaluation as well.<sup>56</sup>

To sum it up one can say that probably even most of the companies themselves do not know their own sustainability status, let alone what that could be. Any academic who tries to evaluate this status and compare it with other companies or with the average status of all companies in a branch or the whole economy, therefore is in a very difficult position. Even if he could clearly explain what corporate sustainability means, he up to now does not have the appropriate empirical instruments to evaluate it in a valid and reliable way.

There are three major steps to go towards evaluation of sustainability:

- First of all the scientific community has to define as exactly as possible the figures and indicators which are appropriate to measure the phenomenon of corporate sustainability.
- Secondly as many companies as possible should voluntarily or forced by law build up a corporate sustainability information system which validly delivers the figures needed for evaluation.
- And thirdly those figures have to be published in order to create the possibility to compare different companies and branches with respect to their sustainability performance without being dependent on the oral or written statements of corporate experts given in questionings.

There exist at least some academic projects, which try to make progress in step one. We mention again the efforts to standardize corporate sustainability reporting of the Global Reporting Initiative as well as other projects in this field.<sup>57</sup> Additionally there have to be mentioned the attempts to use the new strategic management tool 'balanced scorecard', a multi-dimensional strategic information and management system, to implement corporate sustainability management by defining figures and indicators of sustainability as corporate strategic goals.<sup>58</sup> But even if one considers this step as relatively easy to go, future must show if

<sup>55</sup> see e.g. Steven/ Schwarz/ Letmathe 1997, Loew/ Fichter 1998, Lange/ Ahsen/ Daldrup 2001

<sup>56</sup> see e.g. the first international guidelines for sustainability reporting by the Global Reporting Initiative (GRI) ([www.globalreporting.org](http://www.globalreporting.org))

<sup>57</sup> see IOEW/IMUG (ed.) 2001

<sup>58</sup> see Dyllick/ Schaltegger 2001 and Arnold/ Freimann/ Kurz 2001





there is a realistic chance for the next ones. The second step could be prepared by a national or European panel-research project in cooperation between EMS companies and academic research in order to develop and implement corporate sustainability information systems on a trial basis.

## Outlook

The message is clear: EMS are considered economically profitable investments by managers of the participating companies. Despite all the indicated problems of empirical research in this field, the scientists dealing with EMS are also in nearly complete agreement on the profitability. The systems furthermore lead to numerous additional valuable, but difficult to be financially rated, corporate benefits and they seem - even though it is still too early for definite statements - to support their main purpose: the improvement of corporate environmental care.

All empirical findings suggest, that the different normative fundamentals of the EMS, especially the state-run character of EMAS compared to the private norm ISO 14.001, have no significant influence on practical effects. The impacts and benefits of EMS are primary rather conditioned by the general strategic orientations that lead to the implementation of one of the systems.

Considering the different expenses for the implementation of the systems, participating in ISO 14.001 seems to be the better alternative. Higher costs for an EMAS implementation can only become justified, if stakeholders realise the significance of EMS in general and thereby notice the importance of a stronger public commitment as it is integrated in EMAS. As a state-run system the European Union and the national governments are asked for support in matters of publicity. In Germany a marketing programme for EMAS has already been started.<sup>59</sup>

Which strategies predominate in practice and which means are to be employed in order to persuade the corporate actors of their chances offered by environmental management systems, is still too little known. However, there is reason to be confident: In the context of the research-support programme 'Betriebliche Instrumente für nachhaltiges Wirtschaften'<sup>60</sup> by the German BMBF<sup>61</sup> a good many projects have started that may generate answers and solutions.

Moreover, in the future EMS research has to turn its attention more to the – theoretically well-founded – analysis of the internal corporate conditions in which the EMS have to fit in. The findings indicate that especially internally a lot of opportunities have been wasted so far. To utilise these, it is necessary that the structures put into practice harmonise with the organisational culture of the company. Culture contains various informal structures next to the formal ones.<sup>62</sup> This directs the attention to all employees instead of the common expert-orientated handling of EMS. Corporate environmental care has to "live". This means that the protection of the natural environment has to be realised beyond formal structures, with the involvement of every member of the company. Environmental care is a multifunctional task for everyone, managers as well as employees. The success of environmental management systems depends to a large extent on the motivation and participation of the personnel.

Some studies have been carried out on this topic. They are often based on new (sociological and economical) organisational theories like the concept of micro-politics.<sup>63</sup> Another very suitable approach is the modern system theory. The qualitative studies offer more insight into the internal processes, conditions and problems of implementation and the working with EMS. On the other hand they suffer from less generalisable results, as the micro-political and cultural conditions of companies vary considerably. The greater profoundness of the studies is only possible in exchange for the breadth of the studies. In order to offer solutions rather than pointing out problems and wasted opportunities we need to know more about these internal aspects. Further qualitative research is necessary to find patterns that are applicable to most of the companies.<sup>64</sup>

Besides the necessary theoretical analysis of corporate behavior in general and corporate sustainability management in particular, the methodical and informational basis for sustainability evaluation has to be improved. This must be the main purpose of corporate sustainability evaluation research in the near future.

---

<sup>59</sup> see [www.emas-logo.de](http://www.emas-logo.de)

<sup>60</sup> „Corporate instruments for a sustainable economy“

<sup>61</sup> Bundesministerium für Bildung und Forschung, the German Ministry of Education and Research

<sup>62</sup> see e.g. Smircich 1983, Czarniawska-Joerges 1992

<sup>63</sup> see e.g. Birke/ Schwarz 1994, Burschel 1996, Dücker/ Groth/ König 1999

<sup>64</sup> A research project on the effects of employees participation on the cultural fit of EMS is carried out at the moment by our research group.



## References

- Arnold, Wolfgang/ Freimann, Juergen/ Kurz, Rudi 2001: Vorüberlegungen zur Entwicklung einer Sustainable Balanced Scorecard für KMU, in: UmweltWirtschaftForum vol. 4, p. 74-79
- Baum, Heinz-Georg/ Sturm, Anke/ Wittmann, Robert 1996: Pioniererfahrungen mit dem Oeko-Audit, in: Umwelt Wirtschafts Forum vol.4, p.72-77
- Birke, Martin/Schwarz, Michael 1994: Umweltschutz im Betriebsalltag. Praxis und Perspektiven ökologischer Arbeitspolitik, Opladen
- Bundesumweltministerium (BMU)/ Umweltbundesamt (UBA)(Hg.) 2000: Umweltmanagementsysteme. Fortschritt oder heiße Luft? Frankfurt
- Burschel, Carlo 1996: Umweltschutz als sozialer Prozeß. Die Organisation des Umweltschutzes und die Implementierung von Umwelttechnik im Betrieb, Opladen
- Czarniawska-Joerges, Barbara 1992: Exploring Complex Organisations: A Cultural Perspective, Newsbury Park/Ca., London, New Dehli
- Dücker, Thea/ Groth, Thorsten/ König, Susanne 1999: Betrieblicher Umweltschutz und Partizipation, Köln
- Dyllick, Thomas 1995: Die EU-Verordnung zum Umweltmanagement (EMAS-Verordnung) im Vergleich mit der geplanten ISO-Norm 14.001, in: Zeitschrift für Umweltpolitik und Umweltrecht, 16.Jg. p. 299-339
- Dyllick, Thomas 1999: Wirkungen und Weiterentwicklungen von Umweltmanagementsystemen, in: Seidel, Eberhard (Hg.): Betriebliches Umweltmanagement im 21. Jahrhundert, Berlin/ Heidelberg, p. 117-130
- Dyllick, Thomas/ Hamschmidt, Jost 2000: Wirksamkeit und Leistung von Umweltmanagementsystemen. Eine Untersuchung von ISO-14.001-zertifizierten Unternehmen in der Schweiz, Zürich
- Dyllick, Thomas/ Hummel, Johannes 1995: EMAS und/oder ISO 14.001? Wider das strategische Defizit in den Umweltmanagementsystemnormen, in Umwelt Wirtschafts Forum Heft 3 p. 24-28
- Dyllick, Thomas/ Schaltegger, Stefan 2001: Nachhaltigkeitsmanagement mit einer Sustainability Balanced Scorecard, in: UmweltWirtschaftsForum, vol. 4, p. 68-73
- Forschungsgruppe Betriebliche Umwelpolitik (FBU) 1995: Pilot-Oeko-Audits in Hessen - Erfahrungen und Ergebnisse. Ein Forschungsbericht, in: Hess. Ministerium für Wirtschaft (Hg.) Schriftenreihe Technologie in Hessen, Wiesbaden
- Forschungsgruppe Evaluierung Umweltaudit (FEU) 1998a: Vorläufige Untersuchungsergebnisse und Handlungsempfehlungen zum Forschungsprojekt "Evaluierung von Umweltmanagementsystemen..", Oestrich-Winkel/ Frankfurt(Main)
- Forschungsgruppe Evaluierung Umweltaudit (FEU) 1998b: Umweltmanagement in der Praxis - Teilergebnisse eines Forschungsvorhabens zur Vorbereitung der 1998 vorgesehenen Ueberpruefung des gemeinschaftlichen Oeko-Audit-Systems, Teile I - III, in: Umweltbundesamt (Hg.) Texte 20/98 Berlin
- Forschungsgruppe Evaluierung Umweltaudit (FEU) 1998c: Umweltmanagement in der Praxis - Teilergebnisse eines Forschungsvorhabens zur Vorbereitung der 1998 vorgesehenen Ueberpruefung des gemeinschaftlichen Oeko-Audit-Systems, Teile V und VI, in: Umweltbundesamt (Hg.) Texte 52/ 98 Berlin
- Freimann, Jürgen 1999a: Jenseits von EMAS. Umweltmanagementsysteme – Erfahrungen und Perspektiven, in: Seidel, Eberhard (Hg.): Betriebliches Umweltmanagement im 21. Jahrhundert, Berlin/ Heidelberg u.a.O., p. 131-145
- Freimann, Jürgen 1999b: Akteursperspektiven im betrieblichen Umweltmanagement. Eine empirische Annäherung, in: Zeitschrift für angewandte Umweltforschung, 12.Jg. p.492-506
- Freimann, Juergen/ Schwaderlapp, Rolf 1996: Implementation of the EU's EMAS Regulation in German Companies, in: Eco-Management and Auditing, vol. 3, p.109-112
- Gerling Consulting-Gruppe 1995: Oeko-Audits in nordrhein-westfaelischen Unternehmen, in: Ministerium fuer Wirtschaft, Mittelstand und Technologie NRW (ed.) Duesseldorf
- Glatzner, Ludwig 1997: Öko-Audit: Was ist dran am Zertifikat? in: Globus Heft 1 p. 56-60
- Hamschmidt, Jost 2001: Economic and Ecological Impacts of Environmental Management Systems in Companies: Experiences from Switzerland, EURO Environment Conference, Aalborg, 18.-20. October 2000, CD-ROM
- Hartmann, Wolf D. 1998: Oeko-Audit: Ein Dutzend Erfahrungen aus der Unternehmenspraxis, in: Umwelt Wirtschafts Forum vol. 6, p. 32-36
- Hillary, Ruth (ed.) 2000: ISO 14.001: Case Studies and Practical Experiences, Greenleaf Publishing, Sheffield



- Hoepfner, Nils-Oliver/ Sietz, Manfred/ Seuring, Stefan/ Czaja, Frank 1998: Oeko-Audit-Verordnung - Ergebnisse einer Befragung validierter Unternehmen, in: Umwelt Wirtschafts Forum vol. 6, p.73 - 76
- IOEW/ IMUG (ed.) 2001: Nachhaltigkeitsberichterstattung – die Praxis glaubwürdiger Kommunikation zukunftsfaehiger Unternehmen, Berlin
- Issak, Robert/ Keck, Alexander 1997: Die Grenzen von EMAS, in: Umwelt Wirtschafts Forum vol. 5, p.76-85
- Jaeger, Thomas/ Wellhausen, Anja/ Birke, Martin/ Schwarz, Michael 1998: Umweltschutz, Umweltmanagement und Umweltberatung - Ergebnisse einer Befragung in kleinen und mittleren Unternehmen, Koeln
- Klemisch, Herbert 1997: Ergebnisse des Forschungsprojektes 'Oeko-Audit und Beschaeftigtenbeteiligung in KMU', in: Klemisch, Herbert (ed.) Oeko-Audit und Partizipation - Die betriebliche Umsetzung von Umweltinformationssystemen in kleinen und mittelstaendischen Unternehmen Koeln, p.5-22
- Lange, C./ Ahsen, A.v./ Daldrup, H. 2001: Umweltschutz-Reporting, München
- Loew, T./ Fichter, K. 1998: Umweltberichterstattung in Deutschland und Europa, Berlin
- Schnauber/ Kiesgen/ Mangelmann 1994: Analyse des betrieblichen Umweltschutzes unter besonderer Beruecksichtigung des Umwelt-Audit-Konzeptes, in: Ruhr-Universitaet Bochum/ INNOSYS - Gesellschaft fuer Innovative Arbeitssysteme (ed.) Bochum
- Seidel, Eberhard/ Weber, Frank M. 1998: Die EMAS-Praxis in Deutschland, in: Umwelt Wirtschafts Forum vol. 6, 1998, p.22-27
- Smircich, Linda 1983: Concepts of Culture and Organizational Analysis, in: Administrative Science Quarterly 1983, p. 339-358
- Stahlmann, Volker/ Clausen, Jens 2000: Umwelleistung von Unternehmen. Von der Öko-Effizienz zur Öko-Effektivität, Wiesbaden
- Steven, M./ Schwarz, E.J./ Letmathe, P. 1997: Umweltberichterstattung und Umwelterklärung nach der EG-Oeko-Audit-Verordnung, Berlin
- Umweltakademie Fresenius/ IHK Dortmund 1996: Kopplung von Qualitaets- und Umweltmanagement - Auswertung einer Befragung von 3000 Unternehmen in Nordrhein-Westfalen, Dortmund
- Unternehmerinstitut/ Arbeitsgemeinschaft Selbständiger Unternehmer (UNI/ASU) 1997: Öko-Audit in der mittelständischen Praxis, Bonn



**SUSTAINABILITY**

**Edwin Zaccai, Belgium**

From Conceptions of Sustainability to Indicators

**Thiemo Eser, Germany**

Evaluation and the Appropriate Policy Level under  
the Demands of Sustainability

**Markus Langer/Aloisia Schön, Austria**

An Integrated Referential Framework for  
Sustainable Development



## Edwin Zaccai

### FROM CONCEPTIONS OF SUSTAINABILITY TO INDICATORS

#### *Summary and notes*

---

Progresses in sustainable development can only be assessed within a frame of reference. The official expressions that define sustainable development are relatively consensual and do not necessarily provide enough criterions upon which assessment can be based.

"SD is development that meets the needs of the present without compromising the ability of future generations to meet their own needs"  
(WCED 1987, "Brundtland Report")

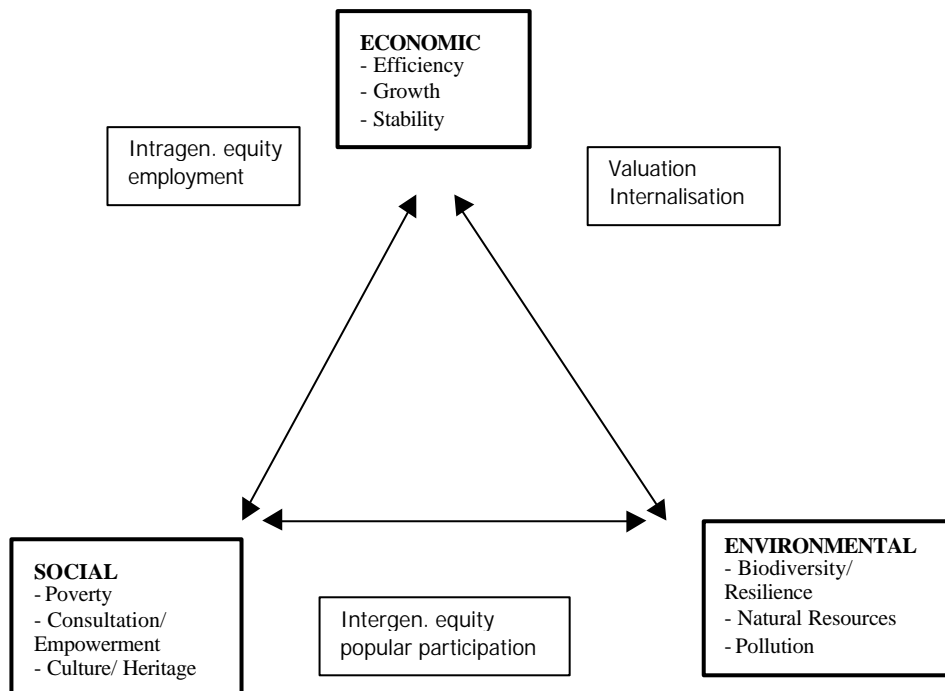
- Stays in the UN development tradition
- Sets conditions on the future (long term)
- (NB. There are other definitions by the WCED)

SD means "improving the quality of human life while living within the carrying capacity of supporting ecosystems"  
(IUCN-UNEP-WWF, 1991, "Caring for the Earth : a Strategy for Sustainable Living").

- Adds a condition on natural environment, inspired by an ecological model

SD " requires dealing with economic, social and environmental policies in a mutually reinforcing way".  
"The Union's Sustainable Development Strategy is based on the principle that the economic, social and environmental effects of all policies should be examined in a coordinated way and taken into account in decision-making"  
(Presidency conclusions, Göteborg European Council, June 2001, § 19 and §22).

- Taking up of this formula for corporate sustainability ("triple bottom line")
- Listing of diverse combinations of achievements belonging to different "dimensions"
- Indicators for different criteria in each "dimensions"
  
- To integrate, to balance the economic – social – environmental dimension
- What means a well-balanced approach, with different dimensions and different criteria
- Must every single actions have these three aspects "balanced" ?
- What for environmental action that are not economically profitable ?



(World Bank, 1993-1995)

SD is "non declining consumption per capita, or GNP, or whatever the agreed indicator of development is"  
(D. Pearce 1998, *Economics and Environment*)

- Indicators for measuring are explicitly introduced

## Conclusions

Plurality of conceptions

Different sources propose different sets of criteria that characterize SD.

One aggregation of these criteria (Zaccai 2002) : importance of environment / global vision / long term / integration / innovative image

Consensus at the first level, disputes and contestations at the second level (on how the concept should be interpreted in practice). "Contestable concept" (M. Jacobs 1998), cf. democracy, justice, etc.

Evaluation concerns the second level

But SD contains also a charge of image, (ethical) principle, project for the Society (alternative to Progress ?), which cannot be reduced to evaluation.

*"The phrase "sustainable development" has staying power because most people want to believe in it. It survives because it appears to build bridges between the demands of environmentalists and developers. It sounds comforting - human well-being and economic security for ever, not brought to heel by ecological collapse or social distress. It is an article of faith, and in that sense almost a religious idea, similar to justice, equality and freedom" (Pearce 1993: 183-184)*



### Ex.1. : European Union, SD Strategy, 2001

Integrating environment into Community policies

Decoupling economic growth from resource use

Priorities :

- Combating climate change
- Ensuring sustainable transport
- Addressing threats to public health
- Managing natural resources more responsibly

*(Presidency conclusions, Göteborg European Council, June 2001)*

### Ex. 2. : World Summit on Sustainable Development, 2002

- Poverty eradication
- Changing unsustainable patterns of consumption and production
- Protecting and managing the natural resource base of economic and social development
- Sustainable Development in A Globalizing World
- Health and Sustainable Development
- Sustainable Development of Small Island Developing States
- Sustainable development initiatives for Africa

*(Commission on Sustainable Development Acting as the Preparatory Committee for the WSSD, Chairman's Text for Negotiation, 9/5/02)*

### Some questions

#### Table of contents

- European level : add the environmental pillar to a socio-economic strategy (Lisbon 1999). Also opening towards a balanced project
- UN level : social dimensions are explicitly involved, but for some topics (poverty, health), and not others (ex. employment, education)

Implications for evaluation : which topics to be included

#### Process

Are we in presence of real coherent plans, or of catalogues of separated actions gathered for ad hoc presentation ?

Coordination : who, at what level, etc.

Implications for evaluation : when and how in the process

### Evaluation by indicators

Today there is a trend favouring indicators in the assessment of policies. But the term "indicator" hides in fact a wide variety of data types.

- Which indicators are to be **chosen** according to the actions to be assessed?
- The indicators bound to **changes occurring physically** (e.g. pollution reduction) have for instance to be differentiated from those relating to the **means used** (e.g. number of legislations, or investment amounts).
- What optimal level of **aggregation** is to be used, and what coordination is to be aimed at between different indicators systems?
- What is the place of analysis on the one hand, and of quantitative aspects on the other?



## See in the following examples

### EU : "Structural indicators"

(Com (2001) 619 final)

5 categories (Economy, social, ...)

Environment :

- Greenhouses gases emissions / Energy intensity of the economy / Volume of transport relative to GDP / Modal split of transport / Urban air quality / Municipal waste

Set of indicators allowing comparison between countries (but with what purpose ?)

Coherency with some objectives of the EU SD Strategy ("decoupling") and not others (natural resources)

6 env. indicators are from sufficient for political evaluation

### Stockholm's Local Agenda 21 Program

(in A. Atkisson 2002)

Miljö	Trend
Energianvändning per person och år	↗
Sopmängden per person och år	→
Tillförseln av tungmetaller	↗
Utsläpp av koldioxid per person	→
Antalet dagar med bra luft	↗
Andelen trafikanter som åker kollektivt	→

Ekonomi	Trend
Bruttoregionprodukt per person och år	↗
Sysselsättningsnivå	↗
Utbildningsnivå	→
Försäljningsvolymen av miljömärkta livsmedel	↗

Social utveckling	Trend
Astma	↗
Andelen personer som känner ekonomisk trygghet	↗
Andelen personer som är rädda för våld	↗
Barns tid ihop med vuxna under sin uppväxt	?

Demokrati	Trend
Andelen ideellt engagerade personer	?
Valdeltagandet bland förstagångsväljare	↗
Antalet personer som känner sig delaktiga i samhället	?
Andelen ungdomar under 25 år som anser sig kunna påverka samhällsutvecklingen	?

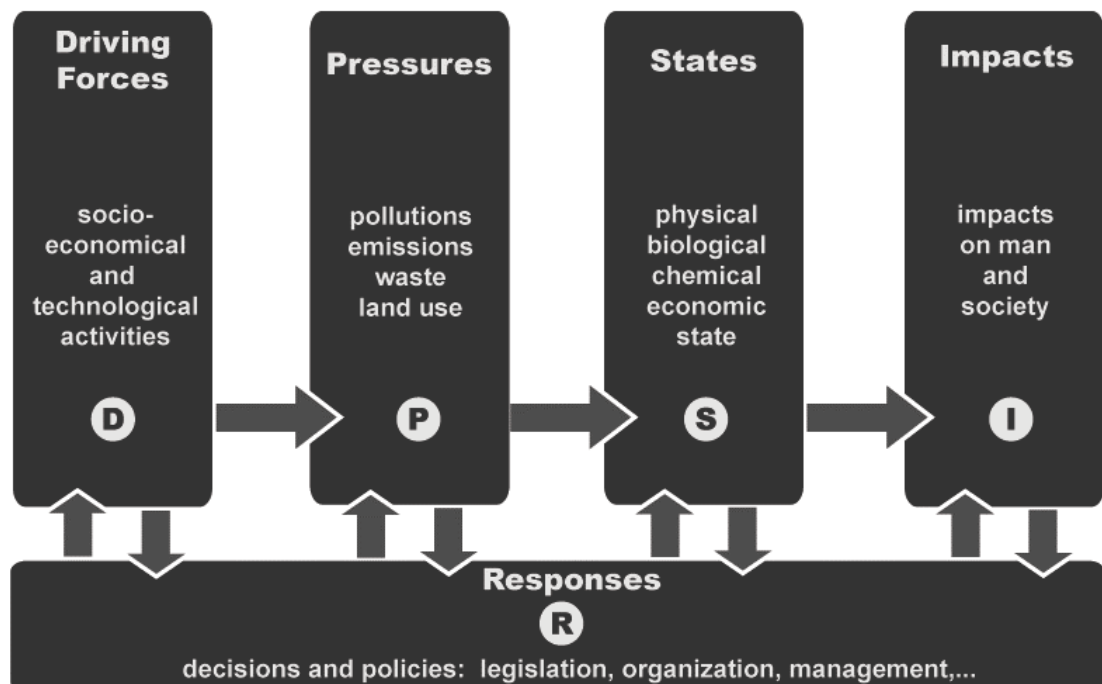




The UK's "Headline Indicators"

<a href="#">H1</a> Economic output	✓	✓	Improved	2000	
<a href="#">H2</a> Investment	✗	✗	No change	2000	
<a href="#">H3</a> Employment	≈	≈	No change	2001	
<a href="#">H4</a> Poverty	✗	≈	No change	2001	
<a href="#">H5</a> Education	⋯	✓	Improved	2000	
<a href="#">H6</a> Health	✓	≈	No new data	1998	
<a href="#">H7</a> Housing	⋯	≈	No new data	1996	
<a href="#">H8</a> Crime	Violent	✗	✗	Deteriorated	2000/ 2001
	Vehicle, burglary	✗	✓	Improved	2000/ 2001
<a href="#">H9</a> Climate change	✓	✓	Improved	2000	
<a href="#">H10</a> Air quality	⋯	✓	Improved	2000	
<a href="#">H11</a> Road traffic	✗	≈	No change	2000	
<a href="#">H12</a> River water quality	≈	✓	Improved	1999	
<a href="#">H13</a> Wildlife (Farmland birds)	✗	✗	No new data	1999	
<a href="#">H14</a> Land use	⋯	≈	Deteriorated	2000	
<a href="#">H15</a> Waste	⋯	✗	No new data	1997/ 98	

Indicators for the Environment : The "DPSIR Model"

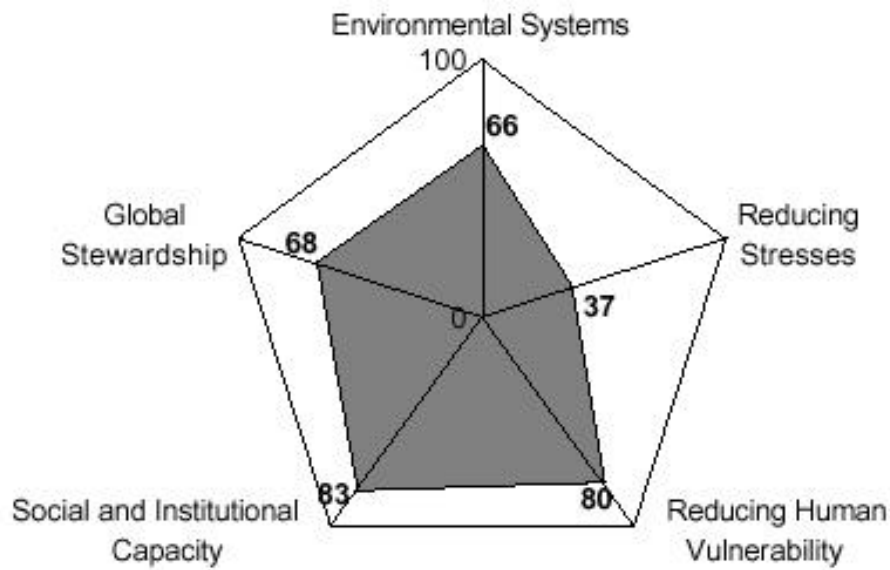




### Environmental Sustainability Index

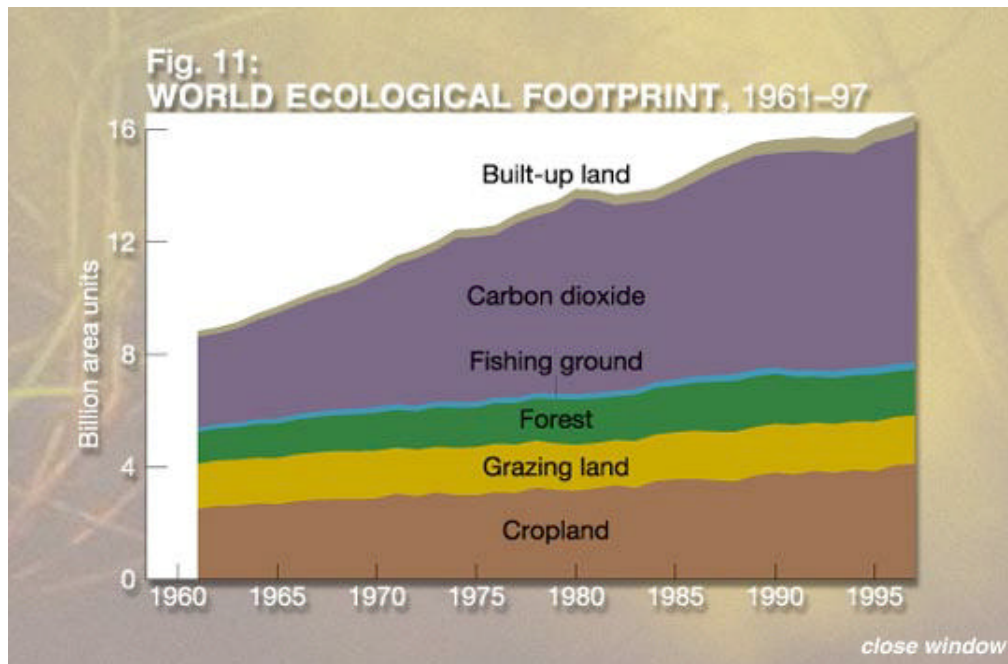
(World Economic Forum)

Ex.: Austria



### World Ecological Footprint

(WWF)





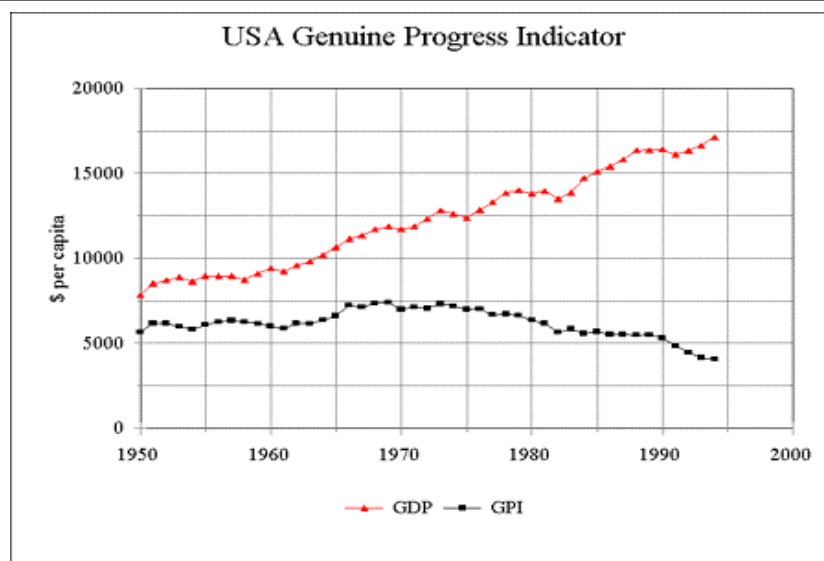
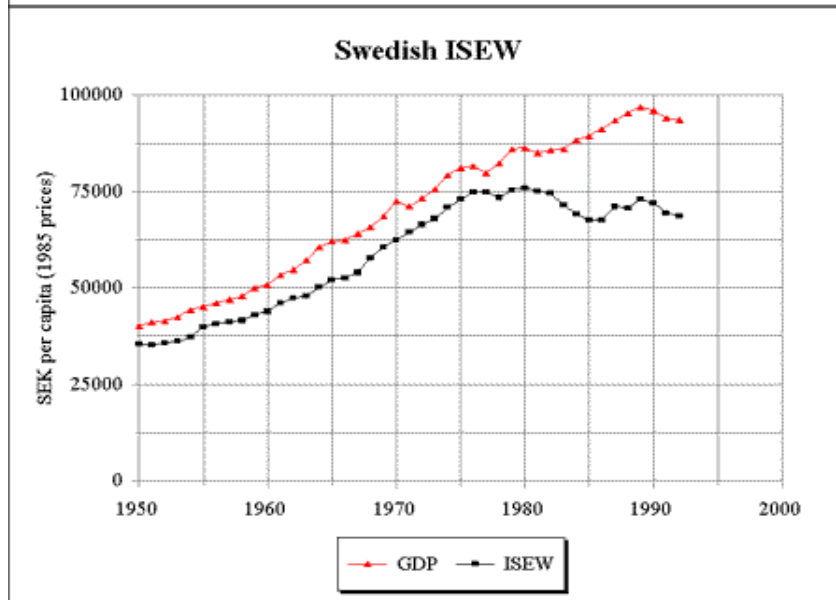
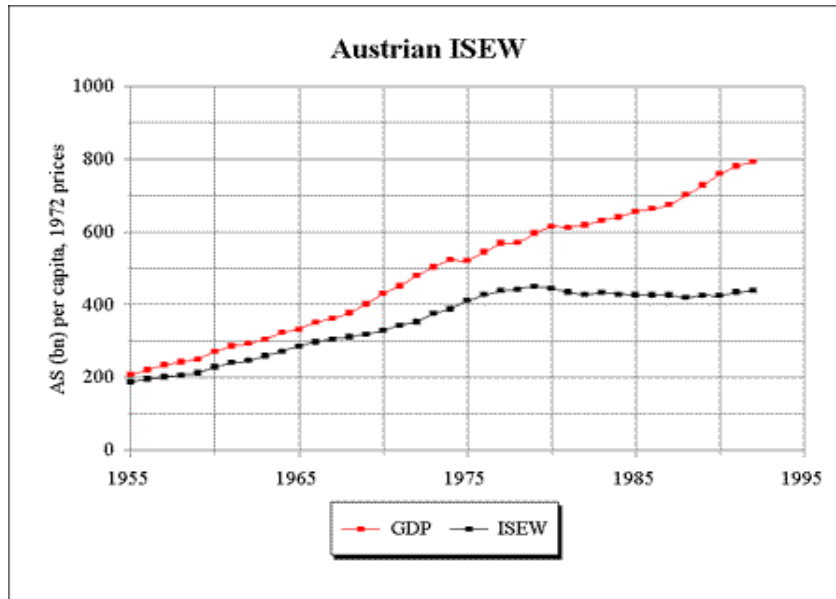
**Human Development Index**  
(PNUD)

6 Examples of ranking

1999		Life expect. at birth	Adult literacy rate	Enrolment ratio	Real GDP p.c. (USD)	Human development index (HDI)	Real GDP p.c. minus HDI rank
4	<b>Sweden</b>	79,6	99	100	22 636	0,936	+13
5	<b>Belgium</b>	78,2	99	100	25 443	0,935	+4
6	<b>USA</b>	76,8	99	95	31 872	0,934	-4
85	<b>Albania</b>	73	84	71	3 189	0,725	+16
86	<b>Dominic. Republic</b>	67,2	83,2	72	5 507	0,722	-16
87	<b>China</b>	70,2	83,5	73	3617	0,718	+7



### Index of Sustainable Economic Welfare (Daly and Cobb)





**Thiemo W. Eser**

## **EVALUATION AND THE APPROPRIATE POLICY LEVEL UNDER THE DEMANDS OF SUSTAINABILITY**

### **1 Introduction**

The discussion of sustainability reaches in all dimensions of daily life and governmental practice. Reviewing the literature it seems obvious that, in principle, it is clear where to the society should develop. However, important measures are not put into place although the necessity is rather obvious. This finding highlights deficiencies in the field of research about the implementation of sustainability in particular on the question how can we best make use of our existing institutions in order to achieve a sustainable development. Therefore, The main points I would like to address under this topic are the following ones:

- Which criteria allow best investigating the appropriate policy level in order to contribute to the achievement of sustainability, and how should these criteria be applied?
- Do federal systems offer better opportunities to deal with sustainability? The purpose of the paper is to clarify in which way sustainability can offer a guiding principle for federal states. For this purpose it needs to be recorded what federal states means (section 2.1) and which specific starting points with regard to sustainability are available. The two aspects responsibility assignment (section 2.2) and co-ordination (section 2.3) between both the levels and the states are in the centre of consideration. The third chapter offers the operationalisation of sustainability with regard to the question of a guiding principle<sup>65</sup>. After the delineation of the different dimensions of sustainability (section 3.1), the thematic dimension of economy (section 3.2), environment (section 3.3) and social (section 3.4) is explained. In section 3.5 the dimension of monitoring by the society is discussed. Finally (section 4) the conclusions for the guiding principle of sustainability for federal structures will be drawn.

### **2 Federal Governments**

#### **2.1 Sketching federal systems**

In the following, federal government and federal system are used as synonyms and treated in contrast to unitary governments<sup>66</sup>. It seems to be advisable to check first the definition of federal governments. Federation means an alliance of states for a temporarily political and economical purpose<sup>67</sup> whereas federalism signifies a principle of a state that allow every single member wide autonomy<sup>68</sup>. These definitions are used as an initial point, because federalism inheres much more that is of pertinence for sustainability: A scope from federalism as a wide social concept to a narrow political instrument, federalism and cross-national relationships as special political phenomena<sup>69</sup>.

Regarding constitutional law the federal state (federation) must be distinguished from the federation (confederation) by the relation between the governmental levels. Members of a federation keep their status as states and the federation represents also a state in itself<sup>70</sup>. Pivotal is the constitutional anchor of the federalism: accordingly levels of the country (Länder) and the federation can be identified for Germany or Austria.

---

<sup>65</sup> Many authors tried to operationalise the keyword *Sustainability* and the result was more or less felicitous. The following article is based on the considerations of the Enquete Kommission of the German Parliament and its survey (cited below). If sustainability should be researched as a guiding principle for federalism, relatively clear outlines are needed. So far the Enquete Kommission provides an anchor, which rest upon a relatively wide acceptability.

<sup>66</sup> Brockhaus online

<sup>67</sup> Brockhaus online keyword „federation“

<sup>68</sup> Brockhaus online keyword „federalism“

<sup>69</sup> See Elazar 1968, S. 354f.

<sup>70</sup> See Münch 1993.



The inclusion of the municipalities is not applicable in this definition. In a broader sense of federalism they should be included in the same way as the constitutional levels. The EU-level should be included as well as already a certain constitutional status is established, even if it is often denied<sup>71</sup>.

The legality of all constitutional states is built on the division of power between the legislative, executive and juridical authorities on the horizontal level stands. However, big differences between the states exist in view of its definition. In contrast to a centralized form of government, where all centralized levels disposes the power of decision in one country, the federalism allows a diversion of powers in also a vertical regard, which means that national functions are separated into two national levels, a central or federal and a regional level.

Motivations to employ federal systems touch a range of issues, which are back-to-back synthesised. The perception outlined above articulates clearly, that a pure national definition for such a discussion is too narrow<sup>72</sup>:

- The most important motivation lies, indeed, in the *vertical division of power*. This creates a high transparency and control in the regard of functional, spatial and factual aspects<sup>73</sup>.
- Furthermore a *better of more efficient fulfilling of public task* can be identified by the application of further criteria. At that point it becomes clear that the focus only on state levels in the narrow sense would not be allow to address the potential of federal systems<sup>74</sup>.
- A more extensive demand lies in the *subsidiarity* within the execution of tasks in federal systems. Apart from the division of powers and a better responsibility assignment – the principle of subsidiarity defines: upper levels only become active, if lower levels are not able to discern their tasks<sup>75</sup>.

These positive motivations of federalism are confronted with special risks, that are mainly related to co-operation between the different levels<sup>76</sup>. The multiplicity of the information-, co-ordination- and co-operation forms throw a glance at the possible risks. A risk of blockades are given in case of:

- various common competences of decision making of the different levels;
- the delineation and limitation of competences between the levels is not clear;
- there are too big differences between the states with regard to their economic strength, size and their political power.

The kind of modalities how and how far consensus must be achieved play a key role in this respect – is it real consensus or imposed, competitive or consensus model of federalism? <sup>77</sup>The disadvantages mirror the advantages of the unitary state, but the advantages persist in principle if the disadvantages are accordingly dealt with.

Now how does the European landscape does look in that respect? It seems to be quite difficult to classify the European States, because every state shows an inimitable and a grown structure<sup>78</sup>. Nevertheless it appears advisable, to make a classification into three categories: a federal system well established on the base constitutional regulation (1) unitary system with regional administration and some kind of regional representation, but which is not of constitutional status (2) and a purely central regime (3). Among all these already mentioned forms, the level of the regional authority is allocated<sup>79</sup>.

---

<sup>71</sup> It becomes obvious, if one keep a view on the significance of the European tribunal, which dispose the increasing competences across the national constitutional courts.

<sup>72</sup> As a fundamental principle of federalism, the following aspects are pointed out: not-centralisation, democracy, “check and balances“ open bargaining processes, constitutionalism and save units of the sub division of state. see Elazar 1995, 475ff.

<sup>73</sup> See Laufer 1991

<sup>74</sup> See for example Beer 1977 and Oates 1972.

<sup>75</sup> See Eser 1996 for different interpretations (catholic social doctrine (Soziallehre), economic perspective). Quite often one forget that to be „in the position of being able“ also includes co-operation on lower levels, where possibly only stimuli from above are given, without adopting the tasks within one step.

<sup>76</sup> See Kilper/Lhotta 1996

<sup>77</sup> Solving problems, bargaining (amicable) Paternalism, imposed, Scharpf 1995, 30.

<sup>78</sup> See European Commission Luxembourg 1997, S. 39.

<sup>79</sup> Own version according to Sauerland 1997.

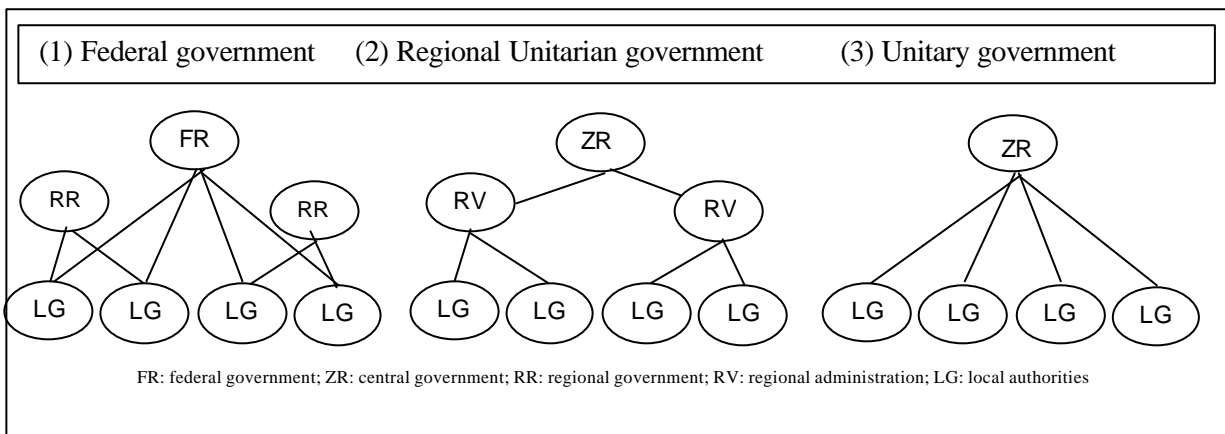


There is a common understanding after the decentralisation efforts in some EU counties the following assignment is most appropriate<sup>80</sup>.

- (1) Austria, Belgium, Germany
- (2) Italy and Spain with longer tradition, but now after more recent decentralisation (in the UK called “devolution”) efforts also France and the United Kingdom
- (3) Denmark, Finland, Greece, Ireland, Luxembourg, the Netherlands, Portugal, Sweden.

It should be mentioned that all groups of countries include strong municipalities and in some cases in particular in unitary countries municipalities dispose of reasonable power.

Figure 1: Federalism and unitary states in practice



For simplification purposes of the theoretical discussion, the two forms (federal and unitary regimes) will be used as two poles of a continuum, whereas the rationalised unitary regime supposed to be in the middle. Federalism allows in contrast to unitarism and centralism that two levels among themselves are in direct competition (vertical competition)<sup>81</sup>, in case the particular constitution does not provide a clear responsibility assignment between the levels. On the other hand there is normally competition on the regional level between the regional authorities (horizontal competition)<sup>82</sup>. Competition means in that case that the particular level is allowed to a certain decision-making and responsibility, whereas within the unitarism no regional competences are available. Competition takes place within the scopes taxation, rational services, regulations and promotion of economic developments. Competition depends from the number of competing units, as well as from the level of autonomy and competition that is fixed by the higher levels<sup>83</sup>.

<sup>80</sup> See Wiehler, F./Stumm, 1995

<sup>81</sup> In terms of competition about competences, whereas a displacement of competences (upturned) is immanent, as long as no counterbalance, such as demanding of the principle of subsidiarity, will be universally accepted.

<sup>82</sup> According to Tiebout, 1956.

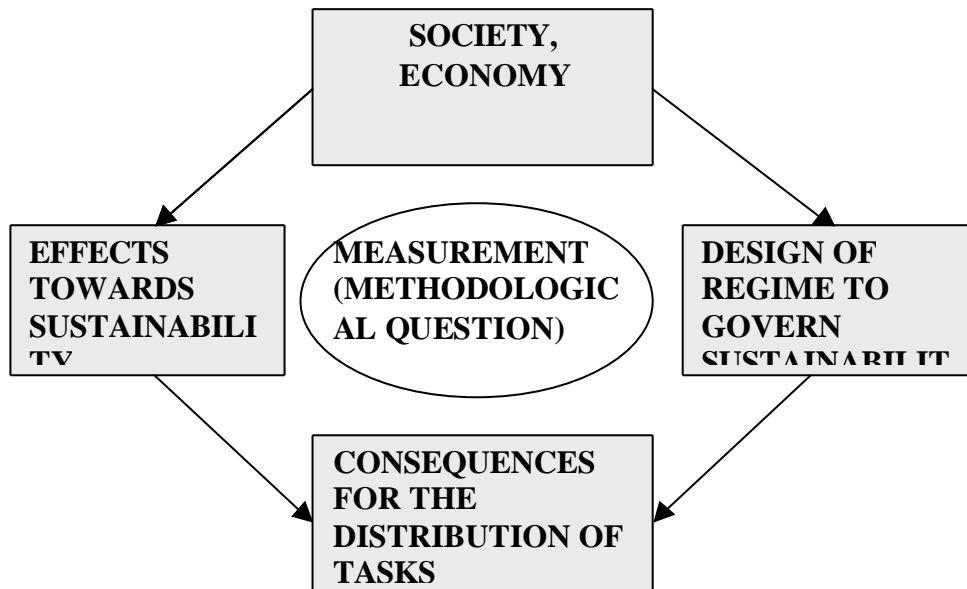
<sup>83</sup> See overview in Keynon 1997, 26.



## 2.2 The allocation of tasks

The allocation of tasks between the national levels within the federalism follows the question, which functions with regard to which tasks according to which norm should be allocated at which governmental level? As shown in figure 2, two dimensions must be considered in order to answer the question.

Figure 2: general framework for the allocation of tasks in a federal system



Within the first dimension it will be observed, to what extent the society defines the public tasks on the basis of pertinent demand as well as how far there are thematic interdependencies of impacts that have consequences of the responsibility assignment between the levels. The second dimension considers the current political and social institutions, which administrate the presentation of these tasks. There the appropriate allocation of tasks should take place. Against a constitutional law and political scientific background, there is a requirement of division of powers and democracy. In principle that means to federal systems, that only those should decide about a public task who are both spatial and factual affected. But often the congruence between those who are affected and those who decide it is not given. Ideally only those decide on tasks who are directly affected. This can only be achieved by a regime, which is capable to allow interventions on different governmental levels. In practice, often decision-, administration- and financial functions are subdivided into different federal levels with regard to a public task. The classification of these functions allows in the case of interactions between the regional authorities on different levels co-decision and frame setting. But in practise it is often necessary to “create” appropriate delineation and level for governance by the cooperation and coordination of the governmental units in order to achieve the congruence required.

## 2.3 Co-ordination

Now institutional arrangements come to the fore, which allow coordinating intersections and conflict of the allocation of tasks between units of one level in federal systems. This kind of coordination usually goes beyond the mandatory processed prescribed by the constitution of a federal (or unitary) country. The institutional forms are multifarious: they range from legal and partial constitutional, such as common tasks, delegation systems<sup>84</sup>, duty tasks to compulsory and voluntarily built associations for specific purposes and (standing) conferences; the latter are in a legal sense barely connected with any obligations<sup>85</sup>. Especially within the regional policy many forms of these co-operation and team-work are applied. Mostly the level between the countries and the municipal level and regional authorities are involved.

<sup>84</sup> See Postlep 1993

<sup>85</sup> See Eser 1997.





Figure 3 reveals institutional options which exist and which allow putting that kind of consensus building processes into operation, which supports the concerned congruence between the effected individuals (in the end the beneficiaries) and those who decide (the ones who usually have to pay). Consequently an appropriate level of co-operation plays a key role in that respect. It should be mentioned that these mechanisms of coordination often create the problem of legitimacy and liability, as the influence of the democratic key institutions such as parliaments and politicians are partly put out of the process<sup>86</sup> and representatives of administrations overtake their places: This was already interpreted as “co-operation in the shadow of the hierarchy”<sup>87</sup>.

Figure. 3: communication in federal systems beyond mandatory processes

	<b>Forum</b>	<b>Conference</b>	<b>Association (in a legal form)</b>
<b>Tasks</b>	Exchange of ideas	Concepts, priorities, management	Implementation, execution
<b>Structure</b>	All interested	Networks organisations	Mandatory bodies
<b>Actors</b>	Administrators, interested public	Administrators, representatives	Mandatory representatives
<b>Processes</b>	Contact, exchange	Coordination, bargaining	Clearly defined tasks implemented

This brief overview only allows indicating the access points for federalism and therefore only pinpoints the frame for the implementation of sustainability. But first it is necessary to better define the demands of sustainability for governmental regimes and structures.

<sup>86</sup> Quite clear displaced in Kregel 1998.

<sup>87</sup> See Benz u.a. 1998, 21.

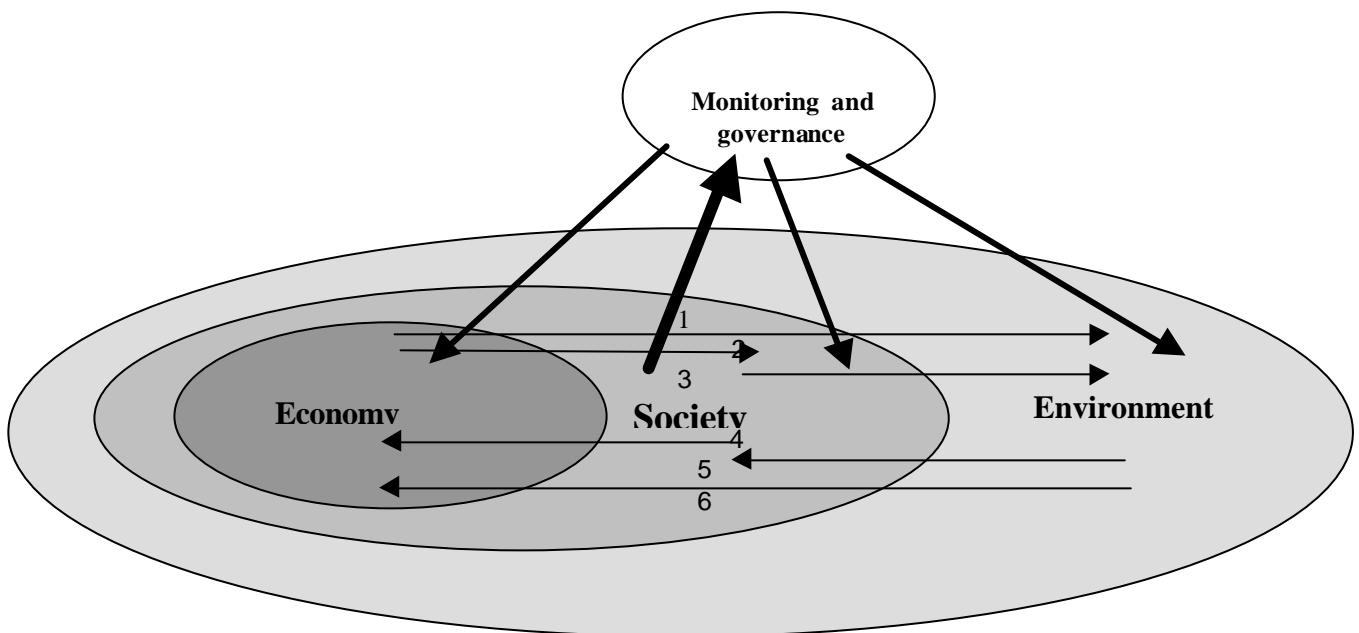


### 3. Sustainability

#### 3.1 Dimensions

Stating from the definition of sustainability in the Brundtland-report<sup>88</sup>, there exists a broad consensus that a distinction of the three dimensions economy, environment and society is useful with regard to the discussion of sustainability<sup>89</sup>. However, there is not any consensus in particular when it comes to down to the last dimension as the quintessence can be circumscribed as social, participation and politics as well<sup>90</sup>. This wide coverage reveals a certain dilemma, as it gathers about everything that cannot be summed up under economy or environment. With regards to a narrow focus of this dimension there is an area of conflict between the thematic aspects (social) and monitoring aspect (policy)<sup>91</sup>. Figure 4<sup>92</sup> sums up the context.

Figure 4.: the relation between dimensions of sustainability



The scheme shows that the three dimensions economy, environment and society lie on one level as thematic dimensions. Therefore a differentiation seems to be appropriate which considers three thematic dimensions devoted to environmental-, economical- and social concerns, which need be in balance. The balance between all three dimensions is achieved via the societal monitoring only, which therefore can be understood as a independent institutional dimension on its own.

<sup>88</sup> „Sustainable Development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs“ World Commission on Environment and Development 1987, 46. Out of it over 70 definitions could be deduced see Kreibich 1996, 40.

<sup>89</sup> In connection with that trisection, it is quite often spoken of aims. In practice it proves that even in those cases where a distinction is made between objectives or aims and dimensions as e.g. at the Enquete Kommission „Schutz der Menschen und der Umwelt“ on the „concept sustainability“ it was not possible to hold a clear distinction through because the aims are dealt with under the topic „Dimensions“ and translated in „requirements for action (Handlungserfordernisse)“ (Enquete Kommission, 1997, S.44). These requirements flow into further elaborated aims or, as in the case of the environment in the Management concepts.

<sup>90</sup> See for example Blower 1993, 6ff.

<sup>91</sup> Wehling 1997, 36 characterizes the phenomena quite aptly: „The idea of a sustainable development consists of a normative postulate, which says that social developments hat to be created in that way, that the natural base is supported to sustain future generations. On the other hand it updated, that the analytic perception social and analytical developments are not possible to be kept apart but that the transformation must be included. On that background sustainability has to be seen as a political concept with theoretical implications. The theoretical implications are still open.“

<sup>92</sup> According to Bush-Lüthy.



As a result, the sustainable development has to be interpreted as a regulative idea, that settles on future-orientated social learning-, searching- and formation process, which features openness and lacking knowledge about the future<sup>93</sup>. With it, the question of monitoring becomes a central subject for the understanding and implementation of sustainability. Now at first the thematic and than the institutional dimension are addressed with reference to sustainability.

### 3.2 Economy

If the society has selected the market-economy as the economic system of their choice, the question of the allocation of public tasks in order to avoid market failure and in order keep the market economy running. The aim and the function of the market economy is the provision of goods and services for the people<sup>94</sup>. It is also assumed that a separation of allocation and other social aspects is possible<sup>95</sup>. Now the relevant question for federalism is, where, or more precise, on which level should be which task allocated in order to achieve best results in the functioning of the market economy.

The market-economy represents the strongest decentralised economic system, because the economic subjects schedule individually their plans – both on the supply and the demand side. If market failure occurs<sup>96</sup>, public intervention is recommended. But firstly it has to be proved, whether it is possible prevent the failure of the market by the self-organisation and co-operation of the economic subjects. The failure of the market because of the lack of information (lemon markets) and with regard to the quality assurance can be reached for example by chambers and industrial unions. The systematical underestimation of risks can be reached through the obligatory insurance for example in health sector and with pension funds. An account of market failures is given in figure 5.

Figure 5: Policy level relevant aspects of economic sustainability

<p><b>A) „Classic“ arguments</b></p> <ul style="list-style-type: none"><li>- Market economy as a decentralised system</li><li>- Market failure demands for more centralised action:<ul style="list-style-type: none"><li>o Regional monopolistic markets</li><li>o Negative external effects: environmental effects</li><li>o Positive external effects: location management</li><li>o Public goods – free rider positions</li></ul></li><li>- Macro-economic interventions, stability, distribution</li></ul> <p><b>B) „New“ arguments</b></p> <ul style="list-style-type: none"><li>- Material flow orientated economic development</li><li>- Valued added chains, import substitution More differentiated approach towards regulation possible in cases of regionally limited market failure, often connected with administrative boundaries</li></ul>
---

In terms of federalism and sustainability it is important to allocate the task at an appropriate level. That means for example to delegate dealing with external effects only to those governmental units which are directly involved on the causes and effects side<sup>97</sup>. The same applies in the case of income distribution; relevant is the on the effect side how far in spatial terms is migration invoked by redistribute measures<sup>98</sup>. Now in order to reach a sustainable economy all these cases of market failure need to evaluate with reference best level of

<sup>93</sup> See Minsch u.a. 1998, 18.

<sup>94</sup> Under the umbrella of the market-economy different kinds of interpretations do exist. See Enquete Kommission 1997, Minsch 1998, 22.

<sup>95</sup> In practice, it is not possible to divide, which is shown within the Coase Theorem, if allocations- and disposition aspects are taken into account.

<sup>96</sup> In the following, only the most important aspects will be activated: external effects, lacking possibility to devide goods, lack of information and a lack of conformation. See Fritsch u.a. 1991.

<sup>97</sup> Furthermore the World Trade Organisation is involved and represents the global level.

<sup>98</sup> That means, that people benefit from the measure of redistribution in local authorities with high degree of redistribution. Those, who lose because of the measure of redistribution, orientate themselves to the local authorities with low levels of redistribution. The resulting segregation leads to the fact, that relatively levels of redistribution cannot barely implemented on a decentral level. See Olson 1999, Oates 1972.



intervention. Overlapping to the other dimensions is already visible so the decision on where to deal with which issues must be considered taking into account also the other dimensions.

### 3.3 Environment

The dimension of environment ties up to the concept of external effects. In principle the aim is to keep the environment within its potential, whereas in particular the capacity of regeneration of the environment and the minimization of irreversible damages as well as withdrawal of the environment are in the foreground. The German Enquete Kommission formulated management rules for material flows<sup>99</sup>. With regard to federal systems it needs to be tested, to what extent it is possible to identify the (spatially) bound systems of the environment and to find the appropriate level and units for the intervention in order protect the environment in the necessary way.

In practice, that means to identify the most important eco-systems as well as their spatial extension<sup>100</sup>. On that basis it seems to be useful to define the greatest common denominator. Figure 6 provides in the first two columns some examples for the delineation of areas on the basis of eco-systems.

Figure 6: political level relevant aspects of environmental sustainability

<b>Ecosystem</b>	<b>Spatial extension</b>	<b>Cooperation with regard to the eco-system</b>
<b>Water</b>	Rivers, lakes, seas	Rhine cooperation, Lake Konstanz cooperation, Baltic Sea
<b>Air</b>	World, regional	World climate conference, „Blue sky over the Ruhr“
<b>Natural and cultural landscapes</b>	Forests, mountain areas	Alpine convention
<b>Antropogene landscape i.n.s.</b>	Cities, rural areas	Municipalities and cooperations

<sup>99</sup> See Enquete Kommission 1997 S. 44ff, it is not undisputable, whether the management rules do suffice with respect to sustainability, but this is not the point. It is about, which consequences result out of the reflection of the environment. Minsch 1998, 21f.

<sup>100</sup> See for example Park 2001, 576ff.



The ideal case is described by so-called “Container regions”<sup>101</sup>, that are closed in some respect. It needs to be added, that the material flows of many eco-systems are not found yet, and that overlapping occurs in any case. Section 2.3. already pointed at out the fact, that we cannot approach reach the an ideal boundaries in any case, but we can try to approximate by making use of the flexibility of the federal system and the current administrative system.

### 3.4 Social affairs

The starting point for the social questions with regard to the sustainability is, which provisions the society has to make, in order to save the long-term survival of people on the world and within their jurisdiction. The most important basic conditions are peace and stability, as well as the protection of the individual freedom- latter at least according to democratic societies. Further aims arise, which influence the social stability. In democratic societies regarding the social dimension a fundamental consensus but different interpretations persists about values, such as human dignity, free development of personality and solidarity<sup>102</sup>. Finally it is a social decision, which aspects represent a contribution for the social stability of each society<sup>103</sup>. It has to be taken into account, that the social aspects should be judged with regard to their impact and possibility to influence them. Here again the interdependencies between the dimensions become obvious.

The reach of social peace and social stability in spatial terms depends factors such as economic integration, social-cultural point of contacts or historical developments. The intensity of the impact is depended on the gravity of problems, as well as of the spatial proximity. It affects both sides: the postulation of solidarity and the responsibility for solidarity. With regard to the federalism, it needs to be proved, how far spatially these tasks range in order to reach the above mentioned congruence. The arguments reach from the protection of human rights to the prevention of undesirable regional disparities such as high migration rates. The reference for the allocation of social tasks to regional authorities and levels is how functional context groups are organized.

### 3.5. Monitoring and governance

Up to now we were assuming that there is a kind of invisible hand, which guides all public institutions to follow a sustainable road. In fact the processes in order to find objectives and to implement a policy which allows an appropriate monitoring of such a kind of sustainable development is not such a clear task. Figure 7 can only give glance of the complex processes in that respect<sup>104</sup>.

In order to provide guidance for the design of appropriate decision and implementation processes some basic functions have been identified which are supportive for the achievement of sustainable development in particular in order to find the right balance between the thematic dimensions described above and to realised the regulative idea of sustainability defined above as the fourth dimension.

---

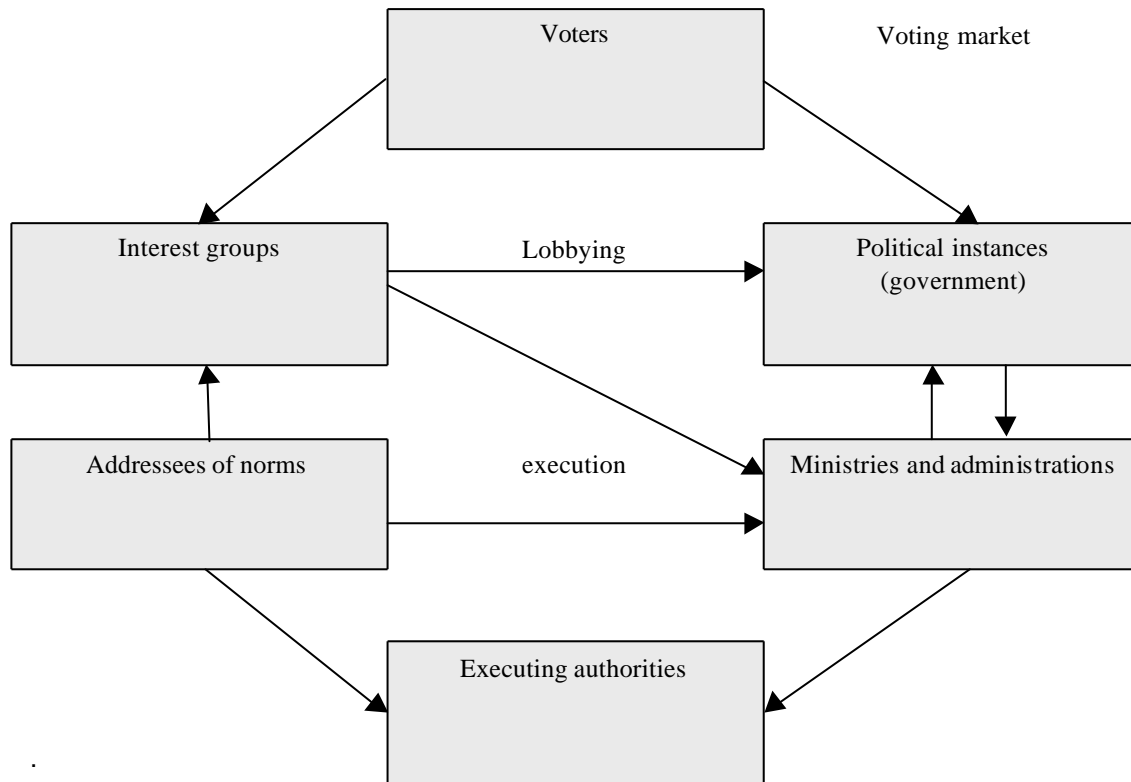
<sup>101</sup> Quoted in Läßle.

<sup>102</sup> See Enquete Kommission 1997, S.51; Minsch u.a. 1998, 20. Bizer 2000 tries to introduce the social dimension of sustainability via the consideration of social capital into the societal production function as side of human capital, real capital and natural capital.

<sup>103</sup> This can be shown by the selection of indicators. Example of the ISÖ: securing the existence of all members of the society, maintenance and development of the social resources, equity in accessing resources and – here another connection to the monitoring is obvious, participation in societal decision processes. Siehe <http://www.isoe.de/literat/kurzdp13.htm>.

<sup>104</sup> See Minsch u.a. 1998, 65 with reference to Gawel 1995, 33 und Endres/Finus 1996,43:

Figure 7: policy level relevant aspects of monitoring and governance of sustainability: actors and interactions



With regard to the interaction between the relevant actors shown in figure 7 the Enquete Kommission has proposed some institutional basic strategies in order to achieve the mentioned design-, searching- and learning process. The strategies manifest themselves in the foundation of respective institutions and have been summarized to an extensive institutional atlas:

- *Reflexivity* serves to increase knowledge of side-effects resulting from the action of actors in politics, economics and society<sup>105</sup>, which is important in that it reduces the orientation to and thereby the dependence of one-dimensional knowledge (experts) and short term orientation. These includes systems of information and counselling by satellite institutions, the improvement and structuring of information and decision making processes as well as a well focused research-, educational- and scientific policy.
- *“Participation and self-organisation contribute to the strengthening of the organisational potentials of the concerned groups of interests”*<sup>106</sup>, which is important as it might, due to the high differentiation of society, enable the differentiation of comprehensive institutions and thereby create new coalition and strategies of solution. These include alternatives of self-organisation of all subsystems of society and beyond its borders, rights of participation from the public to the direct democracy, as well as discursive participation models.

<sup>105</sup> See Minsch u.a. 1998, 143 Usually insufficient demand for information by voters is usually assumed regarding the risk assessment of environmental hazard. The reason for that grounds in the information paradox: the cost of additional information is much easier to assess than the utility.

<sup>106</sup> See Minsch u.a. 1998, 201 (accentuation by the author)



- “*Compensation- and conflict regulation*”<sup>107</sup> become more important as here borders can be crossed, which result from a lacking degree of organisation, insufficient participational rights and rights of decision making and other dependencies of the political-administrative system as well as other actors. These include advocacy institution in for sustainability within the existing institutions, resources – and power compensation, the opening of the processes of norm finding as well as integration strategies on an administrative level.
- “apart from the technical-economic *innovations*, a sustainable society especially relies on social and institutional innovations...”<sup>108</sup> which to that effect is important, as in this way new ways of solutions to cross blockades for a sustainable development can be found. These include new instruments of internalisation, strategies of co-operation, and strategies of information and national intermediary and private institution, informational strategies and supported strategies of the public households.

Now we can search for access points where federal system can support these kinds of functions and indeed there are several issues. Without going deeper into the subject at this stage, e.g. participation is support by the possibility to involve the affected and to exclude the not-affected groups by issues addressed under the thematic dimension, conflict management can be achieved by the monitoring of the higher level in case of disagreement on lower levels, or employment of the competitive process between the lower governmental units in order to simulate innovation processes.

#### 4. Conclusions

Federal systems consolidate the advantages of both unitary and purely independent states in several ways. One the one hand a task adequate allocation is possible, on the other hand the division of powers in a vertical respect can be warranted. In addition a common frame for the co-ordination and execution of societal tasks is available. Disadvantages become obvious, if there is no clear delineation of tasks between units and levels are defined consequently everyone adjudicates on everything.

The fundamental advantages provide starting points for the implementation of sustainability that is, in principle, not news. But it takes supplementary aspects into consideration, which up to now have played an inferior role. In the thematic respect three dimensions and aims, such as environment, economy and social have to be taken into account. With respect to the sustainability they need to be brought in line, especially in cases where conflicts of aims do exist. This takes place within the fourth dimension, the monitoring.

The vision of sustainability insists on an approach, which approximates the spatial circle of impacts monitoring in the above-mentioned thematic dimensions. Concerning the impact circle, examples for every single dimension are mentioned. With regard to the monitoring circle it has to be taken into account that sustainability as a regulative idea imposes new demands on the dimensions of monitoring by the society, which steps out the understanding of division of labour of public and private spheres. Sustainability challenges the design of the welfare state and invokes the potentials for self-organisation of the civil society<sup>109</sup>. Important contributions for the enforcement of sustainability can be obtained, if basic strategies as well as institutional implementation are take into consideration by making use of the potentials of federal states. It needs to be emphasized that some of these essential institutions are up to now available to some extent, which means that a federal system has already shown its advantages concerning innovations in some parts. The social learning-, searching- and design process within the context of federal states has to prove in practice, whether the potential can be put in value.

---

<sup>107</sup> See Minsch u.a. 1998, 264 (accentuation by the author)

<sup>108</sup> See Minsch u.a. 1998, 323 (accentuation by the author)

<sup>109</sup> In the end it is always a matter of the assessment of the society how the circle of impacts is delineated. See Postlep/Döring 1996, 13.



## References

- Beer S.H. 1977: A political scientist's view of fiscal federalism. In Oates W.E. (Hg.): The political economy of fiscal federalism. Lexington 21-46
- Benz A. u.a. 1998: Regionen und regionale Politik – eine Einführung. In: Benz, A.; Holtmann, E. (Hg.): Gestaltung regionaler Politik. Opladen, 15-32
- Blower A. 1993: The time for change. In: Blower A. (Ed.) Planning for sustainable environment. London. 1-18
- Enquete-Kommission "Schutz des Menschen und der Umwelt - Ziele und Rahmenbedingungen einer nachhaltig zukunftsverträglichen Entwicklung" des 13. Deutschen Bundestages: Konzept Nachhaltigkeit. Vom Leitbild zur Umsetzung. Hrsg.: Deutscher Bundestag, Referat Öffentlichkeitsarbeit, 1998
- Elazar, D.J. 1968: Federalism. In: International encyclopedia of social sciences. Vol. 5 353-367
- Elazar, D.J. 1995: Federalism. In: The encyclopedia of democracy. Washington Vol. 2, 474-482
- Endres, A.; Finus M. 1996: Umweltpolitische Zielbestimmung im Spannungsfeld gesellschaftlicher Interessengruppen. Ökonomische Theorie und Empirie. In: Siebert H. (Hg.) Elemente einer rationalen Umweltpolitik. Tübingen
- Eser, T.W. 1996: Ökonomische Theorie der Subsidiarität und Evaluation der Regionalpolitik. Baden-Baden
- Eser, T.W. 1997: Städtenetze - ein Instrument zur Bewältigung öffentlicher Aufgaben der Regionalentwicklung? In: Staatswissenschaften und Staatspraxis, No.3 S. 355-385
- European Commission (ed.): The EU Compendium of spatial planning systems and policies. Luxembourg 1997, S. 39.
- Fritsch u.a. 1991: Marktversagen und Wirtschaftspolitik. München
- Gawel, E. 1995: Staatliche Steuerung durch Umweltverwaltungsrecht. Eine ökonomische Analyse. Die Verwaltung 25, 201-224
- Kenyon, D.A. 1997: Theories of interjurisdictional competition. In: New England Economic Review. March/April.
- Kregel, B. 1998: Kommentar zum Beitrag von Arthur Benz: In: Benz, A.; Holtmann, E. (Hg.): Gestaltung regionaler Politik. Opladen, 123-126
- Kreibich R. (Hg.) 1996: Nachhaltige Entwicklung. Leitbild für die Zukunft von Wirtschaft und Gesellschaft. Weinheim/Basel
- Laufer, H. 1991: Das föderative System der Bundesrepublik Deutschland. München
- Minsch, J. u.a. 1998: Institutionelle Reformen für eine Politik der Nachhaltigkeit. Berlin u.a.
- Münch I. 1993: Staatsrecht I. Stuttgart u.a.
- Oates W.E. 1972: Fiscal federalism. New York .
- Olson M. 1969: The principle of fiscal equivalence. The American Economic Review, 59, 479-487
- Oates W. E. 1999: An essay on fiscal federalism. In Journal of Economic Literature, 37, 1120-1149
- Park, C. 2001: The environment – principles and application. London
- Postlep, R.-D. 1993: Gesamtwirtschaftliche Analyse kommunaler Finanzpolitik. Baden-Baden
- Postlep, R.-D., Döring T. 1996: Entwicklungen in der ökonomischen Föderalismusdiskussion und im föderativen System der Bundesrepublik Deutschland. In Postlep R.-D (Hg.): Aktuelle Fragen zum Föderalismus. Marburg, 7-45
- Sauerland, D.: Föderalismus zwischen Freiheit und Effizienz. Berlin 1997
- Scharpf, F.W. 1995: Optionen des Föderalismus in Deutschland und Europa. Frankfurt a.M.; New York
- Tiebout, C.M.: A pure theory of local expenditures. In: Journal of Political Economy. Jg. 64 (1956), S. 416-424.
- Wehling P. 1997: Sustainable Development – eine Provokation für die Soziologie? In: Brand K.-W. (Hg.): Nachhaltige Entwicklung: eine Herausforderung an die Soziologie. Opladen, 35-50
- Wiehler, F./Stumm, T.: The powers of regional and local authorities and their role in the European Union. In: European Planning Studies. Jg. 3 (1995), S. 227-250
- World Commission on Environment and Development 1987: Our Common Future. Oxford University Press.





**Markus E. Langer, Aloisia Schön**

## **AN INTEGRATED REFERENTIAL FRAMEWORK FOR SUSTAINABLE DEVELOPMENT**

*A project funded by the "Austrian Science Fund"*

### **Why a referential framework?**

Sustainable development is defined as progress towards objectives over which there exists consensus within a social group that they will ensure sustainability of the system under consideration. The objectives may vary between different social groups and may change over time. (Rao 2000) The term sustainable development is attributed a different meaning depending on the context and the objective. As a consequence of the diversity of definitions and objectives, there does not exist a generally applicable system of reference for the evaluation of sustainable development.

An important interface in the process of sustainable development lies between implementation and evaluation. Evaluation sponsors, the project or programme manager and the project team have a vision of sustainable development and a subjective perception of the process in question. For the evaluation, either an internal system of reference constituted by the project objectives and goals or an external system of reference is applied. Due to the great variety of theories and ethics of sustainability, the perception of the stakeholders involved may differ significantly. Similarly, there is a great choice available for the selection of an external system of reference. Clearly, the selection of the system of reference is critical for the quality of the evaluation.

Like a map for orientation, a distinct and comprehensive referential framework is a useful instrument for the transparent and accurate definition of scope and subject of the evaluation of sustainable development.

### **Evaluation with regard to what?**

#### **The role of evaluation for sustainable development**

Sustainable development is a social process that requires comprehensive considerations in decision making and a continuous reflection of the implications of human activity. In this context, evaluation plays a key role. In decision making, evaluations serve as an instrument for the integration of social, economic and environmental policies and for assuring the compatibility of programs on different political levels. During implementation, evaluations ensure the continuous improvement in social processes. In addition to that evaluations play an important role as a catalyst for institutional innovation. Given the wide scope of technical and process related aspects of sustainable development, the assessment and improvement of the process and its results require an evaluation that goes beyond the monitoring of indicators and financial controlling.

The role of evaluation for sustainable development depends on the stage in the cycle of political decision-making where it is applied. Considering the moment of evaluation, a broad distinction between prospective, ongoing and retrospective evaluations can be made (Chelimsky 1995, Bussmann 1997).

Prospective evaluation is an important field for sustainability evaluation. Decisive for political programs is the coherence of their elements. (Bussmann 1997) The evaluation at this stage of the policy cycle helps to identify discrepancies between objectives, operative elements and resources that may impair the desired effects of the program. Prospective evaluations focus on the system of political intervention, its context and its elements. Subject to the assessment are all framework conditions that are defined in the political planning process and will influence the implementation stage.

In the context of sustainable development, prospective evaluation can play an important role in identifying incompatibilities of goal systems between sectoral policies. Evaluations furthermore serve as an instrument to ensure and verify the integration of social, economic and environmental policies. Assessing the consistency of goal systems on different political levels is an evaluation task that supports governance.



Questions of political priorities and the assessment of alternatives are also subject to prospective evaluations. Finally, evaluations are commissioned in order to verify the adequacy of objectives and strategies with regard to a superposed political or scientific system of reference.

Ongoing evaluation, also referred to as process evaluation, is generally conducted in the course of the program or project with the objective of organisational learning and improvement through feedback. Process evaluation helps to identify and eliminate weaknesses in planning and implementation and to find potentials for further innovation. Generally, qualitative methods are applied in order to foster exchange of experience between actors. The target group of the evaluation are the protagonists themselves. The people involved expect from the evaluation an external point of view, an assessment of activities, suggestions for measures and an impetus for their further commitment. As a consequence, the innovative aspects are more prominent. In this context, ongoing evaluation plays an important role as a catalyst for institutional innovation.

In the case of activities for sustainable development, process evaluation provides information on whether the process is on the right track. Deviations are identified and analysed in order to assure the final achievement of objectives and requirements. Clearly, the system of reference used for the assessment of activities and progress is a central issue.

Performance evaluations, also referred to as retrospective evaluations, focus on the results of the evaluated subject including output, outcome and social impact. In many cases, performance evaluations are conducted at the end of the intervention in a summative manner.

The motivation for performance evaluations can often be found in reporting requirements and accountability. In this context, hypothesis testing is a frequently applied approach. Indicators are a means for assessing in quantitative terms whether targets have been achieved. Another criterion for the impact assessment is the effectiveness of the intervention which is defined as congruency of intended and observed behaviour of the target groups. (Bussmann 1997) Effectivity comprises both quantitative and qualitative aspects. Relationships between effects and causes are equally subject to performance evaluation.

In political practice, as it is the case for Local Agenda 21 processes, the questions of effects, impact and efficiency are of foremost interest. Currently a tendency from retrospective output evaluation towards process evaluation focusing on the organisation and protagonists of the system can be observed, which goes along with a shift towards a methodology developed in the auditing context. (Leeuw 1998) However, for the time being, evaluation as an instrument of ensuring political consistency and logical coherence (Bussmann 1997) is not so often applied.

### **Systems of reference – the case of sustainable development**

A core element of evaluations is the assessment of adequacy and progress relative to a set of objectives or criteria. Thus, as a prerequisite, evaluation requires a system of reference, with regard to which the subject of the evaluation, the *evaluandum*, will be assessed. From the perspective of the evaluated system, either an immanent or exmanent system of reference are applicable. Immanent frameworks of reference are deduced from objectives and tasks defined within the evaluated project or programme. Exmanent systems of reference are provided by external political or scientific sustainability concepts.

In practical experience, immanent systems of reference pose several problems. Often the goal system is not clearly defined at the outset of the project or programme. In other cases, the set of objectives only covers a narrow segment of the scope of sustainable development. Besides, information on the achievement of project goals and efficiency does not allow a general assessment of sustainable development.

If evaluations use external systems of reference, sustainable development is often interpreted in a reductionist way, frequently based on a set of indicators. However, a technical assessment of indicators can only provide insight into impacts, leaving aside information on the process. (Martinuzzi, Langer 2001) Besides, politicians, project managers and evaluators do not necessarily share the same perception of sustainable development and thus apply different systems of reference. The differences are hardly discussed at the outset of the evaluation. (Langer, Schön 2001)

However, the quality and accuracy of evaluations can be improved by discussing explicitly the elements and scope of the assessment system. On this basis, an unambiguous system of reference can be defined in a transparent and participative manner.

For this purpose, a distinct and comprehensive referential framework is a useful instrument for the transparent and accurate definition of scope and subject of the evaluation of sustainable development. Like a geographic



map it provides orientation within a greater context, it facilitates finding the focus of the process and the co-ordination of people involved.

Based upon literature analysis and expert workshops, the Research Focus Managing Sustainability has elaborated such a framework.

### Features and limits of the framework

The framework is an orientation tool to facilitate explicit discussion and informed decision making. The target groups of the framework comprise in particular evaluators, managers and institutions commissioning evaluations of sustainable development processes. It thus has to account in a comprehensive way for the different perspectives and approaches to sustainable development.

The framework accounts for the wide range of notions of sustainable development in a systematic way. Like a map, the proposed framework depicts the landscape of potential approaches to sustainable development and thus provides an overview of the aspects of sustainable development (scope) and the different requirements that are associated therewith (forms).

With the help of the framework, the scope, focus and level of complexity of concrete processes can be clearly determined and distinguished from alternative approaches. An explicit determination of the content of projects or processes may help to make the concept of sustainable development more tangible and operational. Similarly, a clear terminology also facilitates communication and co-ordination among protagonists and stakeholders. With regard to the implementation of sustainable development, the framework helps to match the demands and expectations on sustainable development that exist on different policy levels.

Without any assessment, the framework considers different forms to implement the aspects of sustainable development. On this basis, an informed choice for a specific form adequate for the given situation can be made within a project or process. As the framework points out perspectives for the improvement, extension and further development of the process, the framework can provide support in directing processes from a basic approach to more complex forms. Finally, the framework also provides a guideline for identifying approaches that may not be suitable to lead to sustainability.

For analytical purposes, profiles can be drawn from the framework that allow a comparison of approaches and perspectives. However, the framework does not provide a means for consistency testing. Incompatibilities between the choice of a certain form in one aspect and another aspect cannot be automatically identified. In this case the instrument relies on the discussion and participative decision making of the protagonists. Nevertheless, the framework facilitates these discussions.

### Structure of the framework

The structure of the framework follows the key questions that determine a process for sustainable development:

- Systemic aspects: time, space and material system boundaries
- Potential Aspects: What shall be sustained or achieved?
- Process design: How do we get there? What are the rules for decision making, implementation and improvement?

Along these three questions, the framework is structured in aspects and forms. The **aspects** determine the **scope and focus of a project or process** and thus allow a rough orientation like borders on a map.

Following the aspects, a basic match in scope of the concepts applied in planning, implementation and evaluation of the process can be determined. Further more the aspects allow the categorisation of approaches according to their focus on the social process or on impact and results.

<p><b>Systemic Aspects</b>  Integration of aspects especially the triangle of sustainable development  Temporal characteristics of sustainable development  Spatial characteristics of sustainable development</p> <hr/> <p><b>Potential Aspects</b>  Stocks and flows of potentials  Distribution and equity  Valuation of potentials</p> <hr/> <p><b>Process Aspects</b>  Social crosslinking/networking  Risk and dynamic change  Management of processes  Participation  Evaluation</p> <hr/> <p><b>Table 1: Aspects of the referential system</b></p>
--



In a particular project or process, certain aspects suggested in the framework may be attributed negligible relevance. However, the selection of the aspects should be the result of a transparent decision making process where all aspects are taken into consideration.

Whilst the aspects describe the scope of the process, the requirements or approaches associated therewith take on different forms of complexity. Here we refer to the forms as “basic”, “intermediate” and “complex”. In some cases also a “non sustainable” form can be determined.

After some analysis, an **informed choice** for a specific form adequate for the given situation can be made within a project or process. Such an approach also offers the possibility to develop a process into a well-determined direction by proceeding from a basic approach to more complex forms.

For analytical purposes, profiles can be drawn from the framework that allow a comparison of approaches and perspectives. However, the framework does not provide a means for **consistency testing**. Incompatibilities between the choice of a certain form in one aspect and another aspect cannot be automatically identified. In this case the instrument relies on the discussion and participative decision making of the protagonists. Nevertheless, the framework constitutes a helpful instrument to facilitate these discussions.

### Empirical evidence

Within the analysis of four cases of evaluations of local and regional initiatives of sustainable development, the framework described before, depicting the scope and form of approaches to sustainable development, was applied. The investigations show that in the observed cases the choice of certain aspects and forms of sustainable development is being made implicitly. However, such a choice is linked to a number of **shortcomings**:

- The choice is mostly done without specific knowledge of the aspects chosen and those not chosen
- The choice is not done explicitly
- The individual choices by different protagonists do not necessarily match

If an overall view of the cases is being made, it can be concluded that the full **integration of environmental, economic and social aspects** is not disputed by any of the actors within a sustainable development process. However, there is a more differentiated view concerning the **substitution** of resources and the limit up to which a resource might be exploited. For this aspect the protagonists adhere to either a basic, intermediate or even complex understanding of sustainable development. Concerning the **just allocation** of resources, the interviewees either think that equal opportunity is sufficient while there also is the more differentiated view that actual equality is important for sustainable development. The actors of these processes also see the field of action and influence mostly limited to the actual process. Only a minority of the active persons would consider effects beyond the process of sustainable development.

On the **process level**, there also exist rather differentiated views. Aspects with specific social background as participation and networking are considered to be aspects, which have to be regarded in great depth. The management of processes and the awareness and the management of risk are seen as less important.

It can be subsumed that there generally is some **common ground concerning sustainable development**. However, there are some differences in the general perception by the protagonists of local or regional processes of sustainable development, which aspect is part of sustainable development and which is not. There are even larger differences if the aspects of sustainable development are differentiated in levels of complexity – their forms of sustainable development. The fact that personal views of sustainable development differ considerably between project managers and evaluators raises the question of which aspects and forms will be applied in the process of sustainable development and its evaluation.

### What are the implications for evaluations?

There are considerably different views of the aspects involved in a process of local sustainable development. The protagonists of the processes of sustainable development are not explicitly dealing with some of the divergences of these views. Therefore there are **discrepancies** concerning the question which aspect and form of sustainable development is an issue for the process of sustainable development.

The results may imply **inefficiency and non-accuracy** as a result of insufficient coordination between the evaluator and the project manager and further research has to be done in this respect. On the other hand, evaluations may also play an important role in pointing out issues not covered by the process.



## Conclusion

In the context of sustainable development, **evaluations have particularly high relevance** as evaluations enable continuous improvement in social processes. They serve as an instrument for the integration of social, economic and environmental policies and assure compatibility of processes on different political levels. Evaluations play an important role as a catalyst for institutional innovation.

The interface between the different groups of protagonists involved in planning, implementation and evaluation of sustainable development have a particular need for establishing a common ground to ensure consistent and efficient implementation and accurate evaluation.

The investigation of four case studies in Austria have shown that the **choice of certain aspects** of sustainable development is being made implicitly. Among stakeholders of local and regional initiatives of sustainable developments, there is a general but heterogeneous knowledge of potential attributes of sustainable development. A more detailed analysis with regard to the forms of the aspects of sustainable development reveals even greater differences.

In addition to the underlying personal views, perceptions of the evaluated process and the content and priorities of the evaluation also vary greatly. Differences in the importance attributed to aspects of sustainable development by evaluation sponsors and evaluators may impair the effectiveness and accuracy of the evaluation. Even though the interviewed evaluation sponsors generally expressed their satisfaction with the evaluations, in the cases under scrutiny the recommendations have hardly led to action.

Empirical evidence emphasizes the necessity to establish a common **referential framework of sustainable development**. Such a referential framework is especially important for the definition of scope and subject of the evaluation of programmes, processes and projects for sustainable development.

**Further research** has to be conducted to reveal additional details concerning the conduct of evaluations and the methodology for the evaluation of processes for sustainable development to advance towards a comprehensive model of the evaluation of sustainable development.



## Annex

### The referential framework – checklist

#### **Sustainable development is...**

- **Integrating system characteristics of the natural and the socio-economic system**
  1. A unitary approach to sustainable development requires the integrated analysis of the complex structure of effects in a functional and holistic and interwoven way.
  2. A focused approach is characterised by the fact that the goal system comprises of environmental, social and economic aspects, but not necessarily in a balanced way
  3. A partial approach focuses on a singular system or its constituents but also considers some feedback mechanisms from outside the system under scrutiny
  
- **Considering the different time spans and time lags of systems and system interactions**
  1. The time horizon can be determined according to the duration of projects, programs or planned interventions related to a process
  2. The time horizon can be set according to the duration of effects caused by the respective project or processes, which often last beyond the period of interventions within the project as such
  3. The time-horizon can be chosen according to the system influenced by the decision. The adequate time horizon would be consistent with the time these systems need to react and to regenerate
  
- **Considering spatial interlinkages of project, processes, programs and policies**
  1. The spatial boundaries are set to define the immediate area of activity of a project, program, process or an organisation
  2. The delimitation is including more than just the immediate area of activity as relevant interrelations with other systems outside are also considered
  3. The actors of a sustainable development process decide to include far reaching interrelations also on the global level
  
- **Considering material system boundaries**
  1. The partial approach to sustainable development focuses on a singular system or its constituents, for instance natural resources, but also considers feedback mechanisms.
  2. A focused approach to sustainable development is characterised by the fact that priorities are attributed among social, economic and environmental objectives.
  3. With a unitary approach all relevant facets of development including environmental, social and economic issues are considered in a functional and holistic way including the complex structure of effects
  
- **Defining levels and limits of what shall be sustained**
  1. Monetary limits focus on the monetary value of capital, resources and potentials that shall be sustained.
  2. Physical limits and levels, often expressed as safe minimum standards, are independent of monetary valuation but rely upon scientific knowledge of cause and effect relationships.
  3. Qualitative standards describe the desired state of the social, economic and natural environment in quantitative and qualitative terms and reflect both scientific evidence and ethic elements.



- **Valuing and measuring social, economic and environmental resources and capacities**
  1. As market prices alone have proved insufficient to assure the sustainable use of resources and capacities, shadow prices are introduced to value and measure goods that do not underlie market rules.
  2. In addition to use-values expressed as market and shadow prices, the concepts of total economic value also includes option and existence values, thus accounting for a long term perspective.
  3. Beyond the anthropocentric perspective, allowance can be made to so-called intrinsic or inherent values. These values are independent of the function for human beings, they are directly related to the subject in question.
  
- **Organizing the distribution of resources and income in an equitable way**
  1. The potential for compensation is the only assessment criterion.
  2. Equal opportunities are regarded as the basis for equity, taking into account that people with different preferences and values need different framework conditions to achieve their objectives.
  3. Actual compensation is demanded in order to diminish social inequalities in material terms.
  
- **Able to manage projects, processes and programs**
  1. Specific management-structures like steering committees for projects and processes of sustainable development are in place, which are paying attention to basic management instruments and support the creativity of the actors
  2. Complex management-structures like steering committees together with work-groups are in place. These structures are then employing multiple methods of process and project management and support creativity and innovation
  3. Broad, efficient and innovation-oriented institutions and management structures are in place. These are then utilizing comprehensive and especially quality-oriented methods of process and project management
  
- **Supporting societal crosslinking and networking**
  1. Networking and the exchange of information is aiming at a win-win situation, but the benefits derived are confined to a small group of directly involved actors who also form a primarily single discipline information background
  2. The interests and background of multiple disciplines or stakeholders are being recognized and an interdisciplinary or multi-faceted approach is being chosen to benefit people, initiatives or organizations not directly involved. The knowledge and experience from other disciplines is being integrated
  3. An integrated, transdisciplinary and systemic approach, allowing for broad solutions and paths is being chosen. Stakeholder-chains are being extended, and even indirectly involved entities are taken into account, leading to broad solutions
  
- **Dealing with dynamic change and risk**
  1. The ability to react to change and the interaction with risk is primarily confined to dealing with issues of dread
  2. Changes and possible paths of development are and are prepared. Dynamic change, uncertainty, risk and dread are being anticipated and are included in the development of possible scenarios and feasible ways to react
  3. The precautionary principle is being followed to deal with potential risks. The avoidance of irreversible situations, the minimization of risk and participation are in the foreground



- **Supporting participation and responsibility to come to new forms of governance**
  1. The use of democratic instruments as well as the free market ensure basic participation
  2. The ability to participate actively at case-specific decision making as well as implementation processes is ensured and is going beyond traditional democratic or market instruments
  3. The ability to participate freely and extensively at any project and process is guaranteed to single citizens as well as interest groups
  
- **Based on a reflexive process facilitating constant learning**
  1. The objectives and goals are being quantified. The attainment of the goals can be measured by using indicators
  2. The quantitative as well as qualitative aspects are being evaluated if originally set goals or guiding visions were reached
  3. Evaluation is being considered to be an integral part of a process of sustainable development. The objective of the evaluation is organizational learning and the improvement of the process





## References

- BAUEN Rebecca; BAKER, Bryan; JOHNSON, Kirk: "Sustainable Community Checklist", Washington, 1996
- BORN, Manfred (Hrsg. Econtur): "Handlungsleitfaden zur Entwicklung eines kommunalen Nachhaltigkeitsindikatorensystems im Rahmen der lokalen Agenda 21", Bremen, 1997
- BORTZ, Jürgen; DÖRING, Nicola: "Forschungsmethoden und Evaluation für Sozialwissenschaftler", Berlin, Heidelberg, New York, 1995
- BOSEL, Hartmut. „Indicators for Sustainable Development - Theory, Method, Applications", International Institute for Sustainable Development, Winnipeg 1999
- BOSEL, Hartmut: "Earth at Crossroads- Paths to a sustainable future", Cambridge, 1998
- BRAND, Karl-Werner (Hrsg.): "Nachhaltige Entwicklung- Eine Herausforderung an die Soziologie", Opladen, 1997
- BREUEL, Birgit (Hrsg.): "Agenda 21- Vision: Nachhaltige Entwicklung", Frankfurt, New York, 1999
- BUSSMANN, Werner: "Evaluationen staatlicher Maßnahmen erfolgreich begleiten und nutzen: Ein Leitfaden", Chur, Zürich, 1995
- CHELIMSKY, Eleanor; SHADISH, William R. (Hrsg.): "Evaluation for the 21st Century, A handbook", London, New Delhi, 1997
- COSTANZA, Robert and Patten, Bernard C.: "Defining and Predicting Sustainability". Ecological Economics 1995. Vol. 15 pages: 193-196
- COSTANZA, Robert; CUMBERLAND, John; DALY, Herman; GOODLAND, Robert; NORGAARD, Richard: "An Introduction to Ecological Economics", Boca Raton, Florida, 1997
- DEPT. OF FINANCE, Commonwealth of Australia: "Doing Evaluations, A practical guide", Canberra, 1994
- ENQUETE KOMMISSION: "Schutz des Menschen und der Umwelt - Endbericht", Deutscher Bundestag, Bonn 1999
- Europäische Städte und Gemeinden auf dem Weg zur Nachhaltigkeit: "Aalborg Charta- Charta der Europäischen Städte und Gemeinden auf dem Weg zur Nachhaltigkeit", Aalborg, 1994
- EUROPÄISCHES PARLAMENT - GD Wissenschaft-Arbeitsdokument: "Die Praxis der Evaluation bei den Strukturfonds der europäischen Union", Luxemburg, 1997
- FAUCHEUX, Sylvie; O'CONNOR, Martin (Hrsg.): "Valuation for Sustainable Development- Methods and Policy Indicators", Cheltenham, 1998
- HARDI, Peter and ZDAN, Terrence: Assessing Sustainable Development: Principles in Practice, Winnipeg, IISD 1997
- HARDI, Peter; ZDAN, Terrence (Hrsg.): "Assessing Sustainable Development- Principles in Practice", Winnipeg, 1997
- HOLMBERG, J., ROBERT, K.-H. "The System Conditions for Sustainability", Göteborg, 1997
- HOLZINGER; Elisabeth (Hrsg.): "Programm Evaluation- Patentrezept für Effizienz?", Wien, 1994
- HUCHLER, Elisabeth; MARTINUZZI, André: "ÖkoBusinessPlan Wien- Erster Evaluationsbericht Juni 1999", Wien, 1999
- HUCHLER, Elisabeth; MARTINUZZI, André: "Ökoprofit Dornbirn, Erfahrungen, Erfolge, Übertragbarkeit", Wien, 1997
- HUCHLER, Elisabeth; MARTINUZZI, André: "Ökoprofit Dornbirn, Langfrist- Evaluation des Jahrganges 1996", Wien, 1998
- HUCHLER, Elisabeth; MARTINUZZI, André; STÖRMER, Eckhard: "Evaluation des ersten Jahrganges von ÖKOPROFIT München", Wien, 2000
- JÄNICKE, Martin, CARIUS, Alexander, JÖRGENS Helge: „Nationale Umweltpläne in ausgewählten Industrieländern", Berlin [u.a.], 1997
- JÄNICKE, Martin, KUNIG, Philip, STITZEL, Michael: „Lern- und Arbeitsbuch Umweltpolitik: Politik, Recht und Management des Umweltschutzes in Staat und Unternehmen", Bonn, 1999
- JÄNICKE, Martin: „Umweltplanung im internationalen Vergleich: Strategien der Nachhaltigkeit", Berlin [u.a.], 2000
- JÄNICKE, Martin: „Umweltpolitik im internationalen Vergleich", Berlin, 1995
- JÄNICKE, Martin; Kunig, Philip; Stitzel, Michael: „Lern- und Arbeitsbuch Umweltpolitik: Politik, Recht und Management des Umweltschutzes in Staat und Unternehmen“, Bonn 1999
- JÜDES, U.: "Culture of Sustainability". Tagungsvortrag. Wien, Juni 2001



- KLAASSEN, Ger A.J. & Opschoor, Johannes: "Economics of sustainability or the sustainability of economics: different paradigms", in: Ecological Economics, Elsevier 1991, 4, pp. 93-115
- KUFFNER, Astrid: „Evaluation von Nachhaltigkeitsaspekten - nachhaltige Evaluation? Eine vergleichende Analyse von deutschsprachigen Evaluationsberichten die sich mit Nachhaltiger Entwicklung beschäftigen", Wien 2000
- LANGER, Markus, „Nachhaltige Entwicklung – Information, Materialien, Methoden“, Österreichische Gesellschaft für Natur- und Umweltschutz, Wien 1999
- MARTINUZZI, André, HAMMERSCHMID, Gerhard, HUCHLER, Elisabeth: "ÖkoBusinessPlan Wien- Zweiter Evaluationsbericht Februar 2000", Wien, 2000
- MICHELSEN, Gerd, FISCHER, Andreas. "Wissen über Nachhaltige Entwicklung" in: Campus Courier, Institut für Umweltkommunikation an der Universität Lüneburg, Nr. 1, 1999
- MINSCH, Jürg, EBERLE, Armin, MEIER, Berhard, SCHNEIDEWIND, Uwe: „Mut zum ökologischen Umbau: Innovationsstrategien für Unternehmen, Politik und Akteurnetze", Basel [u.a.], 1996
- MINSCH, Jürg: „Institutionelle Reformen für eine Politik der Nachhaltigkeit", Berlin [u.a.], 1998
- MINSCH, Jürg: „Ökologische Grobsteuerung, konzeptionelle Grundlagen und Konkretisierungsschritte: Agenda für eine nachhaltige Entwicklung in der Schweiz", St. Gallen, 1994
- NEUMAYER, Eric: Weak versus strong sustainability. Exploring the limits of two opposing paradigms, Northampton, Edward Elgar 1999
- PEARCE, D.W. and G.D. Atkinson: "Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability". Ecological Economics. 1993, 8:103-108.
- PEARCE, David; Barbier, Edward and Markandya, Anil: "Blueprint for a Green Economy", London, Earthscan Publications 1990
- PEARCE, David; Barbier, Edward et al.: "Blueprint 2: Greening the World Economy", London, Earthscan Publications 1991
- PEARCE, David; YOUNG, Carlos: "Measuring Sustainable Development- Macroeconomics and the Environment", Cheltenham, 1997
- PEZZEY, J.: "Economic Analysis of Sustainable Growth and Sustainable Development". Environment Department Working Paper No. 15. Washington, D.C.: The World Bank. 1989
- PEZZEY, J.: Sustainable development concepts: an economic analysis, World Bank, Washington 1992
- PFISTER, G. RENN, O. "Die Studie 'Zukunftsfähiges Deutschland' des Wuppertal-Institutes im Vergleich zum Nachhaltigkeitskonzept der Akademie für Technikfolgenabschätzung", Arbeitsbericht Nr. 75/Juni 1997, Baden-Württemberg 1997
- POWER, Michael: "The Audit Society- Rituals of Verification", Oxford, New York, 1997
- RAO, Pinninti Krishna (2000): "Sustainable development: economics and policy", Malden and Oxford
- ROSSI, Peter, H.; FREEMAN, Howard, E.: "Evaluation- A systematic Approach", Newbury Park, London, New Delhi, 1993
- SANDERS, James R. [Hrsg.], BEYWL, Wolfgang [Bearb.]: Handbuch der Evaluationsstandards - die Standards des Joint Committee on Standards for Educational Evaluation", Opladen, 1999
- SERAGELDIN, Ismael: "Sustainability and the Wealth of Nations. First Steps in an Ongoing Journey", Environmentally Sustainable Development Studies and Monographs Series No. 5, Washington, The World Bank 1996
- TNS / The Natural Step, Holmberg, J., Robert, K.-H. "The System Conditions for Sustainability", Chalmers University, Göteborg, Sweden, 1997
- UNCED - United Nations Conference on Environment and Development: "Agenda 21: Programme for Action for Sustainable Development", New York , 1992
- WCED (World Commission on Environment and Development): "Our Common Future". Oxford: Oxford University Press 1987
- WIDMER, Thomas: "Meta Evaluation- Kriterien zur Bewertung von Evaluationen", Bern, Stuttgart, Wien, 1996
- WIDMER, Thomas; ROTHMAYR, Christine; SERDÜLT, Uwe: "Kurz und gut? Qualität und Effizienz von Kurzevaluationen", Chur, Zürich, 1996



**FIELDS OF APPLICATION**

**Maria-Angeles Diez, Spain**

Current Approaches used in the Evaluation of Regional Policies: a Critical Review

**Alain Thierstein, Switzerland**

An Evaluation Framework for the Quality Management of Projects on Sustainable Development

**Diane Warburton, United Kingdom**

Evaluating Participation in Sustainable Development

**Clive George, United Kingdom**

Sustainability Impact Assessment

**Kate Theobald, United Kingdom**

LASALA: Self-assessment of Local Agenda 21 for Social Authorities



**Maria Angeles Diez**

**CURRENT APPROACHES USED IN THE EVALUATION OF REGIONAL POLICIES:  
A CRITICAL REVIEW<sup>110</sup>**

**1. Introduction**

Theories and policies for regional development have undergone sharp transformations over the last few years. At present, concepts like 'endogenous development' have been relegated to the sidelines by new theoretical approaches that are more complex and sophisticated: the language they use is of learning regions, areas that are able to stand up for themselves and adapt to the new competitive conditions imposed by globalisation. From this perspective, new regional development policies emerge which devote much of their attention to networks of inter-company co-operation (clusters) and regional innovation systems. These new development policies are focused on the social capital of the region and include in their strategy social, institutional and cultural elements.

In fact, a new agenda has begun to spread throughout many European countries and regions, where the social capital is seen as a new competitive advantage in regional development terms: building knowledge networks, learning mechanisms and formal and informal norms and rules, which can foster reciprocal understanding and mutual confidence among the agents of the regional economy, as a necessary complementary asset to the economic factors traditionally thought to influence economic development. At a practical level, what this means is that regional policy must favour the creation of a framework for interaction: a framework in which companies, organizations and public agencies are able to explore joint solutions to problems shared in common, a framework in which, once dialogue has been established, attitudes are more receptive to the interchange of information and interactive learning are generated (Amin & Hausner 1997; Cooke & Morgan 1998; Storper 1997).

But, the important question here is, how are we going to evaluate these regional development policies? What approach are we going to use? In order to answer this question, I have revised, first the principal characteristics of these policies and the main challenges posed by their evaluation, trying to put forward a number of programmatic elements that, theoretically, could help to improve evaluation practice. As I will argue, classical evaluation models do not adapt to the specific peculiarities of modern regional policies and, besides, their capacity to produce and facilitate the accumulation of knowledge and learning is limited. Second, I have tried to find out how these challenges have been solved in the work done so far and ascertain whether any of the previously mentioned programmatic elements have been introduced yet into regional policy throughout a research review of present evaluation studies of EU Structural Funds interventions.

**2. Why are new regional policies so difficult to evaluate?**

First, from a theoretical point of view, we can identify some specific characteristics of these new regional policies: the presence of intangible, many-faceted and wide-reaching objectives; the complexity of cause-effect relationships; their systemic nature and embeddedness, their dynamism and flexibility, and, finally, the devolution of powers to the region. A detailed re-examination of these elements makes it possible to point to the most problematic areas when undertaking their evaluation. These difficulties can be transformed in methodological and political challenges for evaluation work and lead us towards a search for possible solutions and new approaches, as can be seen in Table 1.

The first characteristic mentioned is the presence of intangible objectives. This means that the aim of these policies is the creation of knowledge, learning and capacity-building, both at a personal and at a collective level. How can these effects be observed and measured?. They may be revealed, not only via economic impact on well-defined concepts such as competitiveness and/or other indicators connected with market exploitation, but also through non-economic effects on the behaviour of companies, institutions and regional organisations.

---

<sup>110</sup> An extended version of this paper will be published in the Journal *Evaluation* in 2002.



So, there are many difficulties involved in quantifying effects and identifying measurement indicators for them. In fact, these effects are not easily reducible to a single criteria, represented by a monetary cost-benefit ratio, as required by cost-effectiveness analyses and other models that have been traditionally applied for the evaluation of regional development policies. To solve this challenge, I propose the use of an approach that includes qualitative information, to capture and observe the intangible effects, paying special attention to the observation of the organizational and cultural changes that arise from these new policies.

**Table 1: Matching regional policies to evaluation approaches**

<i><b>Characteristics</b></i>	<i><b>Challenges</b></i>	<i><b>Evaluating proposals</b></i>
The objectives of the policy are the creation of knowledge, learning and capacity building	Well-defined objectives do not exist and there are numerous difficulties in quantifying effects and identifying measuring indicators	Qualitative information is the most suitable and useful tool for estimating the effects of individual and institutional learning
Innovation is a complex interactive process where continuous feedback is produced	There is no linear causal relationship between resources, activities, results, effects and regional impact	What is needed is a holistic approach and the application of naturalistic, qualitative and interrogative techniques
Systemic nature: at a vertical and horizontal level	Complex interactions are produced between the different regional subsystems and effects at different levels: companies, institutions, regional community	Case studies as a method of observation and analysis
The policies are firmly rooted in their context and embedded in their socio-economic framework	It is necessary to know and understand the cultural and political context in which the evaluation develops	Social, cultural and political elements are an integral component of the evaluation. Evaluation is a socio-political process
Innovation policies are dynamic processes where continuous interactions are produced	Evaluation must be an active-reactive-adaptive process in relation to changes in conditions (context) and the needs of stakeholders	Evaluation design must be dynamic and flexible
Policies are designed via a bottom-up approach and with the active participation of all the regional actors	Evaluation must be opened up to the different actors involved and must recognise the existence of a pluralist society	The participation of the actors involved must guide the evaluating design. Evaluation is a collective learning process

The second characteristic is that these regional policies usually involve complex interactions in which a great number of agents and organisations take part and where continuous feedback loops are produced. As a consequence, it is not possible to identify a linear cause-effect model between inputs and activities, on the one hand, and the results and effects, on the other hand, as needed and demanded, again, by traditional evaluation models, whether they be experimental designs or econometric models. Rather than a linear relationship, we have complexity and circularity. Nevertheless, there is a need to understand the logic behind the functioning of these policies and to improve our knowledge of the working mechanisms. So, one solution might be to use a holistic approach and/or introduce other proposals such as realistic evaluation (Pawson & Tilley 1997) and theory-based evaluation (Weiss 1995) in order to comprehend how these complex mechanisms work and why.



The third characteristic is their systemic nature, acting simultaneously on the different components of the regional system of innovation and/or the regional productive clusters, and favouring the creation of links for horizontal cooperation. These policies do not only affect companies and do not only have an impact on the macro-economic indicators of the regions. Interactions are produced within a full range of contexts: between firms and regional social infrastructure, between producers and users at an inter-firm level and within the wider institutional milieu. In order to identify all these thick interactions, it is helpful to use case studies as an evaluation method, due to their capacity to pinpoint information and contribute to understanding the phenomena that develop in complex contexts.

In fourth, I mentioned the element that I have called 'embeddedness'. This is related to the idea that these policies are firmly rooted in their social and cultural context and spring from and evolve within their own regional situation: identifying the socio-economic conditions, the needs of companies and the region, the political and institutional context within which regional policies take shape, with the aim of achieving a single political design. Each regional development policy is unique as they are specific responses to particular problems (Saraceno 1999). However, classical evaluation models are characterised by the fact that they overlook the economic and social context in which the policy unfolds. The economic impact, the central axis of the evaluation, is estimated, whilst leaving to one side the local framework. On the contrary, I believe that evaluation ought to be more contextualised. The approach should be to integrate exogenous factors of a cultural, social and political nature within the evaluation, instead of attempting to control their effects, to take them out of the evaluation and/or deal with them as 'confounding factors'. Evaluation must be understood as a social-political process.

The fifth characteristic I have mentioned is that these policies look for the development of dynamic processes in the region. The process of learning is dynamic and continuous interactions take place along with flows of knowledge, resources and human capital, which evolve over time. As a consequence, the evaluation must also be capable of evolving, that is, flexible at adapting to changes in the policy under evaluation and in the changing conditions in which this regional policy unfolds. Evaluation must be dynamic, active, reactive, and adaptive in relation to the changing context and also in relation to users' need for the evaluation. Evaluation itself must be understood as a developmental process. Moreover, given the innovative nature of these new development policies, we are not aware of important elements regarding the implementation process itself, this is, the way in which the results are produced. It will, therefore, be vital to evaluate the real process. The design of the evaluation must be capable of breaking down the traditional division between summative and formative evaluation and stimulate approaches whereby both types of evaluation form part of the same exercise.

Our last characteristic refers to the active role the region plays, a new role that implies, on the one hand, a devolution of powers and responsibilities to the region, and, on the other hand, the introduction of new forms of designing regional policy based on bottom-up approaches and on the active participation of all the regional actors. What I am emphasising is that these new policies are not necessarily policies FOR the companies or for the technological centres and/or for the productive clusters. The intention is to design these policies WITH the active involvement of all these regional agents. And, opening up the policy design to companies, to intermediary organizations, to trade unions, ... supposes a radical change in the ways of making regional development policy, introducing dialogue and communication directed towards consensus and helped by democratic representational structures in the regional institutions themselves. So, at this more political level, the design of the evaluation must be guided by the participation of all the actors involved and must be opened to all the stakeholders (Stamme 1999; Kuhlmann 1998). Evaluation must be transformed into an open process of collective learning in order to contribute by helping regional governments to set increasingly better policies in motion. Moreover, the evaluation should serve as a useful tool for mobilizing communities for action. This implies the recognition that policies are developed within a pluralist society and introduces participatory evaluation as an approach capable of generating new knowledge.

### **3. Reviewing the practice: 6 case studies**

Now, I am going to turn my attention to evaluation practice in an attempt to find out how these challenges have been met and ascertain how these regional policies are being evaluated. This involves a review of six evaluations, all of which have been recently requested by the European Commission in relation to Structural Funds.



These evaluations are: evaluation of Business and Innovation Centres (BIC), evaluation of Regional Technological Plans (RTP), evaluation of Regional Innovation Strategies (RIS), thematic evaluation of the impact of the Structural Funds on Research, Technology, Development and Innovation (RTDI) in Objective 2, thematic evaluation of the impact of the Structural Funds on Research, Technology, Development and Innovation (RTDI) in Objective 1 and 6, and finally, the thematic evaluation of the impact of Structural Funds on SMEs. They are all good clear exponents of the most up-to-date practices in evaluation and provide us with an accurate picture of the 'state of the art'.

This review has served to corroborate that the practice of regional policy evaluation has been gradually incorporating some of the elements presented above. So, in the first place, it has been possible to confirm the important impact that regional policy is registering outside the economic sphere, in the form of social, institutional and cultural effects which have been observed thanks to the use of qualitative techniques and data. The use of qualitative information and techniques has allowed evaluators to begin to appreciate the changes taking place among regional agents of innovation systems and productive clusters, in their patterns of behaviour, and in their interactions with other regional actors. Similarly, the evaluation teams have started to ascertain where the results of these changes are leading towards the building of networks and consensus, the acquisition of new knowledge and skills, the reinforcement of mutual trust, the beginning of learning processes and institutional reflexivity.

Secondly, the problems that crop up when establishing linear cause-effect relations have been verified. The difficulties in meeting this challenge occasionally lead to evaluations abandoning the job of attempting to understand the functioning mechanisms of these policies. They did not focus the evaluation on the question of how and why these policies work, because they are not able to reproduce the cause and the effects. Some of the evaluations identify the best practices as a way of understanding the policies and use case studies to gather information concerning these complex realities. Here, case studies have been proved a very useful tool for identifying good practice in innovation and action policies supporting SMEs. This information makes a direct contribution towards enhancing evaluators' and managers' knowledge and help them to reach specific findings and recommendations directed towards improving public interventions. However, stress should be made on the need to develop and construct an overall theory of action (Finne *et al.* 1995) since evaluation must explore and understand the interaction between policies and the net effects of the intervention (Georghiou 1998)

Thirdly, I have been able to discover how evaluations try to resolve the systemic nature of many of these policies: the analysis and observation of data have been tackled from a top-down (the impact on macro-economic indicators) and a bottom-up approach (the impact on micro-economic indicators). This focus leads to the core of evaluation efforts being concentrated upon the two extremes of the chain of effects, that is, upon the region as a whole, on the one hand, and upon the companies and the people as the prime beneficiaries, on the other hand. Consequently, observation of changes produced at the intermediate levels get relegated to a secondary position: we don't know anything about what happens inside this space, between the region and the final beneficiaries (in the clusters, in the networks, in local and regional institutions and their inter-connections). Here, efforts have to be made to go beyond counting the numbers of firms and people served in various ways, trying also to identify and assess the added value that springs from multiple firm partnership and institutional building (organisational learning, creation of networks, cooperative behaviour, ...). These are added benefits that affect the competitiveness of an industry and the development of the whole region as well.

Fourthly, we have pointed out that these regional policies are embedded. Here, evaluations use case studies for making more contextual evaluations, where exogenous factors and the local conditions in which policies unfold are elements that are integrated in the evaluation. Fifthly, one innovation that has been gradually introduced is the inclusion of process assessment, along with an estimate of effects and impact, thus producing a combination of formative and summative evaluation exercise.

Finally, it has to be pointed out that there is no evidence of the introduction of participatory evaluations. The idea of opening evaluation to all the people involved in the policies, all the stakeholders, has not been taken into account. Policies continue to be evaluated from one angle, from the point of view of the managers who define and implement these new regional policies. So, it is noteworthy that policies that seek to introduce new ways of making regional policy, opening them up to the active involvement of regional agents, do not try to do the same in their evaluation exercises. Such an approach would involve the direct participation of business, local and regional staff and service providers who can express their own opinions and make the evaluation process their own.



To comment briefly on the methods used, all the evaluations under review include both quantitative and qualitative information and therefore plump for the combined use of a variety of evaluation methods. In general, there is a clear predominance of qualitative approaches over quantitative ones, and also of case studies, in a double sense, both as a means of information gathering through fieldwork and as a method for analysing and interpreting primary information. However, there are two evaluations that have concentrated on measuring the regional impact using quantitative indicators.. These are the thematic evaluation of the impact of Structural Funds on SMEs and the evaluation of Business and Innovation Centres (BICs). Both of them employ predominantly statistical techniques to handle and analyse the data observed. They are looking for quantified effects and impacts, significant correlations between variables, patterns of regional development and/or similar trajectories. The objective of both evaluations is to quantify the creation of employment, the number of jobs created through the policy.

The other evaluations reviewed use case studies in order to illustrate the effects on the beneficiaries, to show the impact on the regional economy, reveal patterns of behaviour and functioning, to identify best practices and draw lessons. They try to observe and analyse changes produced in the process and also intangible effects. Some of the non-economic effects identified are: an increase in innovation culture, the mobilisation of local actors, the creation of co-operation networks, public-private partnership, the introduction of participatory approaches to policy design, the reinforcement of systemic links, changes in the regional political agenda and steps taken towards institutional learning. Furthermore, in the end, it is precisely this non-economic impact of new regional development policies on social capital and institutional thickness which appears as the most significant impact pointed to in all the evaluations reviewed. So that, even in the evaluations with a more quantitative approach mentioned above, it is recognised that regional policies should not be judged only by their economic results, since "non-economic impacts are more important than purely economic ones" in the words of the evaluators. They also maintain that the most significant effects of these policies are produced on the social, cultural and institutional spheres.

It can thus be seen that a significant change has taken place in the practice of evaluating regional policy, with a shift towards approaches closer to the pluralistic paradigm. But, here an important Paradox arises since the choice of this qualitative approach did not stem from a previous selection process in which the advantages and drawbacks of the different evaluation models had been analysed, nor from a decision based on the degree to which the approach adopted fits the needs and characteristics of the policies. On the contrary, the qualitative approach was used, to quote the evaluation teams, as "the only available option", because of the difficulties involved in using any other evaluation model (such as quasi-experimental designs and/or statistical or econometric evaluations) to estimate the quantitative effects on beneficiaries and the region.

Consequently, evaluators have been forced to adopt qualitative evaluations, despite the fact that they are seen as a second option which in no way constitutes the best evaluation model to be used. Thus, for example, in the thematic evaluation of RTDI Objective 2 the evaluators held that "as the effects on the regional economy cannot be measured, case studies are used to illustrate the type of impact that these interventions may have on the regional economy". But, at the same time, they recommend that in future evaluations, once current deficiencies are overcome, value for money models be adopted, aimed at estimating the net cost per job created. Suggestions in the same direction can be also found in the RTP and RIS evaluation.

In short, these statements unquestionably point to the persistence of two underlying assumptions behind all these evaluation designs. First, the idea that there is one methodological evaluation model superior to all others. Second, this model of excellence turns out to be, precisely, the classical evaluation model based on quasi-experimental designs or value for money studies, and as such is always preferable, whatever the case under study, to the remaining available options. This means, therefore, that within European Structural Funds evaluation culture, what still dominates is the classical model based on quantitative approaches as superior methodological evaluation designs.





#### 4. Some proposals and recommendations

To conclude, I would like to put forward some proposals and recommendations that I think we should bear in mind when designing an evaluation of regional development policies. Firstly, each regional policy must be evaluated using the design and the methods that best adapt to its specificities and the needs of evaluation users. So, each evaluation situation requires a unique and specific evaluation design; a design that is appropriate for a specific and particular action or policy-making context (a combination of people, policy, history, socio-economic conditions, resources, values, needs, interests and opportunities) (Patton 1997). Secondly, there is no single evaluation model amongst all the possible options that can serve as a methodological recipe applicable to each and every evaluation. There is no ideal methodological design, nor do either superior or inferior evaluation methods exist. As the European Commission itself recognises: "The experience of all these years of evaluation demonstrates that a universal evaluation method does not exist" (Commission Européenne 1999). Thirdly, it is necessary to accept that the traditional debate between quantitative and qualitative evaluators has today been overcome. Evaluation is a complex task in which evaluators must have the freedom to adopt the models that prove most suitable for each evaluation. There are even approaches arguing for the use of multiple techniques within the same evaluation as a valid option offering a great potential (Greene & Caracelli 1997). Thus, the combination of techniques (triangulation) comes on the scene as an open field of work and a delicate methodological approach, but there is every chance it will become a powerful tool for evaluation in the future. In this sense, it is vital to defend the view that the final quality of an evaluation does not depend upon the methods used, but rather on the usefulness of the information produced and its capacity for offering up valuable answers. But, as we have seen, many evaluators and users of the evaluation still hold onto the remains of a closed vision in which quantitative methods, techniques and data continue to enjoy greater credibility and value than qualitative ones.

At a more operative level, I would like to say that the pluralistic paradigm presents itself as the most suitable evaluation approach for tackling evaluation of regional development policies. In fact, many of the difficulties and challenges thrown up by these policies receive a better response within this pluralistic approach which appear to suit, in a more natural way, the specific characteristics of these policies and the objectives pursued by regional governments through their design and evaluation.

#### References

- Amin, A. & Hausner, J. (Eds.) (1997): *Beyond market and hierarchy. Interactive governance and social context.* Edward Elgar.
- Autio, E. (1998): "Evaluation of RTD in Regional Systems of Innovation", *European Planning Studies*, Vol. 6. N° 2, pp. 131-140.
- Commission Européenne (1999): *Evaluer les programmes socio-économiques.* Vol. 1. Office des Publications Officielles des Communautés Européennes. Luxembourg.
- Cooke, P. & Morgan, K. (1998): *The associational economy. Firms, regions and innovation.* Oxford University Press.
- Finne, H. et al. (1995): "Trailing Research. A model for useful program evaluation", *Evaluation* Vol. 1 N° 1, pp. 11-31.
- Georgiou, L. (1998): "Issues on the evaluation of innovation and technology policy", *Evaluation* Vol. 4 N° 1, pp. 37-51.
- Green, J. C. & Caracelli, V. J. Eds. (1997): *Advances in mixed-method evaluation: The challenge and benefits of integrating diverse paradigms.* *New Directions for Evaluation*, N° 74. Jossey-Bass Publishers.
- Kuhlmann, S. (1998): "Moderation of policy making?. Science and Technology Policy evaluation beyond impact measurement-the case of Germany", *Evaluation*, Vol. 4. N° 2, pp. 130-148.
- Patton, M. Q. (1997): *Utilization-Focused Evaluation.* 3<sup>a</sup> Edition. SAGE Publications.
- Pawson, R. and Tilley, N. (1997): *Realistic Evaluation.* SAGE Publications.
- Saraceno, E. (1999). "The evaluation of local policy making in Europe", *Evaluation*, Vol. 5. n° 4, pp. 439-457.
- Stame, N. (1999): "Small and medium enterprise aid programs: intangible effects and evaluation practice", *Evaluation and Program Planning*, Vol. 22. pp. 105-111.
- Storper, M. (1997): *The regional world. Territorial development in a global economy.* The Guilford Press. New York.
- Weiss, C. H. (1995): "Nothing as practical as good theory: Exploring Theory-Based evaluation for Comprehensive Community Initiatives for children and families" en *New approaches to evaluating Community Initiatives: concepts, methods and contexts.* Eds. Connell, J. P., et al. The Aspen Institute. Washington, D. C. pp. 65-92.



**Alain Thierstein, Manfred Walser**

## **AN EVALUATION FRAMEWORK FOR THE QUALITY MANAGEMENT OF PROJECTS ON SUSTAINABLE DEVELOPMENT**

### **1. Introduction**

The Rio Declaration on Environment and Development clearly demonstrates that the concept of sustainable development is a normative, global and regulative idea. Sustainable development is a multi-dimensional concept that deals with different levels of action (Thierstein, Walser 1997). Even the final report of the "Brundtland-Commission on Environment and Development" focused on different local and cultural perceptions so as to propose a global problem-solving scheme. Furthermore, many of the 27 principles of the Rio Declaration refer to the fact that in addition to the global level of regulations, both the national and community level play a crucial role in the proper and effective implementation of the concept (UNCED 1992a).

Chapter 28 of Agenda 21 states: "Because so many of the problems and solutions being addressed by Agenda 21 have their roots in local activities, the participation and cooperation of local authorities will be a determining factor in fulfilling its objectives. Local authorities construct, operate and maintain economic, social and environmental infrastructure, oversee planning processes, establish local environmental policies and regulations, and assist in implementing national and subnational environmental policies. As the level of governance closest to the people, they play a vital role in educating, mobilising and responding to the public to promote sustainable development." (UNCED 1992b). So far it is indisputable: sustainable development must be regarded as a political concept. This raises the need for its perception as a political process which in turn raises questions like the following:

- How and by whom are the objectives for specific spatial entities being set within the general framework of the Rio Declaration and Agenda 21?
- Which authorities are delivering means and producing legitimation? Who are their partners in cooperation and at which stage of the implementation process does this cooperation take place?
- What is the relationship between public and private bodies with regard to the implementation of sustainable development policies and respective measures?
- Who is in charge of evaluating and controlling sustainable development projects and programmes?

These questions need to be answered as a precondition for the implementation of the concept of sustainable development at regional and local level. The concept must prompt a "down-to-earth" approach; it must get a handle on day-to-day political activities with clear-cut, legitimate and evaluative questions. Thus, the aim of this paper is to present a method for the evaluation of both sustainable development projects on a regional or urban scale and the initial experiences with such an approach which demonstrate its pros and cons, strengths and weaknesses.

### **2. The key question: how can the concept of sustainable development be implemented?**

The implementation of the concept of sustainable development faces three fundamental challenges:

- The issues at stake: The concept of sustainability tends to be adapted to existing structures and constraints and implementation is often limited to material issues such as the use of resources, social distributive justice and economic development. However, these basic dimensions are interrelated and the main challenges for sustainable development arise at the interface between the three dimensions. Science, in particular, will tell us how it has to be done and how much degrees of freedom of action is left. But beyond that, we need specific transformation knowledge to translate the scientific findings into action.
- The temporal horizon: Time is scarce; this is what scientists familiar with greenhouse effects and global climate change tell us. Demographic patterns change slowly but have fundamental impacts on consumption and investment behaviour. Thus, the question arises as to which issues are most urgent with respect to time and how do they combine with other issues.



- Spatial delimitation: Although sustainable development is a global and normative concept, it will primarily be implemented by nation states and their sub-national levels. Again the question arises as to which sustainability issue will be dealt with at which spatial level. We also need to identify the specific tasks that are best dealt with at local, regional, national and even global level.

The preamble of Agenda 21 offers some valuable hints for an effective process of implementing the concept of sustainable development. Section 1-3 of the preamble starts as follows: "Agenda 21 (...) reflects a global consensus and political commitment at the highest level on development and environment cooperation. Its successful implementation is first and foremost the responsibility of Governments. National strategies, plans, policies and processes are crucial in achieving this. International cooperation should support and supplement such national efforts (...). Other international, regional and subregional organisations are also called upon to contribute to this effort. The broadest public participation and the active involvement of the non-governmental organisations and other groups should also be encouraged".

([gopher://gopher.undp.org:70/00/unconfs/UNCED/English/a21\\_01](http://gopher://gopher.undp.org:70/00/unconfs/UNCED/English/a21_01)).

Thus, the preamble mentions three important requirements for any attempt to implement sustainability policies:

- the selection of the appropriate spatial and institutional level;
- allowing the broad public to participate;
- involving non-governmental organisations and other groups.

The next section sheds some light on one of the appropriate levels of implementation of sustainable development: regions and city-regions, which may both encompass rural, urban and intermediate spaces.

### 3. Implementing sustainable development in regions and city-regions

The implementation process at an intermediate level between national and local level involves taking advantage of four functions of space (Fürst 1993), 302):

- The region is a resource space. The classical impact of resources (minerals, water etc.) loses its meaning through globalisation and technical innovation. Increasingly, the ability to mould one's own living space takes centre stage, including endogenous development potential and soft localisation factors.
- The region gains in importance as a level for public policies. The trend towards the regionalisation of sectoral policies and the necessity for a transparent arena for a mediating planning process are behind this change.
- The region within its institutional boundary is distinguished as a promoting power for identification, because nowadays affiliation is also politically defined. The commons is normally easier to manage because the level of a community is more transparent.
- Even when openly defined, the region is beginning to be regarded as an arena for networks of actors from different social areas that depend on trust and high levels of communication.

The dependence on national and international framework conditions shows that implementation of sustainability policies cannot be achieved solely within the responsibility of a region or a city-region. However, if the existence of a 'local Agenda 21' in many cities is important for implementing sustainability, then collaboration and coordination between localities is essential. This could be achieved through the development of a regional framework of sustainable development policies (Patterson et al. 1995), 776). Firstly, as an area of identity and actor interaction, the region is manageable – a factor which may assist in legitimising concerted action and which increases the likelihood of societal coherence. Secondly the regional level is probably better at bringing together fragmented regional interests and avoiding a strict limitation to local interests, in other words 'parish-pump politics'.

Bearing in mind the need to implement the global concept of sustainable development at sub-global levels and taking into consideration the above mentioned challenges, we would like to present a concept for evaluating projects and programmes that aim to foster sustainable development. In order to create such an evaluation tool, we first had to enlarge and enhance the basic three-dimensional concept by accounting for the political process and ethics.



#### 4. Enlarging the concept of sustainable development

Ethics and politics are often neglected in the sustainable development debate. While much discussion focuses on the ethical principles of sustainability, when it comes to concrete implementation, they constitute a kind of “black box” located on a different problem level. The sustainability concept is normally translated into the three dimensions of ecology, economy and social fabric.

The ecological dimension is easy to grasp using threshold values and legal regulations, for example. In general, the economic dimension in general does not pose major difficulties as it can be expressed not only using a main indicator such as Gross Domestic Product (GDP), but also using various quantitative goals and key figures. However, the social fabric dimension is often handled as a ‘remaining quantity’: it can describe a combination of factors such as social class, poverty, education, health, participation, culture, ethics etc. To achieve greater consistency for the implementation of the concept of sustainable development, the immaterial and procedural objectives should be separated from material aspects of social fabric.

If the fields of ethics and processes are introduced as additional dimensions to the three commonly identified sectors of ecology, social matters and economy, the concept becomes more complex. At the same time, however, there is also greater transparency. Thus, all of the fields can be combined to give the “Ten Dimensions of Sustainable Development” (Schleicher- Tappeser et al. 1997; Thierstein, Walser 2000) (see figure 1). Our evaluation framework is based on these ten dimensions:

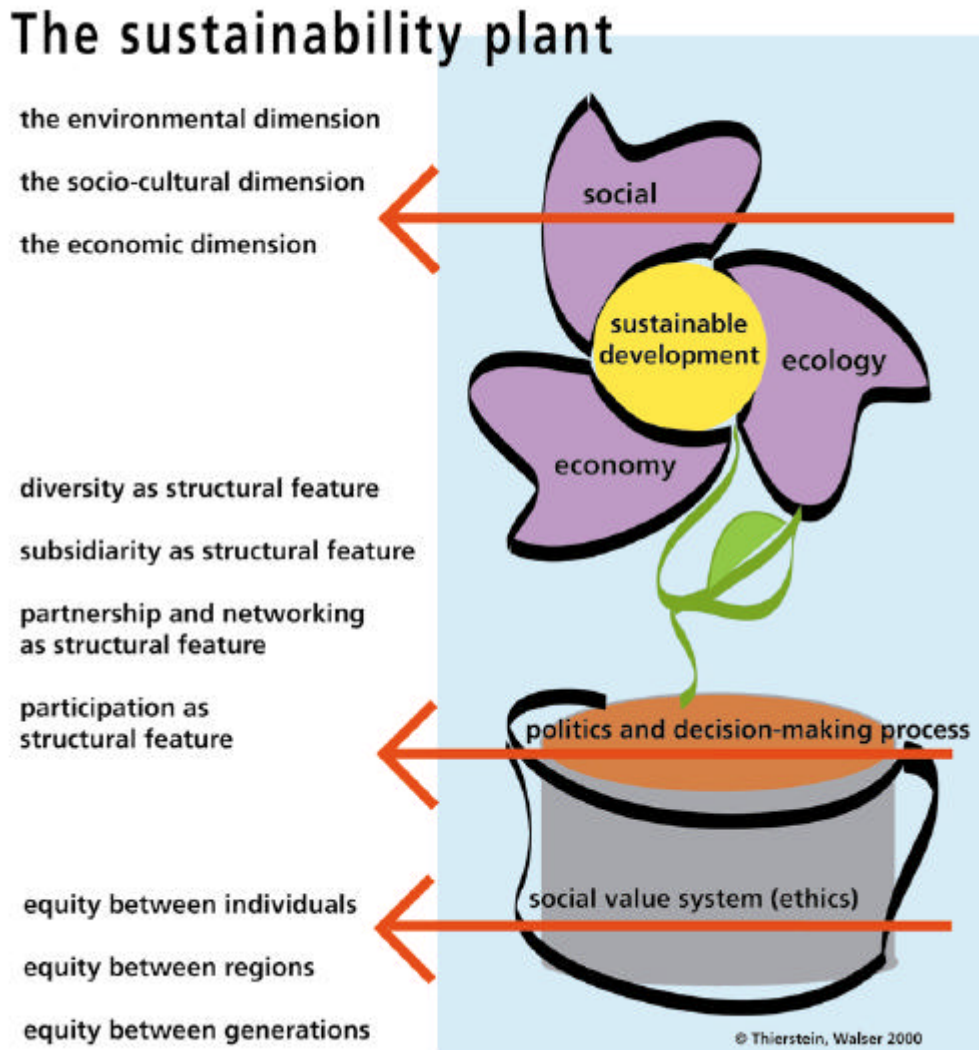
- The three material sectors are the tangible elements of sustainability. These are the petals of a flower, the eye-catcher.
- The decision-making processes are the soil in which the plant grows and from which it draws the sustenance that allows it to bloom. The political system and political culture are part of this. They define who takes, controls and implements decisions. Thanks to the nutrient content of the soil, the plant can grow and develop.
- Finally, the social value system which is based on ethics, is the flowerpot that holds it all together (the plant in question is a pot plant). In democratic states, at least, a political decision that fundamentally contradicts social values is inconceivable.

There are many possible solutions to global threats. They range from eco-dictatorship to a vision of a highly automated state in which the functions of nature are almost entirely replaced by machines. This is only the spectrum of visions that have been developed in our cultural sphere – other cultures and religions often produce completely different attitudes to life, the environment and personal development. In the short and medium term, however, the only political decisions that are conceivable for us are those located within our own value system. Thus, the plant cannot break out of its pot: at best, while growing, it may let some of its roots follow the course of the water and grow out of the pot.

With this “sustainability plant” we can serve the social dimension within the concept of sustainability better: in other words, we take into account not just the objective – the project – but also the way it is achieved, that is the process.



Figure 1: The sustainability plant



## 5. Developing an evaluation tool

Building on these conceptual foundations, our assessment tool comprises ten dimensions that interpret the five main elements of sustainability. The ten dimensions serve primarily as an ex-ante evaluation tool that helps to improve the overall quality of a given project or political programme. The conceptual framework is translated into an evaluation checklist which expresses the ten dimensions in the form of a number of specific questions. The purpose of this checklist is to systematise the discussion of the character and general impact of a given project or programme and to help in the approximate estimation of its level of sustainability.

It is important to recognise that our tool only reaches its full potential when understood as a discursive means for the exchange of opinions and facts on a specific project. The tool is intended for use as a checklist which raises awareness of the depth and breadth of the concept of sustainable development. The evaluation tool is intended to help its users to be as *comprehensive* as possible in their exchange of information and opinions regarding the project under discussion. In other words, our evaluation tool aids in the verification of the comprehensiveness of all of the relevant aspects of sustainable development in a given project. Taken as an ex-ante assessment, the checklist can also be used as a management tool for improving the quality of certain sustainability dimensions that achieve a lower score than other dimensions.



The significant advantage of this tool based on the ten dimensions of sustainability is that it can be used to assess projects and programmes in the form of an ex-ante-evaluation. The three principles of equity, as shown in the lower part of figure 1 above, can be used to assess the long-term consequences of a project by posing the following question: Does this project or measure serve justice? In contrast, the short and medium-term primary objectives of a project, programme or measure are evaluated using the three fields of ecology, social fabric and economy. Combined with the three material elements and the four structural features for the evaluation of organisations, the resulting ten elements facilitate simple assessments as to whether a planned project is pursuing a sustainable path or rather not.

While developing this evaluation tool we were fully aware that a useful evaluation tool must guarantee a maximum level of transparency. Thus, we tried to select questions and indicators that are (1) limited in number, (2) relevant, (3) responsive, (4) simple and (5) policy-relevant.

This assessment process is implemented as follows:

- Projects – be they private or public – should be evaluated prior to being implemented. At this stage, the relevant stakeholders who are involved in or concerned with a specific project are called together to jointly develop and share their assessment about that same project. Both expert knowledge and the experience of laypersons based on the three main fields of ecology, economy and social fabric is processed by three groups focusing on “targets”, “processes” and “ethics”. Each group must provide an answer to the following central evaluative question: Do the objectives of the project in question have a positive or negative impact, or is the impact neutral? In the evaluation, a positive impact is given one plus point, a negative impact one minus point and neutral impact is awarded zero.
- In order to reach a valid answer to the questions or issues posed in the assessment, the group must exchange knowledge and information based on data or experience openly and thoroughly. By doing so, the group will gradually gain an overall impression of the specific sustainability aspects of the project in question.
- By following the order of the checklist, the discursive process of project assessment deals first with more short-term impacts on the level of objectives and eventually culminates with the consideration of the long-term impacts by evaluating the equity dimensions of sustainable development.
- The entire assessment project is time consuming, but it eventually leads to a comprehensive evaluation of the project in question. However, it is unlikely that the result of this assessment will be unanimous.
- Nevertheless, the detailed checklist will enable the identification of certain basic flaws and weaknesses in the project. Quality improvements and compensation measures may then be discussed with the project promoters and funding bodies.

	Field/Criterion	-1	0	+1
<b>1.</b>	<b>Economy</b>			
1.1	Supports regional economic cycles			
1.2	Supports high quality jobs, new work forms and models			
1.3	Supports operational participation processes and the social quality of the world of work			
1.4	Targets innovative products and processes			
1.5	Supports attractiveness of location, an innovative milieu, the networking of economic actors			
1.6	Creates the same economic benefits with reduced consumption of resources (eco-efficiency)			
1.7	Improves the financial situation of public and/or private budgets			
1.8	Minimises material goods traffic			
	<b>Total (minus points subtracted from plus points)</b>			
<b>2.</b>	<b>Ecology</b>			
2.1	Supports near natural landscape and/or near natural elements in urban areas			
2.2	Reduces land use and/or soil sealing			
2.3	Reduces the input of pollutants into air/water/soil			



2.4	Reduces noise			
2.5	Reduces the consumption of (non-renewable) energy			
2.6	Reduces the consumption of (non-renewable) resources, support of material cycles			
2.7	Supports bio-diversity and habitats			
2.8	Supports environment-friendly forms of transport and reduces superfluous mobility			
	<b>Total (minus points subtracted from plus points)</b>			
<b>3.</b>	<b>Society and social matters</b>			
3.1	Contributes to security of material livelihood			
3.2	Supports communication and co-existence in the population			
3.3	Increases quality of living conditions and quality of life			
3.4	Supports the educational offer and opportunities for all residents (empowerment)			
3.5	Supports consciousness-raising and behavioural change towards sustainability			
3.6	Supports cultural, social and regional identity			
3.7	Aids health and supports health care			
3.8	Supports the feeling of safety (avoidance of hazards in personal environment)			
	<b>Total (minus points subtracted from plus points)</b>			
<b>4.</b>	<b>Forms of participation</b>			
4.1	Transparency and accessibility			
4.2	Support of weaker or badly organised interests			
4.3	Consensus orientation and discussion of strategic issues (vision)			

	Field/Criterion	-1	0	+1
<b>5.</b>	<b>Forms of networking</b>			
5.1	Support of dialogue and cooperation between different groups			
5.2	Inclusion of many people and/or many different interests			
5.3	Creation of space and/or media for the exchange of information			
<b>6.</b>	<b>Subsidiarity in interaction between different institutional levels</b>			
6.1	Clearly defined decision-making competencies on different institutional levels			
6.2	Creative freedom at work level			
6.3	Support from higher institutional levels			
<b>7.</b>	<b>Diversity in interaction between different approaches</b>			
7.1	Equal cooperation and decision-making competencies for actors from public and private institutions			
7.2	Combination of different approaches and solutions			
7.3	Cooperation between different perspectives and cultures			
	<b>Total</b>			
<b>8.</b>	<b>Equity between groups</b>			
8.1	Support of social cohesion			
8.2	Support of material balance			



8.3	Support of the integration of disadvantaged groups			
8.4	Support of equity between men and women and/or gender mainstreaming			
<b>9.</b>	<b>Equity between regions</b>			
9.1	No transfer of burdens to the disadvantage of neighbouring regions			
9.2	No transfer of burdens to the disadvantage of other regions world-wide			
9.3	Support of disadvantaged regions through structural changes at home			
9.4	On-site support of poor and disadvantaged regions			
<b>10.</b>	<b>Equity between generations</b>			
10.1	Reduction in the consumption of non-renewable resources of all kinds			
10.2	Reduction in public and private debt			
10.3	Participation of children with decision-making powers			
10.4	Discussion and dissemination of a vision of a sustainable future			
	<b>Total</b>			

The assessment of the five main elements of sustainability is summed up in the following box:

Field	Number of points (total; equal weight)
Economy	
Ecology	
Social matters	
Formation and implementation process	
Ethical bases	
<b>Total</b>	

The next section describes our initial experiences with the assessment tool.

## 6. Experiences with the ten-dimension evaluation tool

Our ten-dimension evaluation tool has already been applied in some projects and the following section briefly describes the circumstances of its application in two cases.

The *first* case was the 'International Conference of the Lake Constance region' (*IBK*), which established a regional Agenda 21 initiative in the year 1999 (URL: <http://www.regio-bodensee.net/agenda/>). This intergovernmental body is organising an annual competition for sustainable development projects or project ideas in this cross-border region. The competition committee used our tool for the 2001 competition, to which 122 projects at varying stages of completion were submitted. The competition committee was composed of six distinguished experts on sustainable development originating from all four member-countries of the *IBK*: Germany, Austria, Switzerland and the Principality of Liechtenstein.

The *second* opportunity for the use of our tool arose in early 2002 when students on the post-graduate course in spatial planning at the Swiss Federal Institute of Technology in Zurich conducted a real-case ex-ante evaluation of a large-scale infrastructure project within the agglomeration of Zurich. Four groups of students from various professional backgrounds assessed a planned light rail transportation infrastructure project intended to link the urban area between the north-eastern periphery of the city of Zurich with the international airport of Zurich-Kloten (URL: <http://www.vbg.ch/glattalbahn/>).





## 7. Lessons learnt and preliminary conclusions

The experience gained so far with our ten-dimension evaluation tool is limited but nonetheless instructive; the authors of this paper both were involved in the two cases described above. In both cases, the evaluating groups using the assessment tool had to become acquainted with a methodology that initially seemed rather odd. In both cases, within a short period of time, the evaluating groups adopted the tool with an open mind, irrespective of whether they were experts in the field, as in the case of the Lake Constance project, or laypersons as in the second case.

Having assessed the process of the adoption and application of our tool, we have reached the following conclusions.

### The strengths of the methodology

- Firstly, our assessment tool is centred round the *participation* of stakeholders. It is they who are likely to implement the concept of sustainable development in the long run. Although – as was the case in the first example described above – the tool is used by sustainability experts, the methodology urges the users to incorporate a broad perspective on the possible outcomes of a project.
- Secondly, our ten-dimension evaluation tool is not a top-down expert approach based on sophisticated technology and indicators. It is a *discursive*, easily grasped and nonetheless systematic method that can be applied in various real-case situations involving the implementation of public or private projects.
- Thirdly, the methodology works well even if there is relatively little *information* available up front. In both of the cases in which the tool has already been applied, the promoters of the projects under discussion were able to provide initial and necessary information. The quality of the assessment results from the combining of "objective" factual information with implicit knowledge based on professional and personal experience.
- Fourthly, our tool appears to be more useful for *ex-ante evaluation* and monitoring, which means the involvement of stakeholders and citizens, than for an ex-post impact assessment. This is mainly due to difficulties in the correct delimitation of time, space and issues. This leads us to the limits of our methodology.

### Flaws and limits of the methodology

- Firstly, the two cases in which our evaluation tool was applied showed that the checklist needs to be accompanied by some additional information, such as a 'manual' on how to handle or implement the tool. For example, it clearly emerged that when using the tool for a large-scale infrastructure project, the evaluators must be aware of the *spatial delimitation* of the project in question. The light rail project, for example, does not only impact in terms of mobility and noise etc. on the actual ground on which it is built and in the immediate areas it serves. A large-scale infrastructure project of this nature must be regarded as a spatial development project that has far reaching consequences, or *outcomes* over time, for example, for the modal split of transportation, the structures of the urban economy, the prices of real estate and the labour market.
- Secondly, the same problem arises with regard to *temporal reach*. When examining the consequences for the time horizon, current evaluations regularly distinguish three moments in time: the output of a project, the impact and the outcome. Thus, as chiefly a discursive instrument for checking the comprehensiveness of measures our ten-dimension evaluation tool — will not produce an easy consensus with regard to temporal consequences because the people involved in assessing a planned project will have greater problems evaluating its impacts the further into the future these impacts accrue.
- Thirdly and interestingly, the regional Agenda 21 case demonstrated that the use of experts on sustainability is an advantage in the ex-ante assessment of projects. The small group was first instructed on how to use the evaluation tool properly for the 122 projects submitted. Having assessed a certain number of projects, the evaluators began to cut our rather lengthy assessment list of questions and criteria short. The experts then developed a kind of shorthand procedure to win time in the evaluation of the 122 projects submitted. During an intermediate break in the course of the evaluation, discussion among the experts showed that they had made rather similar assessments of each project using their shorthand method as compared with the use of the systematic and comprehensive check list. This finding does not mean that an elaborate evaluation tool is not necessary. It probably reflects the experts' high level of



personal experience and knowledge, which again reinforces the quality of our evaluation tool. In contrast, the second real case showed that the post-graduate students used the full version of the questions on the checklist and their evaluation was, therefore, more time-consuming. The students came to the conclusion that every item in our checklist prompted an equally helpful level of discussion about the sustainability impacts of the light rail transportation system. Thus, they were also forced to develop an agreement on how strictly they should apply the evaluation tool. On the other hand, however, the overall results of the evaluations carried out by the four student groups were quite similar and indicated, therefore, that our tool can achieve a satisfactory level of performance.

### **Preliminary conclusions**

Our own preliminary conclusion on the use of our ten-dimension tool for the evaluation of sustainable development projects is thus largely promising and positive. Two main aspects dominate in our preliminary conclusions:

Firstly, the checklist has proven to be a well-developed and well-balanced instrument for verifying the comprehensiveness of all of the relevant sustainability dimensions of a project. Needless to say, there is always room for improvement on individual questions and criteria. However, the overall impression prevails that our standardised tool covers a wide variety of project characteristics that can present for evaluation. Nevertheless, certain limitations remain which call for further experimentation and research.

Secondly, having gained some experience with the application of our evaluation tool, the question remains as to who should apply our method and in which context it should be applied. So far we believe that our evaluation tool, which is centred round a participative and discursive approach, is suitable for use by both professionals who deal with sustainability questions and laypersons who are motivated and involved, but do not necessarily possess an expert status. Thus, our tool is primarily intended for people and experts who work for regional development agencies or are involved in urban development and regeneration processes, for example dealing with the reuse of brownfield sites as in the case of 'Zurich West', the city's old industrial quarter (URL: [http://www.stzh.ch/fste/pro\\_impulsgruppe.htm](http://www.stzh.ch/fste/pro_impulsgruppe.htm)) (Glanzmann, Thierstein 2002). Some experiences with 'Local Agenda 21' processes show that the use of our tool in the assessment of the sustainability of such projects would necessitate the development of a tutorial by external experts. Otherwise, local actors tend to interpret or misuse the tool in favour of their own objectives in order to reach a positive assessment for 'their' project.

## **8. Looking forward and future research**

The two main improvements that need to be made to our assessment tool involve: (1) the management of the evaluation tool itself and: (2) the need for the integration of an evaluation module that covers the process of implementation of a project.

### **Managing the evaluation tool**

As already indicated above, the application of our evaluation tool requires additional information on how to handle the instrument in a specific situation. A short manual needs to be developed with a description of a '*standard operating procedure*' (SOP) for the tool to ensure that maximum advantage is made of the tool.

- Firstly, the introduction of the SOP should stress the tool's approach as a discursive instrument that requires a fairly high standard in terms of the quality of the discussion process.
- Secondly, the SOP needs to be aware of and discuss in detail the spatial and temporal delimitation of the impacts of the project under discussion.
- Thirdly, the SOP should mention links to available indicator systems for sustainable development. It is well understood that our discursive evaluation tool cannot be used without indicators that provide information on the state, development and response of sustainability aspects that are relevant for the evaluation of a specific public or private project.
- Fourthly, the SOP must clearly state that all projects evaluated must comply with legal and other regulations as a minimum requirement for passing the 'sustainability test'. Where available, corresponding indicators should be used and followed.



## Evaluating the implementation process

Our sustainability approach with the ten-dimension evaluation tool places great emphasis on the implementation process and on participation. These facts mean that it is necessary not only to evaluate the 'product' – that is the private or public project – but also the 'process', that is the manner a specific project is designed, implemented, monitored and – in most cases – replaced by a follow-up activity.

In principle, our ten-dimension evaluation tool accounts for these procedural aspects, which include 'participation' (element number 4 in our checklist), 'networking' (number 5), 'subsidiarity' (number 6) and 'diversity of approaches' (number 7). However, we need to go beyond that and develop a well-founded indicator system for assessing the implementation process. Adequate process indicators are needed in order to assess the quality of a project's implementation process in its entirety. The main elements of such process indicators must deal with:

- the process of the formulation of the objectives relevant to a specific project;
- the composition of the project leaders in terms of the representation of the main stakeholders;
- the degree of (formalised) legitimation of the proposed project;
- the degree of public transparency of the proposed project.

To our knowledge such an indicator system has not yet fully emerged but first elements are showing up – see for example the University of Applied Sciences in Erfurt where process indicators were proposed to evaluate the process of local Agenda 21 (URL: <http://www.fh-erfurt.de/vt/>).

Today, the concept of sustainable development displays the same characteristics of a regulative idea as those displayed by the concept of human rights over two hundred years ago. Its implementation takes time, knowledge and expertise. In 2002, we are only at a stage ten years after the global launch of the concept in Rio de Janeiro. The concept needs not only a thorough discussion on objectives and indicators but also on the management process of its implementation. It is our hope that our evaluation tool can make a small contribution to that long-term endeavour.

## References:

- Fürst, D. 1993: Raum - die politikwissenschaftliche Sicht. In: Staatswissenschaften und Staatspraxis. Heft 2. S. 293-315.
- Glanzmann, J.; Thierstein, A. 2002: Zürich West. Nachhaltige Entwicklung auf Quartierebene. Discussion Paper des Fachbereichs Raumordnung No. 03-2002. ORL-Institut. ETH Zürich.
- Patterson; Alan, T.; Kate, S. 1995: Sustainable Development, Agenda 21 and the New Local Governance in. In: Regional studies. 'Vol. 29.8'. 'S. 773-778' 995.
- Schleicher- Tappeser, R.; Strati, F.; Thierstein, A.; Walser, M. 1997: Sustainable Regional Development. EURES-Institut. Freiburg i.B.
- Thierstein, A.; Walser, M. 1997: Sustainable regional development: the squaring of the circle or a gimmick? In: Entrepreneurship & Regional Development. 9. 1997. 159-173.
- Thierstein, A.; Walser, M. 2000: Die nachhaltige Region. Ein Handlungsmodell. Schriftenreihe des Instituts für Öffentliche Dienstleistungen und Tourismus. Beiträge zur Regionalwirtschaft. Bern, Stuttgart, Wien:Haupt



## **Diane Warburton**

### **EVALUATING PARTICIPATION IN SUSTAINABLE DEVELOPMENT**

The importance of public participation to sustainable development has long been recognised. Agenda 21 stresses how the successful implementation of sustainable development is 'first and foremost the responsibility of Governments [but] ... the broadest public participation ... should also be encouraged' (Agenda 21, para 1.3). Subsequent European, national and local programmes to promote sustainable development have continued to promote participatory approaches. The principle of increasing participation by people in the decisions which affect them is now widely accepted in both sustainable development and other public policy areas.

This paper describes these trends, and examines the implications for evaluation processes. It distinguishes between the evaluation of participation and participatory approaches to evaluation, examines some current approaches, and discusses how a new 'virtuous circle' of evaluation methodologies can begin to be developed which can address some of the potential tensions and contribute to the effective implementation of sustainable development.

#### **Participation in sustainable development**

Participatory approaches have been core to implementing sustainable development from the very beginning. The Brundtland Commission (the World Commission on Environment and Development - WCED) argued that the first step in pursuit of sustainable development required 'a political system that secures effective citizen participation in decision-making' (WCED 1987). The rationale for the priority given by the Commission to participation was that 'The law alone cannot enforce the common interest. It principally needs community knowledge and support, which entails greater public participation in the decisions which affect the environment.' (ibid). The Commission suggests that the best way of doing this is by 'decentralising the management of resources upon which local communities depend, and giving these communities an effective say over the use of these resources. It will also require promoting citizens' initiatives, empowering people's organisations, and strengthening local democracy.' These approaches have remained central to all projects and programmes promoting wider participation in sustainable development.

Agenda 21 takes up the theme, with paragraph 3.2 stating that 'An effective strategy for tackling the problems of poverty, development and environment simultaneously should begin by focusing on resources, production and people and ... a democratic participation process in association with improved governance' (United Nations 1992). These general principles are repeated throughout Chapter 28 of Agenda 21, which stresses the importance of involving all sectors of society, especially those which are often excluded from policy processes (eg women, young people, etc). Paragraph 3 of Chapter 28 defines the process of Local Agenda 21:

"Each local authority should enter into a dialogue with its citizens, local organisations and private enterprises and adopt a Local Agenda 21. Through consultation and consensus-building, local authorities would learn from citizens and from local civic, community, business and industrial organisations and acquire the information needed for formulating the best strategies. The process of consultation would increase household awareness of sustainable development issues."

The importance of participatory approaches to sustainable development continues to grow. For some, this is linked to growing government and other institutional understanding of sustainable development since 1992, with a greater recognition now that sustainable development is much more than environmental issues. New priority issues for sustainable development at local level are now beginning to include poverty reduction, equity, social justice and security. In order to tackle these issues, and to tackle the causes as well as the effects of social, economic and environmental problems in order to make development sustainable (through aiming for prevention as well as cure), the participation of various constituencies, stakeholders and citizens, is vital to ensuring governance systems are seen to be legitimate, credible, respected and trusted by those they seek to serve. As a ICLEI report suggests: 'Governments cannot hope to achieve sustainability without



the active and willing participation of their citizens and their trust that government is acting for their best interests. Good governance has been held back by sceptical views of government, including a lack of accountability to constituents, insufficient involvement of citizens in the political process, inadequate representation of all stakeholder interests, insufficient transparency in the governing process, and corruption.' (ICLEI 2001).

The UN Commission for Sustainable Development has recently stated its renewed commitment to participatory approaches, saying that 'Participation generates shared values, mutually reinforcing commitments, joint ownership and partnership, which are crucial to achieving sustainable development' (CSD 2002). Local government clearly has a key role as champions and facilitators of sustainable development, with a crucial part of their role being to have 'developed participatory, multistakeholder strategies to implement sustainable development. They have promoted local governance involving the recognition of the importance of transparency, accountability and participation in governance' among other priorities (eg the importance of integration of policy and practice, strategic partnerships and information, knowledge and capacity building) (ICLEI 2001). Local governments have also 'come to recognise the importance of information-sharing, enhanced roles for civil society and other partners, and a participatory and integrated approach to the incremental implementation of sustainability' (ibid).

The continuing emphasis on participatory working does, however, place new responsibilities on government at all levels to develop their capacity to evaluate how and why these new approaches are working. And at present there is little guidance available to help them. As the CSD points out, 'Capacity development in monitoring and evaluation approaches, to support the learning and to improve public sector management and performance, including the use of participatory approaches and sustainable development indicators and complementary qualitative techniques' (CSD 2002, para 241).

### **What is meant by participation?**

Participation has been an essential element of anti-poverty and community development strategies since the UN Decade of Development in the 1950s. Then, participatory development was seen as an antidote to the failure of conventional development programmes to tackle what were seen as the root causes of poverty. The motivation then, as now, was to give more respect to the knowledge and experience of ordinary people in identifying the basic problems they were facing, and possible solutions, and not just relying on professionals, scientists, academics and technicians. In the UK, there was a major period of growth in participatory working, and wider engagement in policy debates, from the mid-1960s to the late 1970s. The current wave of expansion in participatory projects and programmes began again in the mid-1990s and (to date) that growth continues.

Examples in the UK now include extensive public participation in policy programmes on regeneration, health management, social exclusion, environmental management, social welfare and youth programmes, and others. The types of participation range from providing more information about public programmes, through simple surveys, focus groups and questionnaires designed to test public opinion, through extensively resourced deliberative processes (such as citizens' juries) which take evidence over an extended period of time and come to conclusions which may (in some cases) be binding on the institutions which commission them, formal and informal partnerships on boards or panels for single projects or major programmes, to community-led programmes resourced by public authorities. There are also independent activities in the 'third sector' comprising groups which range from established voluntary organisations (NGOs) to small informal groups which operate 'below the radar' of most public institutions but which provide services, undertake practical neighbourhood improvement projects, provide self help support and engage in policy debates, campaigns and protests.



The most useful simple analysis of participation is that undertaken by Shelly Arnstein in 1969, which summarised the most common forms of participation (Arnstein 1971).

Level 1	Manipulation	These levels assume a passive audience, which is given information which may be partial or constructed
Level 2	Education	
Level 3	Information	People are told what is going to happen, is happening or has happened
Level 4	Consultation	People are given a voice, but no power to ensure their views are heeded
Level 5	Involvement	People's views have some influence, but traditional power holders still make the decisions
Level 6	Partnership	People can begin to negotiate with traditional power holders, including agreeing roles, responsibilities and levels of control
Level 7	Delegated power	Some power is delegated
Level 8	Citizen control	Full delegation of all decision-making and action

Arnstein's ladder, as it is commonly known, has some important omissions, including those activities which contribute to democratic activity but which are not commissioned, led or controlled by government or other institutions, including those which are essentially about protest. It also focuses only on participation in decision-making, rather than including participatory action. This raises the important issue of what exactly is being participated in. In most cases, participation is analysed in terms of participation in public policy programmes led by government, NGOs or other institutions, but this remains only part of the picture.

Participation in sustainable development tends to be understood as essentially an engagement in policy debates in governmental policy development arenas. These range from global level stakeholder involvement (eg NGOs and business) in international policy discussion (eg WSSD planning meetings); European, OECD and other international policy forums; national forums, panels, commissions, inquiries, task forces etc (of which there are many); and, within the UK, regional round tables on sustainable development, and Local Agenda 21 activities. All these initiatives use various techniques to involve the public, voluntary organisations, business and other sectors of society for various reasons, including in order to gain access to their knowledge, gain legitimacy from consulting a particular constituency, and to promote knowledge and understanding of particular policy debates through access to these extra-governmental channels.

### Why has participation become so important?

The reasons why participation has grown so fast can be grouped under four basic categories (although these do overlap):

- ethics (to do with rights)
- effectiveness (to do with better projects and programmes, including access to wider range of expert knowledges)
- strengthening governance and democracy (civil society, legitimacy of government decisions etc)
- opportunities for learning and change (social and personal change, including through capacity building).



The ethical arguments centre around the belief that people have a right to participate in the decision which affect them, based on the discourse of human rights and responsibilities which underpins UN resolutions (and other international treaties). The most obvious current manifestation of this is the UN ECE's Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters, known as the Aarhus Convention. The Convention will upgrade public participation in relation to EU Directives on relation to environmental impact assessment, waste, water pollution, air quality, landfill etc. In commenting on the Convention, EC Commissioner Margot Wallstrom said 'Real environmental progress can only be achieved with the participation of the citizens concerned. The legislation is not only based on the belief that it is right for the public to be involved in decisions which affect them but also on the objective of making environmental legislation more effective and work better in practice' (European Commission 2001a).

The same approach is taken up in the EU's 6th Environmental Action Programme, entitled *Environment 2010: Our Future. Our Choice*. In setting out environmental priorities for the coming ten years, it acknowledges that 'This is not just an issue for politicians and industry; it concerns all of us' and recognises that 'People want more of a say in how decisions are made which affect the environment, and that means access to clear and trustworthy information' (European Commission 2001b). The same sentiments appear in the latest UK Strategy for Sustainable Development, which states that 'Public involvement is essential for a truly sustainable community' (DETR 1999, para 7.87), and that 'Opportunities for access to information, participation in decision-making, and access to justice should be available to all' (ibid, para 4.1).

The arguments for participation on the grounds of 'effectiveness' centre around the practical benefits of including users, beneficiaries, stakeholders and constituencies (defined in many ways) in designing, implementing and maintaining projects and programmes. These practical benefits have been summarised as follows (InterAct 2001):

- financial costs saved or avoided in the long term by establishing appropriate solutions at an early stage which met 'felt needs'
- increased user satisfaction
- reduced conflict
- greater transparency and accountability
- increased equity by involving excluded groups
- strengthened capacity of stakeholders
- improved relationships between stakeholders
- improved public image, and greater public acceptance of projects and programme
- improved communications saving staff time
- greater stakeholder ownership, commitment and responsibility, leading to willingness to sustain benefits leading to, for example, reduced vandalism and thus reduced repair and maintenance costs
- less demand on control services (eg police)
- stronger communities
- greater willingness and ability of participants to be involved in other voluntary and community activities including wider civic and democratic engagement.

The growing interest in participatory approaches for reasons of strengthening governance and democracy stem largely from challenges to the role of government, both in terms of government seeing itself as less about 'delivery' and more about 'enabling' (certainly in the UK), and about the changing role of national governments in a globalised economy, in which government is seen as less willing (and able) to act on behalf of citizens (Macnaghten et al 1995). The situation has been summed up by British Minister Peter Hain as follows: 'as governments face diminishing control over events, those they govern want more control over their lives ... people want more say in the decisions that affect them. And they want to have this say more directly' (Hain 2001). However, the rhetoric about a reduced role for national governments can be seen as a tactical discourse in the globalisation debate, or simply as wishful thinking on the part of specific international lobbies (Christie and Warburton 2001), and this rationale must therefore be questioned.

What is not in doubt is that the legitimacy of governments has been undermined by the loss of trust in politicians and others (including scientists, cf Beck 1992). In the UK, the Parliamentary Select Committee on Public Administration has recognised that 'politicians are increasingly mistrusted and representative government is adversely affected' (SCPA 2001). At the same time, some in public institutions see the need for a stronger civil society as a counterbalance to the growing power of the market, and institutions dominated by private sector interests.



The most obvious manifestation of this decline in public engagement with representative democracy is in voting turnouts. In the UK, the turnout for the General Election in 1997 was 71%, the lowest since 1951; it was even lower in 2001, at which point it was the lowest turnout since the universal franchise. In the UK local government elections in 1999/2000, turnout was only 28% (although it was slightly higher in the most recent elections in May 2002), and for elections to the European Parliament turnout was 24%. In the UK, the Government sees public disengagement in electoral politics as part of wider social trends including reduced social capital and social cohesion, and growing alienation, all of which reduce the legitimacy of governments to make decision on behalf of their citizens: "The more people vote, the greater the democratic legitimacy of actions taken by those elected. Participation in elections is therefore crucial to, and a good barometer of, the health of democracy" (DETR 1998).

In the UK, these trends are prompting Government to consider new measures for increasing public engagement including deliberative democratic mechanisms (eg citizens juries), experiments in electronic governance (eg voting via the Internet), and the creation of a People's Panel. The Parliamentary Select Committee believes that 'the health of representative and participative democracy are intertwined' (SCPA 2001). There is also growing emphasis on improving the quality of dialogues which do take place between Government and citizens, including formal guidelines being introduced for how the Government consults on new policies and programmes. In the introduction to these guidelines, Prime Minister Tony Blair says 'Everyone gains by effective consultation. It improves policy and services by harnessing the experience and ideas of the whole community, opening the work of government to greater innovation and creativity. It is critical to the existence of trust between the Government and the people.' (Tony Blair 2001).

The rationale behind the 'learning' arguments for participation are also very broad. They range from Government interests associated with the previous point, with the aim of increasing public understanding of, and engagement with, public policy processes and thus contributing to capacity building and a stronger sense of citizenship, through to radical arguments for political education such as Freire's conscientisation processes (Freire 1996). In this view of participation, it is less the practical outcomes of the projects and programmes which are the priority but rather the personal and social transformations which can result from participating in an essentially political activity.

It can be seen from the above analysis that participation has become an essential element of public policy programmes, including sustainable development, for a range of complex reasons. It will also be noted that this is an approach which has support from across the political spectrum - from the right wing who see it as linked to self help and minimal government intervention, to the left wing who see it as a way in which oppressed groups can gain some power over their lives. However, in spite of its broad political appeal as a concept, in practice the majority of public programmes are still not participatory. Various reasons have been advanced to explain this gap between rhetoric and reality, but key among them (alongside simple inexperience with the methods) is a failure to develop effective evaluation mechanisms which allow policy makers to develop a shared understanding of what 'success' in participation looks like, and to assess participatory initiatives on that basis. This is clearly a complex, and highly political, field of analysis, but progress is now beginning to be made, as will be outlined below. First, however, a brief analysis of the extent and nature of participation in Local Agenda 21 will be presented, to provide some relevant context.

### **Participation in Local Agenda 21**

Participatory approaches have been very important in the design and delivery of Local Agenda 21 (LA21) processes, both in the UK and more widely. Recent research (IDeA 2000) has shown that, of the 72% of local authorities in England, Scotland, Wales and Northern Ireland which had completed LA21s at the time of the survey, 70% had provided support for community and voluntary groups as part of their awareness raising approaches, 42% had multistakeholder forums co-ordinating their strategies, and 30% had involved the community. These are self-reported findings, and the practice on the ground may not always have lived up to the full potential implied in these results. Nevertheless, as the latest UK Government research suggests that 93% of local authorities had completed LA21 strategies by December 2001 (SDU 2001), with still more under development, this is a major activity for local government and more research to assess the full picture of levels of participation would be worthwhile.

Similar results seem to be appearing in wider surveys, with the recent ICLEI survey suggesting that, worldwide, 70% of local governments with LA21 strategies in place solicit multistakeholder input, and many have established formal partnerships across different sectors (ICLEI 2001). The LASALA research shows more details, with 74% of LA21 co-ordinators reporting information being provided to the wider public, 67%





having an LA21 Forum or Group in their areas (87% in the Western Europe region), and 49% having active and representative community involvement in the LA21 implementation process (LASALA 2001).

The types of activities undertaken to encourage participation in LA21 are varied. LASALA research found that 80% of LA co-ordinators ran workshops and seminars, 72% did work in schools, 66% had working groups, 63% had an LA21 Forum or similar, 34% ran visioning exercises and 24% ran Planning for Real exercises. Earlier research into LA21 in the UK (Young 1996, 1997 and 1998) had already shown that LA21 had promoted many innovative approaches to community participation including visioning, community profiling and village appraisals, focus groups, Planning for Real exercises, forums, round tables, citizens juries and advisory committees. LA21 processes had also encouraged local community partnerships to undertake projects with social, economic and environmental benefits including on recycling, housing co-ops, LETS, credit unions and environmental improvements. These local projects, in turn, emphasised local democracy and citizen involvement, especially of excluded social groups. While a full analysis of the implications of all these results remains to be undertaken, the interim conclusions are that 'LA21 participatory programmes have become a conduit for the unleashing of energy and ideas into the wider area of the attempts to regenerate local democracy' (Young 1998).

LA21s may not (yet) have succeeded in making the UK more sustainable in terms of social, economic or environmental development, but LA21 exercises have clearly contributed to the development of policies and practices on governance: 'whilst many of the claims about LA21 are intractable to test, there is some evidence of genuine attainment. This relates mainly to processes of strategy production, stimulation of environmental citizenship, inclusion of various sectors, challenging traditional assumptions and actions, and assisting local democracy' (Selman 1998).

This is no mean achievement in terms of achieving sustainable development. As Kofi Annan has pointed out 'Good governance is perhaps the single most important factor in eradicating poverty and promoting development. By good governance is meant creating well-functioning and accountable institutions - political, judicial and administrative - that citizens regard as legitimate, through which they participate in decisions which affect their lives, and by which they are empowered.' (Annan 1998).

Various hypotheses have been offered to explain the emphasis on public participation in LA21 programmes. Young (1998) suggests that LA21 may have been seen as a unique opportunity for local government at a time of major resource constraints on other programmes, alongside some genuine fascination with the potential for innovation through sustainable development programmes. He also points to pressure from NGOs (eg WWF UK, see Webster 1998) and of practical guidance from local government support bodies (eg LGMB 1994) as possibly important stimuli. Young also suggests that the focus on participation developed as a result of the backgrounds and interests of those individuals who are responsible for delivering LA21 at local level, especially bringing in experience from community development, adult education and involving people in environmental activities (again NGO-led eg BTCV and Groundwork trusts). These approaches were well understood in planning departments, where LA21 activities were most often located, and where participatory working was widespread as a result of years of policy guidance and practice dating back to the 1960s (Warburton 2002).

In the UK, these developments have gone alongside a powerful impetus from national Government to 'modernise' government activities at all levels (local, regional and national), including a new and much greater focus on public and community participation. As Beverley Hughes, then Minister for Local Government, said of the Local Government Act 2000: 'At the heart of our modernisation agenda was the promise to give local people a better deal, a bigger say in how their communities are run.' (Hughes 2001). The new legislation introduced community strategies, local strategic partnerships, Best Value (which has its own requirements for consultation) and opportunities for local government to restructure their decision-making processes (with directly elected Mayors, cabinets etc), all of which have been designed to promote greater community involvement in planning and delivering services.

There have been some concern in the UK that the introduction of community strategies will mean starting all over again, ignoring previous community participation exercises, particularly LA21 strategies. Certainly sustainable development and LA21 staff often remain marginalised in many local authorities. However, the Government has made clear that they 'expect community strategies to build on the best of the work done to prepare Local Agenda 21 strategies, both of which have the aim of sustainable local communities at their heart' (Hughes 2001).



## Emerging tensions around participation in LA21

There are some tensions emerging as a result of the growing focus on participation in LA21. There are the actual costs of running participatory exercises (which can be considerable, depending on the techniques used) and include staff time as well as financial costs for print, exhibitions, meeting rooms, facilitators etc. These costs need to be found at a time when existing budgets are already overstretched. But there are other problems too.

Stephen Young's assessment (1998) of participation in LA21 programmes in the UK found that some saw the results of participation as very limited in terms of impacts on policy decisions, as a result of resource limits, structural constraints (such as the organisational barriers of hierarchy and 'silo' departmental boundaries), and the fact that local priorities could be undermined by conflicts with, or shifts in, regional or national policy (eg on issues such as transport or waste management). This finding is supported by more recent evidence from LASALA that, although 73% of LA21 co-ordinators said decision-making processes were influenced by their multistakeholder forum to a 'high' or 'some' extent, and that possibilities clearly did exist for citizens to participate in decision-making, 'in practice, this is more complex, as accessing certain groups is difficult, and local authorities do not necessarily have the institutional capacity or resources to create real opportunities for individual citizens to participate' (LASALA 2001).

Research suggests an absence from participatory processes of certain groups, especially low income families, young people, ethnic minorities, older people, women's groups, disabled groups, unemployed people, businesses and unions (Young 1998, Craig et al 2001, LASALA 2001). The absence of the private sector and industry is not often noted, but 'in general, the participation of business and commerce in LA21 across Europe appears to be very limited' (LASALA 2001). When participation is seen as an awareness, learning and change exercise in pursuit of sustainable development, the absence of the private sector becomes particularly crucial.

Other problems include that one-off participation exercises (the norm) do not allow participants to accumulate the experience, knowledge and confidence which would encourage them to go on to other forms of citizen engagement - nor those running them to become confident with participatory working; there is lack of appropriate skills and experience among staff in public institutions in working with the public and other stakeholders (even at quite basic levels), and sometimes a lack of motivation among those who see their status as professionals being undermined by listening to 'lay' people; some elected councillors feel that public participation undermines their representative role, as do some NGOs who are uneasy about some direct government / public consultations which 'leapfrog' established organisations; potentially contentious problems have sometimes been placed outside the participatory exercises (eg proposals for new roads); and it is often difficult to integrate participatory initiatives either with established democratic institutions (eg elected local governments) and with more strategic planning at different spatial levels (eg regional planning and waste strategies).

For others (eg Brown 1999; White 1996), participatory processes can ignore the wider contexts of power and politics in which they operate, and can entrench and reproduce existing power relations rather than challenge patterns of dominance; they can be a vehicle for development on the cheap and support the retreat of the state from service provision, rather than being about democratisation and empowerment; they can be based on assumptions about the homogeneity of communities with shared interests rather than reflecting respect for diversity and complexity; and they can inappropriately minimise the role of scientific rigour and technical expertise to the extent that it can be ignored (Pollitt 1999).

Many of these problems can be addressed in well-designed participation programmes, but two key issues remain. Firstly, there are currently no commonly agreed methods of assessing the success (or otherwise) of participatory projects and programmes, and no commonly agreed criteria or indicators which can be used to test it. There is some experience of evaluating participation in specific fields, such as regeneration and community development, but only part of this work can be applied more widely. Secondly, as a result, there is little hard evidence of the effectiveness of participatory working compared to conventional, external, expert-led projects and programmes, and little understanding (beyond the anecdotal) of what works in different circumstances. Even in the current 'evidence-based' policy making circles in the UK, much of the impetus for increased participation comes from arguments of principle and of knowing what does not work (ie top down regeneration programmes which prioritise physical improvements rather than community engagement and social development), rather than what does. The remainder of this paper addresses these two issues in more detail.



## Evaluating participation in sustainable development

The complexities of sustainable development, and the evolving understanding of the interrelationships between the issues, in theory and practice, have resulted in the development of many indicators and evaluation processes related to the social, economic and environmental issues. There are quality standards, such as EMAS and TQM, which often concentrate on the results from processes; standardised tools or models to measure public sector success or efficiency, such as the Public Administration Excellence Model, which do evaluate processes but have problems and weaknesses with adaptation for sustainable development; local, regional, national and international indicators, although these are primarily environmental and very few are associated with participation - even those associated with 'education' are rare: recent research found only 38% had educational targets in Local Agenda 21s, and only 24% had indicators for education (LASALA 2001). Certainly, to date, the governance elements have received far less attention than other elements of sustainable development. There is now some more detailed examination of the issues at international governance levels, with the publication of the EU White Paper on Governance, and the UNCHS Global Campaign on Urban Governance, but these have yet to impact upon practice and evaluation.

At national level, in the UK, there are also developments. The Modernising Government programme (mentioned above) is creating pressures for effective assessment methods for participation, and there is some initial work by the Audit Commission (December 2001) to develop performance indicators for community involvement. These developments are essential for the successful implementation of sustainable development for all the reasons outlined above in the summary of views from the CSD, UN and others.

Essentially sustainable development is a process, through which appropriate behaviours, policies and priorities are debated and determined. As Paul Selman says: 'sustainability is concerned as much with process as with product ... the journey is as important as the destination' (Selman 1995). The quality of the process will determine the legitimacy, effectiveness and acceptability of the policy outputs which result from the process, and thus the extent to which the public (and other stakeholders) will value, accept and 'own' the changes that will be required to lifestyles and aspirations.

It is therefore vital that clear methods and criteria are developed so that participatory sustainable development processes can be assessed effectively, so that lessons can be learned and progress made both in improving those processes and the outputs and outcomes that result. If no-one knows what success looks like, they cannot repeat it or improve on it.

## Evaluating participation in sustainable development: benefits and limitations

Before outlining the benefits and limitations of participation in sustainable development, it is important to make clear the distinction between participatory evaluation and the evaluation of participation. A participatory programme can be evaluated using non-participatory evaluation methods, and a non-participatory programme can be evaluated using participatory methods.

Both are needed, for different reasons. More details are given below but, in brief, evaluation of participation is needed to begin to ensure that data is available on participatory working, which will allow assessments of good practice to be made (ie to achieve better practice in sustainable development in future). And participatory evaluations are needed to ensure the input of stakeholders into assessments of programmes of which they have knowledge and in which they have an interest (ie to achieve better evaluation and to meet ethical concerns about rights to involvement, as well as contributing to active citizenship and better governance systems).

The benefits of evaluating participation may include:

- Improving the practice of participatory working by capturing, analysing and sharing experience of good practice, and what works in different circumstances.
- Building support for participatory ways of working, by providing evidence of effectiveness and achievement, and learning processes to support often isolated workers.
- Contributing to the development of the theory and analysis of participatory working, including creating new theoretical models, methodologies and criteria for success which cover process outcomes (eg trust, ownership, understanding) as well as product outcomes (eg physical improvements, better air quality or greater biodiversity).
- Helps develop the sophisticated social science methodologies which are needed for sustainable development. The social science perspective explicitly recognises the particularity of context (including



constantly shifting policy and political contexts and resource constraints), the complex dynamics of the social world (including human motivations as well as social institutions) and the heterogeneity of settings for sustainable development decision making. Social science methodologies have helped transform the way sustainable development is understood by introducing issues which were barely recognised ten years ago (GEC 2000), including:

- uncertainty and complexity
  - recognition of a diversity of 'publics' with diverse values, knowledges, cultural identities
  - creating different ways of framing environmental risks and potential strategies to resolve problems
  - recognition that different sectors have different abilities to tackle problems
  - recognition that trust is a vital element in public perceptions of science and institutions, and that the development of inclusionary processes can help revitalise trust in science and policy.
- Helps develop new measures of processes to improve governance, an essential element of sustainable development. Conventional measures of performance, value for money etc in public services tend to be restricted to terms of economy, efficiency and effectiveness, but governance issues can introduce three additional e's: excellent, equity and empowerment (Jackson 1991).
  - Go beyond 'consumerist' measures of quality to identifying 'social' measures eg of environmental goods, which are not just to do with consumption but also to do with altruism (or at least enlightened self interest).
  - Value other forms of knowledge beyond the purely technical and scientific.

The limitations of evaluating participation may include:

- Evaluating participation can be costly in terms of time and money, in the short term (although they should save resources in the long term by increasing effectiveness). Participatory programmes are often run by organisations with limited resources, such as NGOs and community bodies (and some local authorities), who can barely fund programmes let alone research. New sources of funding may need to be identified if effective evaluations are to be developed in any coherent way.
- They require new methods of working which may be unfamiliar even to experienced social researchers. New methods always take more time, and will need to be developed and honed before they are fully effective.
- There are particular problems in evaluating any 'human service' programmes, as these never exist in isolation and it is therefore extremely difficult to assess the specific contribution of a particular programme to any given outcome, especially in view of the constantly changing policy and political contexts. Alcock et al suggest that the best that can be hoped for is to articulate what else is going on which may affect (or have affected) outcomes by using various methods of data collection and ensuring a range of perspectives is trawled (Alcock et al 2000).
- Timescales are critical. Many of the full impacts of participatory programmes are not apparent until many years after the initial activity, and mechanisms may need to be identified for capturing results over a long period (ten years or more). Ten years is now the timescale for community-based regeneration programmes in the UK, under the New Deal for Communities programme.

The benefits of participatory evaluation include:

- Providing an initial step to engaging stakeholders if the programme being evaluated is not participatory, or an additional step and continuing engagement if the programme is participatory.
- Better understanding of findings by funders, beneficiaries and other stakeholders as a result of them sharing discussions on the implications of the complexities of political and policy contexts, as well as more straightforward lessons.
- Greater ownership of findings by stakeholders as a result of helping to design and manage the evaluation process.
- Greater understanding among all stakeholders of individual and shared choices, values and assumptions, through jointly developing criteria for success.
- Access to a wider range of knowledges.
- Opportunities for reflection and joint learning, depending on the level of participation in the evaluation.
- Opportunities for 'reality testing' of findings if stakeholders can feed back on findings at early stages (or at least before conclusions are finalised).
- Capacity building and learning opportunities for researchers and other stakeholders as a result of joint working and sharing experience.



- Ensures a recognition that social scientists (like all scientists) are not value-free, by requiring that values, choices and assumptions are made explicit to stakeholders.

The limits to participatory evaluations may include:

- As with the evaluation of participation, there may be additional costs (time and money) at various stages.
- Also like the evaluation of participation, but possibly even more so, participatory evaluation is likely to require the use of new methods, unfamiliar even to experienced social researchers.
- The legitimacy of the research will depend to some extent on ensuring the representativeness of stakeholders, and appropriate mechanisms for accountability, which can be complex to identify and manage.
- Researchers are likely to have to manage conflicting pressures from different stakeholders, including potentially different value frameworks which will affect definitions of the success of the work.
- Results may be compromised if evaluators are not seen to be sufficiently objective, and appropriately distant from all parties: funders, commissioners and participants. Researchers are likely to have to develop relationships of trust with those whose work is being evaluated (to ensure there is access to all relevant data and to avoid secrecy and misinformation), while maintaining sufficient distance to be trusted by those commissioning the research (if they are different).
- The status of findings may be compromised if audiences for findings feel the results have been unduly influenced by those with a vested interest in the project or programme's 'success'.

### Emerging tensions and models

Lessons from practice of evaluating participation in sustainable development are now beginning to emerge, which have revealed some complex tensions as well as some models. The central tension is around choices between participatory and non-participatory approaches to evaluating participation in sustainable development.

Research for the local government Improvement and Development Agency (IDeA) in the UK, on poverty and social inclusion programmes, suggests that the approach to evaluation should reflect the objectives / values / criteria of the programme being assessed, eg participatory methods of evaluation for participatory programmes (Alcock et al 2000b). However, this seems to undermine the potential for developing participatory evaluations of non-participatory programmes, which can be a positive introduction to participatory working for those who have not worked in that way before. It also limits the potential for some evaluations of participatory programmes to include only limited participation in order to meet instrumental objectives, where consultation rather than a fully participatory evaluation is appropriate (Johnson undated). In addition, two evaluations of participatory programmes (evaluations which were not fully participatory according to Rebien's criteria) have produced some of the little 'hard' evidence on the effectiveness of participatory working that exist.

The first of these, by the World Bank, is a simple classic cost benefit analysis which compared the costs and benefits, over time, of participatory and non-participatory programmes funded by the Bank. Their findings showed that, overall, participation by beneficiaries was 'the single most important factor in determining overall quality of implementation', and made a significant contribution to project effectiveness including resulting in lower operational costs eg maintenance (World Bank 1994).

The second is an analysis, by the UK Department of Health, into the health benefits of participation. This research used traditional social science methodologies and concluded that there was an increased likelihood of people reporting poorer health generally, higher stress levels, higher prevalence of smoking, and poorer diet quality among those who felt they had a lack of control over decisions affecting life and a lack of influence over neighbourhood decisions, those with no involvement in community activities, those living in more deprived areas with a low 'neighbourhood social capital' score and, especially for men, those with no personal support group. Direct cause and effect are difficult to prove in research on health, but this research shows that those who feel empowered to make their own decisions, are engaged in community activities and live in places with strong neighbourhood social capital, are less likely to report poor health and less likely to have unhealthy lifestyles (eg related to diet and smoking) than those who do not.



The willingness of audiences sceptical of the value of participation to accept the findings of these two research projects clearly suggests another issue which needs to be taken into account in evaluating participation, which is to do with the status and reputation of the body commissioning the research, and the appearance of 'objectivity' of those undertaking the research and analysis. The relevance of these additional dimensions need to be taken into account in identifying appropriate methods for any specific evaluation study: the balance between credibility among users and beneficiaries (as a result of greater control over process, outputs and outcomes) and credibility among funders and other target audiences.

The choices between using a participatory or non-participatory approach to evaluating participatory or non-participatory programmes will depend on the objectives of the evaluation. The objectives (what the evaluation is expected to achieve) are, in turn, likely to be based on the same four elements outlined earlier as the usual rationales for participation: ethics, effectiveness, strengthening governance and democracy, and opportunities for learning and change. Evaluation objectives, like those for participation programmes, may be instrumental or transformative (O'Riordan et al 1999):

- **Instrumental** evaluation is a means to an end (sometimes referred to as Habermasian after Jurgen Habermas). Instrumental approaches focus on practicality (eg achieving goals more effectively) and the legitimacy of decisions. Outcomes such as strengthened civil society and democracy could also be seen as instrumental. Relevant processes would focus on public reason, persuasion by the best arguments, production of consensus, and implementable and legitimate decisions. Evaluation criteria would be around the extent to which outcomes are do-able, and their legitimacy.
- **Transformative** evaluation is an end in itself, as well as a means to a better product (sometimes referred to as Arendtian, after Hannah Arendt). Transformative approaches would focus on a learning, responding, capacity building, citizenship process, and aim for empowerment, creative agency and self esteem. Relevant processes would be about a sense of recognition and agency, and of being part of a shared society through shared stories. Evaluation criteria would include the extent to which a process generated opportunities for learning, a sense of empowerment and agency, social intelligence and self fulfilment, as well as a sense of belonging to a shared society.

Sarah White expands these categories to four (White 1996): nominal (evaluation for 'display' eg PR purposes); instrumental (as above); representative (opportunities for groups to have a voice and express their own interests); and transformative (as above). Objectives such as strengthening civil society, enhancing democracy and enlargement of citizenship (or facilitation of other (eg state) agenda) could fit into either the instrumental or transformative models, depending on the motivations of whoever is promoting (and defining the objectives of) the participatory process and/or its evaluation.

Hunt and Szerszynski (1999) suggest associated tensions which can result between instrumental and transformative objectives for evaluations, including between problem-solving and relationship building approaches, cultural empowerment and structural change, digestibility and authenticity (ie between preserving the authenticity of participants own words and creating outputs which can be digested by institutions in the form of reports and recommendations, requiring 'translation'), ambivalence and consistency (recognising shifting policy and political contexts while also coming to some general conclusions which can be meaningful to decision makers).

Even when a participatory approach to evaluation is considered appropriate, further tensions arise. Firstly, it has been suggested that all evaluations are participatory, because they need to at least take into account the views of users, beneficiaries, stakeholders etc (Rebien 1996), but the degree to which they are participatory varies. Arnstein's ladder of levels of participation could be used to analyse the levels of participation in participatory evaluation, as well as in participatory exercises themselves, as Arnstein's analysis addresses the key issues of power and control. In evaluation of participation, the key questions are around value and judgements (Alcock et al 2000a), and the issues of power and control arise in addressing whose assessment of the work is valued and why and how that value is measured: what Robert Chambers summarised in the question 'whose reality counts?' (Chambers 1997).



Secondly, there are various different approaches to participatory evaluation. The first is 'fourth generation evaluation'. First generation evaluation is seen to be about measurement by a 'technical' evaluator; second generation about describing patterns of strengths and weaknesses by an evaluator operating as a 'describer' (and covering technical aspects); the third generation about judgement, with the evaluator operating as a 'judge' (as well as a describer and technical). Fourth generation evaluation is 'responsive constructivist evaluation', which is essentially 'participatory evaluation' in which the evaluation's parameters and boundaries are set through an interactive negotiated process with stakeholders (Guba and Lincoln 1989).

The 'theory of change' approach (which has been used in community development by CDF) is essentially a participatory planning process in which the goal is to generate a theory of change which is plausible, doable and testable and which makes explicit the pathways of change the project is expected to follow. Here, theorising happens in advance and is then tested as the process unfolds, through 'theory surfacing' rather than imposing theory on a body of data (Connell and Kubish 1996).

The model designed by InterAct is designed to be participatory, but could be adapted for an evaluation which was only consultative: InterAct is an alliance of experienced practitioners, researchers, writers and policy makers in the field of public participation and stakeholder engagement. They aimed to produce a simple practical framework for evaluating participatory, deliberative and co-operative ways of working, to provide some immediate support to practice, and to increase the sharing of information about methods. The InterAct framework was trialled in an EU LIFE funded project on participatory river basin management, and was launched (with the Institute for Public Policy Research) in June 2001 in order to start a debate on the issues, leading to an updated framework later.

The InterAct framework provides a basic checklist covering both what needs to be examined when evaluating participatory processes, and how it should be done. In terms of what to look for, InterAct proposed the following issues need to be covered: objectives (how they were set, what they were, extent to which they were met etc); context (external and internal factors affecting the initiative); levels of involvement (based on Arnstein's ladder of participation); methods and techniques used; stage reached (not all evaluations are done at the end of projects); inclusiveness (extent to which efforts are made to reach traditionally excluded groups); commitment to using the results of the evaluation; baseline, inputs, outputs and outcomes (inputs in terms of time and money; outputs such as newsletters, reports; outcomes such as personal and organisational change - as well as practical changes such as reduced maintenance costs or increased leverage of support).

In terms of how to undertake evaluation processes, InterAct proposed that the following need to be taken into account: clear objectives; clear principles; designing for use (including identifying the needs and expectations of target audiences for the results); methods (finding appropriate methods depending on circumstances); stakeholder involvement (level and processes necessary to involve stakeholders in design, implementation and dissemination); indicators (including setting appropriate indicators with stakeholders); timing and timescales; cause and effect (recognising difficulties of proving direct links); embedding the process (ensuring evaluation is linked to wider management structures and processes, and reporting procedures); presenting the findings and using the results.

InterAct has also begun to share criteria and indicators which may be useful for evaluating participatory programmes, purely as examples to encourage others to develop the most appropriate indicators for their own programmes. Some examples are given in the Annex to this paper.



## Summary and conclusions

In summary, this paper has described the centrality of participation to implementing sustainable development, described how participatory approaches have been used in delivering LA21s in the UK and more widely, and examined some of the tensions which have emerged. Some of the reasons why participation has become so important to public policy have been examined (focusing on issues of ethics, effectiveness, strengthening governance and democracy, and creating opportunities for learning and change), and the implications of these drivers for evaluating participation in sustainable development. Some of the tensions emerging from the practice of evaluating participation in sustainable development are identified, particularly the relevance of participatory evaluation, and some current key models are outlined: 'fourth generation evaluation', the 'theory of change' model, and a practical framework.

In conclusion, it is apparent that evaluating participation and participatory evaluation are still very new, and there remain many issues to be resolved in finding methods which can command wide support. Not even all those who are deeply engaged in participatory projects have necessarily been convinced of the value of evaluating participation at all - some still see evaluation as a doomed attempt to simplify and 'quantify' a rich, complex and highly political process through the use of crude criteria and indicators. Even apparently much more established processes of learning from good practice remain a developing art, with wide recognition that the evidence cannot show 'what works' but rather 'what works better than something else in particular circumstances'.

What is clear is that participation will be evaluated, and that those with the most experience and understanding of participation need to be engaged in designing those evaluation processes to ensure their expertise and knowledge helps shape appropriate methodologies. It is also clear that participation needs to be evaluated, to assess the achievements of current approaches and create hard evidence based on rigorous research methods which can show balances between inputs, outputs and outcomes. Without this evidence, and the learning that can follow, the current boom in participatory working will start to wane, as it did in the 1980s after the previous boom in the mid 1960s to 1970s. This would not only be disappointing to practitioners and academics engaged in and committed to these approaches, but would have major consequences for the successful implementation of sustainable development, the credibility of governments, the strength of civil society, and the expansion of democracy. There is a lot at stake.

Tim O'Riordan suggests that 'The best evaluation is instructive, collective, continuous and appropriately correcting' (O'Riordan 1999). And the ideal situation may be to establish a balance between instrumental and transformative objectives, clear ethics and principles, participatory and non-participatory methods, qualitative and non-qualitative indicators that are appropriate (according to various audiences) and verifiable (ie numerical but also explanations of why and how), and agreement on timescales. Even where the ideal is not possible, some kind of balance needs to be struck to ensure that achievable objectives, ethical principles, appropriate methods and learning from results can at least be aimed for in a new 'virtuous circle' of learning from experience in ways which also help to develop better methods of assessment. These evaluation activities also need to be undertaken in an appropriate way, and 'need to support the process whilst at the same time understanding and evaluating it - evaluation should ideally be linked to building capacities' (LASALA 2001).

A more consistent and strategic approach to public and stakeholder engagement is fundamental to sustainable development, democratic renewal, social inclusion, and a vibrant civil society, and such challenging and complex issues can only be tackled by processes that take full advantage of the added value that comes from wide ranging participation and collaboration. Effective evaluation methods are an essential part of that strategic approach.





## References

- Alcock, Pete; Barnes, Cathy; Craig, Gary; Harvey, Aidan and Pearson, Sarah (2000a) *Evaluation: issues and concepts*. Discussion paper 1, for What Counts? What Works?
- Alcock, Pete; Barnes, Cathy; Craig, Gary; Harvey, Aidan and Pearson, Sarah (2000b) *What Counts? What Works? Evaluating anti-poverty and social inclusion work in local government*. Improvement and Development Agency (IDeA), London.
- Blair, Tony (2001) Foreword to Cabinet Office draft Code of Practice on Consultation; part of the Service First programme.
- Brown, Katrina (1999) 'What can the North learn from the South? Deliberative inclusionary process and development studies', in O'Riordan et al (1999) *Deliberative and Inclusionary Processes*. A report from two seminars. CSERGE Working Paper PA99-06, Centre for Social and Economic Research on the Global Environment, University of East Anglia.
- Chambers, Robert (1997) *Whose Reality Counts? Putting the First Last*. Intermediate Technology Publications, London.
- Connell, James P. and Kubish, Anne C. (1996) *Applying a Theory of Change Approach to the Evaluation of Comprehensive Community Initiatives. Progress, prospects and problems*. Aspen Institute, New York.
- CSD (2002) *Implementing Agenda 21. Report of the Secretary General*. Commission on Sustainable Development (United Nations Economic and Social Council) acting as the preparatory committee for WSSD. Advance unedited text.
- DETR (1999) *A Better Quality of Life. A strategy for sustainable development for the UK*. Department of the Environment, Transport and the Regions, May 1999.
- DoH (1998) *Health in England 1998: Investigating the links between social inequalities and health*. Department of Health. The Stationery Office, London.
- Freire, Paulo (1996) *Pedagogy of the Oppressed*. Penguin, London (originally published in 1970).
- GEC (2000) *People in the Environment. Ten years of Environmental Social Science*. Global Environmental Change, an ESRC Research Programme, Swindon.
- Gregory, Amanda (2000) 'Problematizing Participation. A critical review of approaches to participation in evaluation theory', in *Evaluation Vol 6 (2)*.
- Guba, Egon G. and Lincoln, Yvonne S. (1989) *Fourth Generation Evaluation*. Sage Publications, London.
- Hain, P. (2001) *The End of Foreign Policy? Britain's Interests, Global Linkages and Natural Limits*. Fabian Society, Green Alliance and Royal Institute of International Affairs, London.
- Hughes, Beverly (2001) 'Putting councils in touch with local people: the role of community strategies and the wellbeing power', in *EG Magazine*, January 2001.
- Hughes, Mike and Traynor Tish (forthcoming) 'Reconciling process and outcome in evaluating community initiatives', article for *Evaluation*.
- ICLEI (2001) *Accelerating Sustainable Development: 'Local Action Moves the World'*. Local Government Dialogue Paper for the World Summit on Sustainable Development. Prepared by the International Council for Local Environmental Initiatives (ICLEI) in collaboration with the International Union for Local Authorities (IULA) and others. December 2001.
- InterAct (2001) *Evaluating participatory, deliberative and co-operative ways of working*. InterAct Working Paper, London.
- Jackson, J. (1991) *Dials or tin-openers: the role of performance indicators*. Paper to NCVO conference on evaluation in the voluntary sector, NCVO, London.
- Johnson, Deb (undated) *Participatory Evaluation? Yeah, That's Fine. But we want a conventional report! A case study of the perils of designing and implementing a participatory evaluation*. Sikiliza International Ltd, Kampala, Uganda. Unpublished paper, via ICA.
- LASALA (2001) *Accelerating Local Sustainability - Evaluating European Local Agenda 21 Processes*. Volume 1. Local Authority Self Assessment of Local Agenda 21 (LASALA). ICLEI, Freiburg, Germany.
- LGMB (1994) *Community Participation in Local Agenda 21*. Jeff Bishop for Local Government Management Board, now the Improvement and Development Agency (IDeA), London.
- IDeA (2000) *Progress with Local Agenda 21 in England and Wales - Spring 2000*. Improvement and Development Agency for local government (IDeA), London.



- Macnaghten, P., Grove-White, R., Jacobs, M., Wynne, B. (1995) Public Perceptions and Sustainability in Lancashire. Indicators, Institutions and Participation. Lancaster University for Lancashire County Council.
- O'Riordan, Tim; Burgess, Jacqueline and Szerszynski (1999) Deliberative and Inclusionary Processes. A report from two seminars. CSERGE Working Paper PA99-06, Centre for Social and Economic Research on the Global Environment, University of East Anglia.
- Pollitt, Christopher (1999) 'Stunted by Stakeholders? Limits to Collaborative Evaluation', Public Policy and Administration, Vol 14 (2), Summer 1999.
- Rebien, C. C. (1996) 'Participatory Evaluation of Development Assistance. Dealing with power and facilitative learning' in Evaluation Vol 2 (2).
- United Nations (1992) Report of the United Nations Conference on Environment and Development, Rio de Janeiro. United Nations, New York.
- Warburton, Diane (1998) Community and Sustainable Development. Participation in the Future. Earthscan, London.
- Warburton, Diane (2001) 'Evaluating participation: making a start' ARVAC Journal, October 2001
- Warburton, Diane (2002) 'Policy context for community involvement in countryside planning', in From Global to Local: Developing comprehensive approaches to countryside and nature conservation, Kevin Bishop and Adrian Phillips (eds), Earthscan, forthcoming.
- WCED (1987) Our Common Future. Oxford University Press, Oxford.
- Webster, Ken (1998) 'Developing Community Local Agenda 21s', in Community and Sustainable Development. Participation in the Future, Diane Warburton (ed), Earthscan, London.
- White, Sarah C. (1996) 'Depoliticising Development: the uses and abuses of participation', in Development in Practice, Vol 6 (1), February 1996.
- World Bank (1994) The World Bank and Participation. World Bank Learning Group on Participatory Development, Operations Department, September 1994.



## ANNEX A. EXAMPLES OF INDICATORS AND CRITERIA

### Example 1. Barclays Site Savers core indicators

The New Economics Foundation devised a set of indicators to test levels of trust in the Groundwork UK's programme of urban regeneration known as Barclays Site Savers. The idea was to identify some simple core indicators which could then be extended depending on local circumstances. The indicators included:

- 'I feel I could help change attitudes and improve things around here'.
- 'I have learned new skills on the project in the last 6 months'.
- Percentage of respondents saying: within the last 6 months I have enjoyed several conversations with a new person from a different age and/or background.
- Percentage of respondents saying: Neighbours around here look out for each other.
- Percentage of respondents saying: I think the project/facility will survive.
- How many new friends have people made through the project?
- Percentage of respondents saying: I know who to contact to help me change locally.
- Percentage of respondents saying: I have benefited from being involved with Groundwork.
- Number of people (previously unknown to Groundwork/the lead agency) involved in the project over the last 6 months.
- Number of agencies working with Groundwork (or working together) on the project.

For more information, see *Prove it! Measuring the effect of neighbourhood renewal on local people*, by Perry Walker, Julie Lewis, Sanjiv Lingayah and Florian Sommer. Published by Groundwork UK, Birmingham, and the New Economics Foundation, London, June 2000.

### Example 2. Evaluation of Rural Action for the Environment

The criteria devised for this evaluation of a national programme included the following:

- Total funding from the programme, compared to match funding from elsewhere, to show levels of leverage.
- Types and numbers of projects funded, to assess extent and breadth of work undertaken.
- Types of groups receiving support, to assess 'reach' and inclusiveness of the scheme, and the extent to which the scheme reached 'new audiences' for environmental work.
- Capacity building, assessed by examining:
  - the amount of training and advice provided, and learning achieved
  - the extent to which groups have developed from their initial projects
  - the number of new groups supported by the scheme
  - the extent of participation amongst groups supported, calculated by assessing:
    - number of groups involved in the scheme
    - types of groups involved in the scheme
    - numbers of people involved in those groups
    - voluntary action person days
    - extent and quality of participation for those involved
  - personal testimony from those involved.
- Extent of Rural Action influence on others, assessed by examining:
  - examples of how mechanisms pioneered by Rural Action were taken up by others
  - examples of how certain organisations and institutions had changed priorities over the time Rural Action had been running (eg parish councils), with statements
  - examples of how local authorities had changed practices over the time the scheme had been running
  - examples of change to individuals who had been involved.

For more information, see *The Achievements and Effectiveness of Rural Action: An Evaluation*, by Diane Warburton. For The Countryside Agency on behalf of the Rural Action Steering Group, July 1998.



### Example 3. LITMUS Project, south London

The following criteria were used for monitoring and evaluating impacts of the participatory process of the LITMUS project in Southwark, south London:

Evaluation of outcomes: qualitative criteria

- level of understanding about LITMUS
- level of trust / faith in LITMUS approach and consultation process
- involvement perceived as useful
- level of encouragement / facilitation
- level of ownership regarding LITMUS
- empowerment of the people / groups involved

Evaluation of outcomes: quantitative criteria

- number of individuals / organisations participating in LITMUS
- number of volunteers engaged
- number of volunteer hours/days spend
- continuity of involvement
- number of independent actions
- number of individuals / organisations acting as facilitators for LITMUS.

### Example 4. Comedia

A Comedia study provides a useful example of how to measure the impact of participatory processes on personal change. The following questions were asked, inviting a yes, no or don't know answer.

- Since becoming involved, I have ...
  - .. become interested in something new
  - .. learnt about other people's cultures
  - .. become interested in something new
  - .. learnt about other people's cultures
  - .. been to new places
  - .. tried things I haven't done before
  - .. become more confident about what I can do
  - .. decided to do some training or course
  - .. felt healthier or better
  - .. become keen to help in local projects
  - .. been happier
- Has taking part had any bad effects on you?
- Do you feel differently about the place where you live?
- Has taking part encouraged you to try anything else?
- Have you learnt any skills by being involved?
- Could you do it better than you could have before?
- Was doing something creative important to you?

For more information, see Comedia's *Use or Ornament: The Social Impact of Participation in the Arts*.



**Clive George**

## **APPLICATIONS OF SUSTAINABILITY EVALUATION AT THE NATIONAL AND INTERNATIONAL STRATEGIC POLICY LEVEL**

Institute for Development Policy and Management, University of Manchester, UK<sup>111</sup>

### **Introduction**

This paper describes two applications of evaluation techniques for the promotion of sustainable development in national and international policy-making. The first is based on a programme of work being undertaken for the European Commission, to carry out Sustainability Impact Assessments of the current round of WTO Trade Negotiations (Kirkpatrick et al 2002). The second application described is the development of a framework and criteria for evaluating the effectiveness of national sustainable development strategies, undertaken for the UK Department of International Development and the OECD Development Assistance Committee (Kirkpatrick, George and Curran 2001).

### **1. Sustainability impact assessment of WTO trade negotiations**

Any evaluation of the implementation of sustainable development requires interdisciplinary expertise covering all three of the concept's social, economic and environmental pillars. This is particularly so for the complex task of evaluating the impact on sustainability of world trade negotiations. The work described here has been carried out by a consortium led by a multi-disciplinary team in the University of Manchester's Institute for Development Policy and Management, which includes the CarlBro Group, the Overseas Development Institute, the British Institute of International and Comparative Law, the Centre for Agriculture, Food and Resource Economics, the Environmental Impact Assessment Centre and BMT Cordah. Contributions to the project have also been made by individual specialists in many other organisations.

### **Background**

Sustainable development has been adopted as a fundamental pillar of the European Union's development policy:

*All [EU] policies must have sustainable development as their core concern...careful assessment of the full effects of a policy proposal must include estimates of its economic, environmental and social impacts inside and outside the EU...*

European Commission, A Sustainable Europe for a Better World:  
A European Union Strategy for Sustainable Development  
com (2001) 264

*... mechanisms to ensure that all major policy proposals include a sustainability impact assessment covering their potential economic, social and environmental consequences*

Göteborg European Council meeting  
June 2001

*[a] system of SIA should be in place within the Commission by end 2002*

SECGEN paper to be presented to the Seville European Council  
June 2002

---

The work on which this paper is based was carried out with financial assistance from the Commission of the European Communities and the UK Department for International Development. The views expressed are those of the author, and do not represent any official view of the European Commission or the UK government.

<sup>111</sup> Institute for Development Policy and Management, University of Manchester, Manchester M13 9GH e-mail: clive.george@man.ac.uk



In accordance with these general policy statements, sustainable development has been embraced as main objective of the EU's trade policy. The work described has been carried out for the Commission's DG Trade, which has committed approximately ten percent of its budget to sustainability impact assessment.

### **Sustainability impact assessment**

In this context, sustainability impact assessment is a systematic process for the ex-ante assessment of the potential economic, social and environmental impacts of policy proposals, carried out in parallel with policy development and approval. The prime aim is to identify potential impacts in all three spheres, such that positive effects may be enhanced, and action taken to counter negative ones, either through selection of appropriate policy alternatives, or through parallel mitigation and enhancement measures.

In the first instance, this type of evaluation does not require impacts in one sphere to be balanced against impacts in another, nor any detailed interpretation of the sustainable development context. A degree of balancing is necessary in the evaluation of alternative policies and mitigation measures, but the prime aim of the assessment is to identify potential impacts.

The assessment is based on consultation of relevant stakeholders, and requires involving the public in the development of policy, with the aim of enhancing transparency and better governance.

Both the USA and Canada have also undertaken strategic impact assessments of trade policy, in relation to the North American Free Trade Agreement. These assessments have focused on environmental impacts, under the countries' laws and policies for environmental impact assessment. However, even these assessments have required an evaluation of economic and social impacts as well as environmental ones. The high degree of inter-dependence between effects in the three spheres is such that it is generally not possible to assess impacts in one without evaluating all three. This is an integral part of the approach adopted by the European Union.

### **European Commission's SIA of WTO trade negotiations**

The work described began in 1999, as part of the EC's preparations for the World Trade Organisation's Ministerial Conference in Seattle.

Phase I of the project consisted of the development of a methodology for undertaking a preliminary SIA, which was carried out in Phase II. Both phases were completed in advance of the Seattle meeting in November 1999. The difficulties encountered at the Seattle conference delayed further work, but the EC subsequently commissioned the further development of the SIA methodology, as appropriate for more detailed studies to be undertaken in Phase III.

Following the WTO's fourth ministerial conference in Doha in November 2001, Phase III of the programme has now started. Its aim is "to provide an analysis of the sustainability impacts of agreed policy options or scenarios, and to present this analysis in such a way as to give a concrete input for negotiators in their search for a balanced set of policies, including any necessary flanking measures".

The Phase III SIA programme is intended to consist of:

- a preliminary global SIA
- detailed sector studies
- a global SIA of provisional agreements

Because of pressure of the negotiation timescale, some of the sector studies have begun in parallel with initiation of the preliminary global SIA. Otherwise, this global overview study will contribute to defining the depth of investigation needed in the sectoral studies. The second global study at the end of the programme is intended to draw together the results of the sectoral studies, investigate cross-sectoral effects, and provide an overview of the impacts of the full set of proposals before final agreements are made.



Four sector studies have been initiated to date:

- Agricultural products
- Market access
  - pharmaceuticals
  - non-ferrous metals (especially aluminium)
  - textiles
- Environmental services
  - water
  - waste treatment
- Competition policy

The first of these studies (agricultural products) was carried out by the Stockholm Environment Institute in parallel with finalisation of the Phase III methodology, and is now complete. The other three sector studies and the preliminary global SIA are being undertaken by the Manchester led consortium.

## **Main stages in the assessment process**

### Screening and Scoping

The first stages of the SIA process are intended to target the available resources onto those aspects the proposed trade measures whose impacts are likely to be most significant.

Screening involves an initial review of each measure, to determine which measures, and which components of each measure, are most likely to give rise to significant impacts.

Scoping is undertaken to determine the terms of reference for the assessment of each measure which is to be appraised, based on the likelihood of significant impacts. It examines the components of each measure to identify impacts that are likely to be significant, and excludes those which are unlikely to be of serious concern. Scoping involves simplified forms of causal chain analysis to identify the cause-effect routes by which significant impacts may occur. It may also identify assessment methods, consultation procedures etc. which could be used at later stages in the assessment process

### Detailed Assessment

Causal Chain Analysis (CCA) is used to trace the links between trade measure/policy change and final sustainability impacts.

Judgements need to be made on the 'best' procedure for detailed assessment, based on:-

- the needs of the effects being studied
- appropriate assessment methods (qualitative, quantitative)
- data availability and quality
- time and budget constraints.

Although time and budget constraints are fairly severe, tied to the timescale of negotiations, the principal constraints on the evaluation are the availability of reliable data, along with a limited understanding of cause and effect relationships between an initial change in the trading regime and its ultimate consequences, and uncertainty in predicting the policy responses that may be adopted in the different countries likely to be affected. Even with an infinite budget and no constraints on time, the level of uncertainty in impact predictions would remain high. It is therefore essential that the assessment acknowledge and identify the degree of uncertainty inherent in its results. This can then be taken into account in identifying those potential impacts that require mitigation, or a programme for impact monitoring and contingency planning.

The findings of the evaluation are summarised in assessment matrices, indicating the anticipated significance of the impacts, in terms of core indicators (themes). It is important that these matrices are supported by textual explanation and an evidence-based justification for the principal findings.



### Mitigation and Enhancement Analysis

One of the main aims of the SIA is to identify and assess types of mitigation and enhancing (M&E) measures that might be used and ameliorate or enhance the significant impacts identified by the detailed assessment.

M and E measures include:-

- trade-related measures which might be built into a WTO agreement itself
- side or parallel agreements between WTO member countries or in regional agreements
- collaborative agreements between international organisations and relating to relationship between WTO agreements and other types of international agreements
- international and regional initiatives to promote technical cooperation and capacity building in developing countries
- measures by national governments

The SIA is expected to be of particular value to the EC in identifying measures in the fourth group, which it can incorporate into its development aid programme and those of member states, to assist developing countries' capacity to implement measures in the last group. Measures in this last group are the responsibility of individual governments, to ensure that the net benefit which they expect to gain from a trade agreement is distributed in such a way as to avoid adverse impacts on particular sections of society or the environment.

Causal chain analysis can be used to identify where M and E measures are likely to be cost-effective, feasible and effective in mitigating/enhancing impacts.

Assessment of 'best' M and E measures in each of the core economic, social and environmental themes should be introduced into the detailed assessment findings, as a modified scenario for the relevant trade measure.

### Monitoring and Post Evaluation

Where uncertainty in impact prediction remains relatively high, monitoring and evaluation can be an important mitigation and enhancing measure. It may include:

- monitoring implementation of the provisions of WTO agreements
- monitoring and ex-post evaluation of the sustainability impacts of the New Round agreements, as implemented
- contingency plans to mitigate impacts identified by monitoring.
- ex-post evaluation of the SIA studies

### **Assessment components and methods**

Scenario Analysis is used to identify those sets of trade policy measures which should be assessed, in order to indicate the likely effects of alternative agreements.

Country Grouping is necessary to identify the different impacts likely to occur in different types of country. For an SIA conducted for the European Union, the EU itself is one group. Developing countries and least developed countries receive particular attention in negotiations, and form two other important groupings. For the purpose of assessment it is also necessary to distinguish between net importers and net exporters for a particular product or service, and other country characteristics may also need to be taken into account.

Individual Country Analysis enables a detailed evaluation of causal chains in specific circumstances, and provides case studies of actual impacts. Countries chosen for this analysis give an indication of the impacts which may be expected in the country groups to which they belong.

Sustainability Indicators need to be selected with two purposes in mind. First, they serve to present the SIA findings in a readily assimilable form. For this purpose, nine core indicators (or themes) have been identified, three for each of the economic, social and environmental spheres. A specific impact identified in the assessment (e.g. a gender effect) is then reported under the appropriate theme (e.g. equity).





### Core indicators or themes

Economic	Social	Environmental
Real income	Poverty	Biodiversity
Fixed capital formation	Health & Education	Environmental quality
Employment	Equity	National resource stocks

Within each of the detailed SIA studies, a number of second tier indicators is identified, to give greater precision in the reporting of potential impacts. These may also be used for a second purpose, the subsequent monitoring of impacts. For the purpose of monitoring, indicators may be chosen according to those impacts identified in the assessment as being of particular importance.

### Significance Criteria

The significance of a predicted impact is judged according to the influence it is likely to have on the negotiation decision, ranked on a scale of:

- 0 non-significant,
- +/-1 marginally significant to the decision, but if the impact is negative, a potential candidate for mitigation
- +/-2 likely to be significant to the decision, and if negative, merits serious consideration for mitigation.

In judging significance, the following factors are taken into account:

- extent of existing economic, social and environmental stress, in affected areas
- direction of changes to base-line conditions
- nature, order of magnitude, geographic extent and reversibility / duration of changes
- regulatory and institutional capacity to implement M and E measures

Causal Chain Analysis is a central component of the assessment, providing an understanding of the cause-effect links between a proposed change in an existing trade agreement and its eventual economic, environmental and social impacts.

Modelling. The results of econometric modelling studies reported in the literature are an important source of information, and may be augmented by specific studies carried out within the assessment. Such models are however highly simplified representations of reality, which, along with limitations in data availability and reliability, requires that their results be used with caution.

Statistical Estimation tests for statistically significant relationships between parameters related to a proposed trade measure, generally from a wide range of countries. This can inform the assessment by giving a broad indication of possible causal links.

Case Studies of observed impacts provide a valuable source of information. It is however important to recognise that the observed effects may have multiple causes, which may entail a high degree of uncertainty in identifying those that can be ascribed to the trade measure being analysed.

Expert Opinion can make effective use of methods, knowledge and data already available, to help fill gaps in knowledge and data within the components of a comprehensive SIA methodology.

Data Sources suffer from limitations in availability, quality and compatibility. Available data will typically be an approximation of the ideal data requirements for SIA.

Consultation Arrangements. The Doha Ministerial Declaration provided a mandate for consultation, transparency and the effective participation of civil society in the negotiation process. This is implemented in the SIA methodology by incorporating consultation as an integral element of the process, through the use of stakeholders as experts through network established for the purpose, the publication of reports on the programme's website (<http://idpm.man.ac.uk/sia-trade>), the encouragement and analysis of comments received, and public meetings held at key stages of the process (after publication of scoping and mid-term reports).



## 2. Evaluating the effectiveness of national sustainable development strategies

Although sustainable development is widely regarded as requiring bottom-up processes, it is strongly conditioned by strategic policy-making at the national level. All aspects of development, from the local to the global, are both enabled and constrained by national policy decisions. Evaluation at the national level can therefore, in principle, make an important contribution to making sustainable development a practical reality.

The work described in this part of the paper was commissioned by the UK Department for International Development, in its role as lead institution of the OECD Development Assistance Committee's working group on national sustainable development strategies. The work is a contribution to the preparations for the 2002 UN World Summit on Sustainable Development, to be held in Johannesburg.

### International commitments to national sustainable development strategies

At the Rio earth summit of 1992 (the UN Conference on Environment and Development), governments committed to "adopt national strategies for sustainable development [which should] build upon and harmonise the various sectoral, economic, social and environmental policies and plans that are operating in the country" (Agenda 21).

At the UN General Assembly Special Session of 1997 (Rio plus 5), this commitment was confirmed, and a target date of 2002 was agreed (Rio plus 10) for introducing national sustainable development strategies.

In 1996 the OECD established a number of International Development Targets, one of which required that "there should be a current national strategy for sustainable development in the process of implementation in every country by 2005, so as to ensure that current trends in the losses of environmental resources are effectively reversed at both global and national levels by 2015".

The second part of this OECD target, for reversal of environmental losses, appears to offer promise of a practical goal against which national strategies might be evaluated. It has not however been adopted by the UN, and cannot be implemented without much more detailed international agreements. The target depends for example on a wide range of enforceable agreements under both the Climate Convention and the Biodiversity Convention, neither of which has reached this level of ambition. The OECD target cannot therefore be used as a basis of evaluation, and the prime definition of a national sustainable development strategy is the original one set out in Agenda 21.

### Objectives of the evaluation

The first step in developing the methodology was to clarify the objectives of the evaluation:

1. to provide a common framework by which countries can report progress in developing strategies for sustainable development
2. to help countries identify and address shortcomings in their own processes for achieving their own objectives for sustainable development

In the first instance, it was intended that evaluations be carried out by countries' own governments, with a view to identifying shortcomings in their sustainable development strategies and improving them. No framework currently exists for carrying out independent evaluations or peer reviews, but should such a framework be established, a similar approach may be used.

In this context, the assessment approach is designed to be supportive of national efforts, not critical of them.



### Characteristics of an effective national strategy for sustainable development

The Rio commitment requires national strategies for sustainable development to “build upon and harmonise the various sectoral, economic, social and environmental policies and plans that are operating in the country”. The OECD has reinforced this:

*Putting a sustainable development strategy into operation would, in practice, most likely consist of improving existing strategic planning processes and their co-ordination rather than establishing a new process. The latter is not recommended.*

OECD/DAC 2001

This was taken as a key requirement against which strategies would be evaluated.

Beyond this, both OECD and the UN have developed a number of principles which nssds should follow if they are to be effective:

### OECD and UN principles of strategic planning for sustainable development

OECD principles of strategic planning for sustainable development	UN principles for effective national sustainable development strategy
<ul style="list-style-type: none"> <li>• Consensus on long-term vision</li> </ul>	<ul style="list-style-type: none"> <li>• Shared strategic and pragmatic vision.</li> <li>• Link the short to the medium and long term.</li> <li>• Ensure continuity of the strategy development process.</li> </ul>
<ul style="list-style-type: none"> <li>• Comprehensive and integrated</li> </ul>	<ul style="list-style-type: none"> <li>• Integrated and balanced across sectors and territories.</li> </ul>
<ul style="list-style-type: none"> <li>• Targeted with clear budgetary priorities</li> </ul>	<ul style="list-style-type: none"> <li>• Set realistic but flexible targets.</li> <li>• Coherence between budget and strategy priorities.</li> </ul>
<ul style="list-style-type: none"> <li>• Based on comprehensive and reliable analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Anchor the strategy process in sound technical analysis.</li> </ul>
<ul style="list-style-type: none"> <li>• Incorporate monitoring, learning and improvement</li> </ul>	<ul style="list-style-type: none"> <li>• Build mechanisms for monitoring follow up, evaluation and feedback.</li> </ul>
<ul style="list-style-type: none"> <li>• Country-led and nationally-owned</li> </ul>	<ul style="list-style-type: none"> <li>• Nationally owned and country driven process.</li> </ul>
<ul style="list-style-type: none"> <li>• High-level government commitment and influential lead institutions</li> </ul>	<ul style="list-style-type: none"> <li>• Strong political commitment at the national and local levels.</li> <li>• Spearheaded by a strong institution.</li> </ul>
<ul style="list-style-type: none"> <li>• Building on existing processes and strategies</li> </ul>	<ul style="list-style-type: none"> <li>• Built on existing processes and strategies.</li> </ul>
<ul style="list-style-type: none"> <li>• Effective participation</li> <li>• People-centred</li> </ul>	<ul style="list-style-type: none"> <li>• Participatory and the "widest possible participation" ensured.</li> </ul>
<ul style="list-style-type: none"> <li>• Link national and local levels</li> </ul>	<ul style="list-style-type: none"> <li>• Link national and local priorities and actions.</li> </ul>
<ul style="list-style-type: none"> <li>• Develop and build on existing capacity</li> </ul>	<ul style="list-style-type: none"> <li>• Built on existing knowledge, expertise and capacity.</li> </ul>

There is a high degree of correlation between the OECD and UN principles, which between them provide a basis for evaluation. However, all the principles are strongly process-oriented. If the evaluation were based on these principles alone, it would be difficult to demonstrate that the strategy would be effective in delivering sustainable development objectives.

This potential difficulty was addressed at a conceptual level, by identifying two distinct requirements of an effective sustainable development strategy. The development must be sustainable, and the strategy must be effective.



### Principles of sustainable development as a process

*Development is a process. Sustainable development is distinguished from other forms of development by the nature of the development process.*

*An effective strategy for development must deliver clearly defined objectives for development, in the social, economic and environmental spheres.*

On this basis, the first of these requirements could be tested against the internationally agreed OECD/UN principles, while the second could be tested against nationally defined objectives for development. The evaluation need not itself define development objectives. It should instead check that such objectives have been defined, and that the strategy is likely to be effective in delivering them.

### The effectiveness of planning - comparison with ISO 9000

In order to evaluate whether the strategic planning process is likely to be effective in achieving its goals, reference was made to the international standard for quality systems, ISO 9000. This is a generic management standard, applicable to any set of management processes, which must be such as to achieve identified objectives

The combination of the OECD/UN principles for a sustainable development strategy, and the ISO criteria for an effective planning process, provided the conceptual framework within which a set of evaluation criteria could be developed.

### Development of criteria for an effective nssd

In order to undertake an evaluation of a nssd, a set of criteria were developed to ascertain whether the principles outlined above are complied with effectively. To do this, direct use was made of the Rio Declaration and Agenda 21, for further elaboration of the principles of sustainable development. Additionally, a literature review was carried out of previous evaluations or reviews of national sustainable development strategies. This provided many useful lessons from evaluation experience, both in the criteria used, and the methodology of the evaluation itself.

In developing the criteria, reference was again made to ISO 9000, as a well established example of how a relatively complex set of evaluation criteria can be set out. As in ISO 9000, a distinction was drawn between the criteria themselves, and guidance on their interpretation.

The various inputs were collated into five key themes:

#### Principles of sustainable development

- A. Integration of economic, social and environmental objectives
- B. Participation and consensus

#### Principles of strategic planning

- C. Country ownership and commitment
- D. Comprehensive and coherent policy process
- E. Targeting, resourcing and monitoring



A total of twenty criteria were developed, four for each theme. Each of these includes a number of requirements, as in the examples below.

**Example criteria: Principle A - integration of economic, social and environmental objectives**

*Criterion A1 - integration*

Strategic planning in the country is based on a comprehensive and integrated analysis of economic, social and environmental issues, which clarifies links between the three spheres, resolves conflicts between them where practicable, and negotiates appropriate trade-offs where conflicts remain.

*Criterion A2 - social and poverty issues*

Strategic planning in the country integrates poverty eradication, gender issues and the short- and long-term needs of disadvantaged and marginalised groups into economic policy.

*Criterion A3 - environmental and resource issues*

Strategic planning in the country integrates the maintenance of sustainable levels of resource use and the control of pollution to maintain a healthy environment into economic policy.

*Criterion A4 - international commitments*

Measures are in place to ensure compliance with international agreements which the country has entered into, on environmental and social issues.

Within the particular circumstances of any individual country, evaluators need to use a high degree of judgement in deciding what information they need to gather, and in deciding whether a criterion is met. To assist with this, a number of guidance questions were developed for each criterion, again drawing on the experience of previous evaluations.

**Example of guidance on application: criterion A1**

A1.1. Do planning systems define priorities in environmental, economic and social terms, and are these systems compatible with those for analysis and participation?

A1.2. Are there mechanisms for dealing with trade-offs between environmental, economic and social objectives, including poverty eradication and the needs of future generations?

A1.3. Do fiscal and regulatory incentive frameworks internalise external environmental and social costs in order to correct market or policy failure, without choking off desirable private investment? Are these frameworks adequately enforced?

do effective laws and regulations exist for building permits, land-use permissions, environmental standards, pollution permits and environmental impact assessment?

are appropriate non-compliance penalties enforced?

are there local environmental by-laws established by local consultative processes?

A1.4. Has action been taken to identify and adopt appropriate economic instruments and other policy tools to integrate economic, social and environmental objectives, including action to identify and undertake appropriate reforms to economic policies, resource allocation and property rights policies, and sectoral policies and practices for environmental protection, natural resource management and development?

**Reporting evaluation results for each criterion**

A simple scheme for reporting was devised, in order to ensure consistency between different evaluations. The evaluation of each criterion is reported as follows:

- A all of the requirements are fully met
- B all of the requirements are satisfactorily met, although some further improvements are desirable
- C some requirements have been satisfactorily or fully met, but others have not yet been satisfactorily met
- D few of the requirements have, as yet, been satisfactorily met



Further guidance was also developed on how evaluators should form judgements on whether a criterion is met satisfactorily. In particular, the assessment should take into account the following factors:

- is the action being taken relevant?
- is it appropriate?
- is it effective?

### Organisation of the Evaluation

The effectiveness of the evaluation depends critically on how it is carried out. Guidance was therefore prepared on undertaking an evaluation, based on the literature review, and on the project team's own experiences of conducting similar assessments. This is summarised below.

#### Evaluation methodology

*steering committee*

high ranking representatives of ministries most involved in strategic planning processes

*assessment team*

senior experts: economic, environmental, social

*desk study of key documents*

national strategies, budget policies, laws, state of the environment reports etc.

*interviews and discussion fora*

cover full range of stakeholders

*examination of further documents*

development plans, by-laws, planning procedures, monitoring data etc.

*supporting evidence*

departmental records, numbers and qualifications of responsible staff etc.

*attendance at planning meetings*

observe the process in action

*draft assessment report*

sent to all participants and published for comment,

*public consultations*

incorporate results and responses into final report

Much of this guidance is based on established practice for evaluating any type of planning process. The first item in the above table, the establishment of a high level steering committee to oversee the evaluation and approve the reports that are published, is considered critical for this type of assessment. The prime aim of the evaluation is to identify shortcomings within the country's strategic planning processes, such that they can be rectified. There is no international "policing" body which can enforce such action, and so the success of the process depends entirely on the commitment and determination of high level decision-makers to improve national planning processes. Without such commitment, the evaluation cannot succeed, and the goal of achieving sustainable development cannot be met.



## References

Full reports on both applications are available on the websites of the trade impact assessment programme and IDPM's Impact Assessment for Sustainable Development Unit.

Kirkpatrick C, Lee N, Curran J, Franklin J, George C and Nomura H (2002) *Further Development of the Methodology for a Sustainability Impact Assessment of Proposed WTO Negotiations* Institute for Development Policy and Management, Manchester, UK  
<http://idpm.man.ac.uk/sia-trade>  
<http://europa.eu.int/comm/trade/miti/envir/sia.htm>

Kirkpatrick C, George C and Curran J (2001) *Development Of Criteria To Assess The Effectiveness Of National Strategies For Sustainable Development* Institute for Development Policy and Management, Manchester, UK  
<http://idpm.man.ac.uk/iasdu>  
<http://www.johannesburgsummit.org> (Ghana experts' meeting)

The following are the principal citations from the above reports which have been used in this paper.

- Carew-Reid, J., Prescott-Allen, R., Bass, S. and Dalal-Clayton, B. (1994). *Strategies for National Sustainable Development. A Handbook for their Planning and Implementation*. ICUM, IIED, Earthscan.
- Chomo, G. and Farrentino, M. (1999) 'NAFTA environmental impacts on North American fisheries'. Paper presented at Conference on Trade, Poverty and the Environment: Methodologies for Sustainability Impact Assessment of Trade Policy. IDPM, University of Manchester. (<http://idpm.man.ac.uk/iasdu/confer.html>).
- Dalal-Clayton, B. and Bass S. (2000) *National Strategies for Sustainable Development: the Challenge Ahead*. IIED. London.
- DFID (2000) *Achieving Sustainability, Poverty Elimination and the Environment*. DFID. London.
- Earth Council (2000) *Guidelines for NCS Rio+10 Assessment*. Stockholm.
- European Commission (2001) *European Governance. A White Paper*, COM (2001) 428 final. Brussels.
- Fauchald, O.K. (2000) 'Assessing regulatory effects of new trade rules' in OECD *Assessing the Environmental Effects of Trade Liberalisation Agreements: Methodologies*, OECD, Paris.
- FFLA / WWF (2000) *The International Experts' Meeting on Sustainability Assessments of Trade Liberalisation. Full Meeting Report*. WWF. Gland.
- George, C. (2000) 'Environmental monitoring, management and auditing' in Lee, N. and George, C. (eds) *Environmental Assessment in Developing and Transitional Countries*. Wiley. London.
- George, C and Ahmad, B (2001) *Evaluation and Future Development of EIA Systems in Albania, Croatia, Egypt, Jordan, the Palestinian Authority, Tunisia and Turkey*. Mediterranean Environmental Technical Assistance Programme, University of Manchester, Manchester, and the World Bank, Washington DC
- IIED (1995) *Participation in Strategies for Sustainable Development*. International Institute for Environment and Development. London.
- IIED (2000) *National Strategies for Sustainable Development. A Guide to Key Issues and Methods for Analysis. A prompt for status reviews and dialogues*. Rolling Draft. London.
- IISD (2001) 'Summary of the International Forum on National Sustainable Development Strategies'. *Sustainable Developments*, November.
- Imbach, A., Dudley, E., Ortiz, N. and Sanchez, H. (1997) *Participatory and Reflective Analytical Mapping for Sustainability (PRAM)*. IUCN, Gland, Switzerland and Cambridge, UK.
- International Monetary Fund (2001) *Poverty Reduction Strategy Papers – Progress in Implementation*. April. IMF. Washington DC.
- International Waters Programme (2000), supported by the Global Environmental Facility (IWP/2000) 'A Review of the Transboundary Diagnostic Analysis (TDA) Approach to the Preparation of Strategic Action Programmes (SAPs)', Annex 8, *International Water Program Study*.
- ISO (2000) *International Standard for Quality Systems ISO 9001:2000* International Standards Organisation
- Jha, V., Markandya, A. and Vossenaar, R. (1999) *Reconciling Trade and the Environment: Lessons from Case Studies in Developing Countries*. Edward Elgar, Cheltenham.



- Kirkpatrick, C. and Lee, N. (1999) *WTO New Round: Sustainability Impact Assessment Study* (Phase Two Report) IDPM, University of Manchester.
- Kirkpatrick, C. Lee, N. and Morrissey, O. (1999) *WTO New Round: Sustainability Impact Assessment Study*. (Phase One Report). IDPM, University of Manchester.
- Maurer, C. (1999) *National Councils for Sustainable Development: Do they Matter?* World Resources Institute. Washington. D.C.
- McCulloch, N., Winters, L.A. and Cirera, X.C. (2001) *Trade Liberalization and Poverty: A Handbook*. DFID and Centre for Economic Policy Research. London.
- McGee, R. and Norton, A. (2000) *Participation in Poverty Reduction Strategies: A synthesis of experience with participatory approaches to policy design, implementation and monitoring*. IDS Working Paper 109. Institute of Development Studies. Brighton.
- OECD (1996) *Shaping the 21<sup>st</sup> Century: the Contribution of Development Co-operation*. OECD. Paris.
- OECD (2001) *Strategies for Sustainable Development: Guidance for Development Cooperation*. (The DAC Guidelines). OECD. Paris.
- OECD (2001a) *Strategies for Sustainable Development: Practical Guidance for Development Co-operation*. Development Assistance Committee. March. OECD. Paris.
- OECD (2001b) *Policy Guidance for Strategies for Sustainable Development*. Final Draft. Development Assistance Committee. OECD. Paris.
- Rubio, G. Prenzushi, G and K. Subbarao (2000) "Monitoring and Evaluation" in the Poverty Reduction Strategy Source Book. World Bank. Washington DC.
- Trachtman, J.P. (2000) 'Assessing the effects of trade liberalisation on domestic environmental regulation: towards trade environment policy integration ' in *OECD Assessing the Environmental Effects of Trade Liberalisation Agreements: Methodologies*. OECD, Paris.
- UN (2001) *Indicators of Sustainable Development: Guidelines and Methodologies*. Division for Sustainable Development. UN. New York.
- UNEP (2000) *Global Environmental Outlook*. London. Earthscan.
- UNEP (2001) *Reference Manual for the Integrated Assessment of Trade Related Policies*, Geneva.
- United Nations (1997) *Earth Summit + 5. Programme for the further Implementation of Agenda 21*. Special Session of the UN General Assembly, June. United Nations, New York.
- United Nations Conference on Environment and Development. (1992) *Agenda 21*. Chapter 40 Information for Decision-making. United Nations. New York.
- United Nations Division for Sustainable Development (2001) *From Theory to Practice: Indicators of Sustainable Development*. United Nations. New York
- United Nations Division for Sustainable Development (2001) *Indicators of Sustainable Development: Guidelines and Methodologies*. United Nations. New York.
- World Bank (1999) *Poverty Reduction Strategy Papers - Status and Next Steps, Revised*. World Bank, Washington DC.
- World Bank (2000) *Comprehensive Development Framework: Country Experience*. March 1999 – July 2000. Report for the Development Committee. CDF Secretariat. World Bank.
- World Bank (2001) *Tracking progress. Comprehensive Development Framework. Poverty Reduction Strategies*. World Bank, Washington DC.
- WTO (2001) *Ministerial Declaration*. Ministerial Conference Fourth Session. Doha 9-14 November. WT/MIN/(01)/DEC/W/1.





## **Professor Bob Evans and Dr Kate Theobald**

### **LASALA: SELF-ASSESSMENT OF LOCAL AGENDA 21 FOR LOCAL AUTHORITIES**

The research findings in this paper relate to the LASALA project which was co-funded by the European Commission's Directorate General for Research: Fifth Framework Programme, 1998-2002, thematic Programme: Energy, Environment and Sustainable Development, Key Action: "City of tomorrow and Cultural Heritage". Contract number: EVK4-CT-1999- 00011

#### **Abstract**

This paper reflects on the findings of a recently completed EU research project entitled LASALA or Local Authorities Self-Assessment of LA21, focusing in particular on the difficulties and benefits of policy evaluation in this area, which themselves are a reflection of the nature of the process. The objective of encouraging and supporting 'soft governance', the attempt to foster capacity building and the need to encourage a process of 'thinking globally' whilst 'acting locally' are neither easily quantifiable nor accessible. The paper provides an evaluation of the LASALA methodology for evaluating LA21 (and more broadly local sustainable development processes, and considers the opportunities for monitoring of sustainable development policies presented by this innovative policy evaluation tool.

#### **Evaluating policy processes and outcomes**

The evaluation of past policy is of crucial importance to the policy process at all levels of government. Only through a process of regular feedback and policy review resulting from an evaluation of policy implementation can policy instruments and approaches be fine-tuned to ensure that objectives are achieved. Moreover, objectives and policy priorities change over time, therefore policy evaluation needs to reflect the continual shift in these priorities. However, in the field of UK local environmental policy, there has not been a systematic and qualitative evaluation of the Local Agenda 21 initiative. There have been studies, for example by the UK's Local Government Management Board (LGMB, 1995; 1996; 1997; 1998; Tuxworth, 1996), which have sought to assess which local authorities have participated in the LA21 process, which officers and members have been principally responsible, and the extent to which LA21 strategies link to other local authority policies and plans. Other countries, for example the Netherlands and the Baltic States have conducted similar exercises (e.g. Joas, 2000) and there have been several studies of LA21 activity and progress across Europe most notably Lafferty & Eckerberg (1998); Lafferty, 2001 and the International Council for Local Environmental Initiatives (1998; 2002).

However, although these studies collectively represent a valuable source of information on the LA21 project, there has been only limited monitoring of LA21 to assess the extent to which declared policy objectives have been secured, and moreover, to investigate the political, socio-economic, administrative and cultural contexts in which local sustainable development processes are developing. Research conducted has been by those in the academic or policy community, rather than by local policymakers and stakeholder groups themselves.

This paper now considers the innovative policy tool of self-assessment, used for evaluating local sustainable development policies, developed within the framework of the European Union funded LASALA project (Local Authorities Self-Assessment of Local Agenda 21). This firstly necessitates a brief review of the European LA21 initiative, to provide the context for the self-assessment of LA21 across Europe.

#### **The Europe-wide Local Agenda 21 Initiative**

More than 1500 local authorities across Europe (figures correct at April 2002) have endorsed the need to prepare LA21s, through joining the European Sustainable Cities and Towns Campaign. The Charter for LA21 (also known as the Aalborg Charter, after the Sustainable Cities and Towns Conference which took place in Aalborg, Denmark, 1994), specifically refers to the need for local consensus between all stakeholder groups in the production of a local sustainable development plan, and sets out key principles for this.



These include extensive public consultation, the recognition of existing planning and finance frameworks, participation by all sectors of the local population, and long term local planning, supported by measurable targets. These ideas are reinforced by the International Council for Local Environmental Initiatives (ICLEI) 'model communities' programme. ICLEI (an association of local governments dedicated to the prevention and solution of local, regional and global environmental problems through local action) has set up a Local Agenda 21 Guidance and Training Programme, to support local authorities in developing and implementing LA21. Thus, much has been done, both by local government associations and by European organisations, in support of LA21. However, the focus has been on how to 'do' LA21, with comparatively little time spent on reflecting on the process and dealing with the problems experienced by local authorities across Europe in implementing policies for local sustainability. The LASALA research methodology presented in this paper is a reflection of the need to evaluate the process to date of LA21 and more broadly local sustainable development policies, across a range of national, regional, political and socio-economic contexts.

### **Local Authorities Self-Assessment of Local Agenda 21 (LASALA)**

LASALA was funded by the European Union, under the Fifth Framework 'Cities for Tomorrow' programme. It was a 20 month project (conducted between March 2000 and October 2001) with six partners, covering the whole of the continent of Europe, and with ICLEI as the co-ordinator.

To date over 2,000 European cities are engaged in a LA21 process of some description, and although there are distinctive national and local contexts, it is becoming clear that there are substantial commonalities in experiences of developing LA21 processes and policies. This paper argues that if the main principles of sustainable development in towns and cities - eco-efficient urban management, and new models of urban governance (the 2 themes of LASALA) are to be achieved, it is essential to evaluate past practice and performance in order to ensure future policy is directed in an effective manner.

Eco-efficient urban management is a collective term used to refer to the use of effective management instruments, processes and practices in the pursuit of urban environmental sustainability, in particular the thrifty use of natural resources. This approach to environmental sustainability requires integrated and holistic approaches to policymaking, which may involve the erosion of traditional administrative and professional boundaries. Such changes are necessary to take account of emerging mechanisms for formulating, implementing, and evaluating policies such as environmental capacity, sustainability indicators, environmental impact analyses and ecological footprinting.

The adoption of new models of urban governance emphasises the importance of developing approaches to 'soft governance' which include those sectors of civil society that have been largely excluded from policy decisions. A particular aim of LASALA was therefore to evaluate whether new models of urban governance, for instance in the form of improved participatory mechanisms, inclusion of citizens, and the continuous flow of information between local authorities and 'stakeholder' groups can provide improved conditions for effective, eco-efficient city planning and resource management systems.

Before exploring the tool of self-assessment as applied to LA21 processes, it is appropriate to set out the aims and objectives of LASALA. The LASALA project had a number of key objectives, both for local authorities themselves in the development of their local sustainability processes, and for the development of LA21 across Europe.

The main aims of LASALA were to:

- i) provide a framework to help LA21 co-ordinators with reporting to their Local Agenda 21 Committee (or other appropriate sustainability Committee) on progress with local sustainability;
- ii) assist local authorities in reflecting on their Local Agenda 21/other local sustainability process and identifying problems and opportunities;
- iii) provide feedback to all local stakeholders as part of a regular evaluation of LA21
- iv) assist in the creation of new ideas for the development of the LA21 process.



In Europe, within the framework of the European Sustainable Cities and Towns Campaign (to which the development of the self-assessment tool is directly linked), the aim of LASALA was to:

- i) provide a common framework for the evaluation of Local Agenda 21 and more broadly of local sustainability across Europe, which can be fed into the Rio+10 process;
- ii) help establish a monitoring and evaluation process on progress with local sustainability processes throughout Europe;
- iii) assess progress and outcomes of the implementation of the Aalborg Charter commitments and additional health/social and economic issues;
- iv) enable the identification of 'good practice' across a diverse range of municipalities, and the dissemination of good practice to local authorities across Europe;
- v) provide feedback into the European Sustainable Cities and Towns Campaign and other international LA21 evaluations and help in their future development.

The project involved the recruitment of approximately 230 local authorities from across Europe, to take part in a self-assessment of their LA21 processes. Initially the 800+ signatory authorities to the Aalborg Charter (the number of signatories in March 2000) were contacted, to request their participation. This initial contact was supplemented by intensive networking with national local government organisations to encourage participation from local authorities, who, whilst not being signatories to the Aalborg Charter, did nevertheless have a LA21 process underway.

### **Self assessment through the World-Wide Web**

A self-assessment module, with supporting information and guidance materials was set up by ICLEI on the web, to provide distance training for those local authorities participating in the LASALA self-assessment. The module used for self-assessment was based on a pilot module developed by ICLEI during 1999 to assist local authorities in evaluating their LA21 strategies. It was however revised substantially to meet the requirements of the LASALA project - specifically to deal with the demands of a large-scale, pan-European programme, and to incorporate the qualitative elements that represented a key element of this evaluation. The utilisation of such a distance-training tool, supported by 'tele-guidance' from each of the partners in their respective regions, represents a new and innovative approach to evaluating LA21 for two principal reasons. First, it provides a comprehensive training manual on-line, for local authorities to access and utilise, both for the purposes of the LASALA evaluation, and for regular monitoring of their LA21/local sustainability process. Secondly it supports local authorities in utilising web-based resources for the evaluation of policy processes.

### **The Self-Assessment Module**

The LASALA Self-Assessment Module therefore contains two Exercises:

#### **Exercise One**

The first Exercise is in the form of a quantitative questionnaire, for the LA21 Co-ordinator or equivalent to complete (in conjunction with colleagues). This Exercise requires information relating to achievements in the local sustainability process, and is in two Sections. Section 1 focuses on both baseline, contextual information on a local authority and on the progress it has achieved in developing its Local Agenda 21/local sustainability process. Section 2 asks questions relating to a local authority's LA21 progress and process vis-à-vis a number of sustainability criteria (based on the 13 Aalborg Charter Commitments).

It was therefore intended that Exercise 1 provide both the local authority and the LASALA project team with a basic knowledge and understanding about the steps a local authority has taken in order to implement sustainable development policies, through, for example, Local Agenda 21 processes and other sustainability policies. It was also intended to assist in developing an understanding of the obstacles to, and potential for a local authority (town or city) to progress towards sustainable urban development.

A further aim was to learn about the role and capacity of a local authority within the wider community, that is the political, ecological, social, and economic base of a local authority, as well as the level of autonomy it enjoys. The information gained from Exercise One permitted the assessment of the overall level of progress towards sustainable development achieved by different towns and cities, taking account of the particular context in which they function. Underlining this was a recognition by the research team that there is no single correct way to develop a LA21 or local sustainability process.



## Exercise Two

The second Exercise was also in the form of a questionnaire, but based solely on the 13 Aalborg Charter Commitments, and responses were required from stakeholders representing different sectors of the local community, including business, education institutions, community groups, utilities (gas, electricity, water), health authorities, trade unions, and the police. This Exercise was particularly important in the evaluation of a LA21 process, as it had to be completed through the conducting of a workshop, (this could be the LA21 Forum or other pre-existing LA21 grouping, or a workshop designed specifically for the LASALA evaluation) involving these stakeholder groups.

Stakeholders, led by a facilitator, were required to discuss the questions in the Exercise, under the headings for each of the 13 Aalborg Commitments. This part of the self-assessment process provided the qualitative data for the evaluation, and was cross-referenced with the Exercise One responses.

The findings of this self-assessment process were therefore intended to be of value to the local authorities themselves, in assessing the effectiveness of their LA21 policies and processes, and have been presented (in aggregate form) in the LASALA Evaluation report (Evans and Theobald, 2001a). A second report identifying cases of 'good practice' in 24 of the participant local authorities has also been published (Joas, Gronholm & Matar 2001).

## Evaluating the Self-Assessment Method

One important feature of the LASALA project was the evaluation of the self-assessment method as a tool for monitoring local sustainability processes<sup>112</sup>. To this end, a feedback questionnaire was devised, for participating local authorities to complete once they had returned both Exercises (or at the minimum Exercise One). The responses from the feedback questionnaires form the basis for the following section. In addition, each of the project's regional co-ordinators were required to maintain contact with the participating local authorities, and to maintain a record of comments on the self-assessment method. Where appropriate, these comments are also included. The key areas focused on for this evaluation of the method were the use of the LASALA website and distance training manual, and the experiences of local authorities in conducting the two Exercises. A broader issue of central importance to the project was the extent to which local authorities felt self-assessment as a method was useful to them in reflecting on past and current policy, and in developing future policies for sustainable development.

## Local Authority Experiences with LASALA

The approach offered by the self-assessment method (SAM) was seen as beneficial both for individual local authorities but also for providing information to localities on how other local authorities (whether neighbouring, in the same country, or in another part of Europe) were developing their LA21s:

**'We joined the LASALA process because we felt that comparative self assessment was an appropriate way to evaluate our LA21 process'.**

Many of those participating in the LASALA project commented that the LASALA project had proved helpful in developing a local network for LA21s, through local authorities establishing or strengthening links with neighbouring municipalities. This related in particular to the conducting of the workshop for Exercise 2 – in a number of cases neighbouring local authorities assisted each other by providing facilitators for the workshops. This is clearly a positive outcome of this approach to policy evaluation and highlights the importance of local authorities working with and learning from other municipalities in addressing the requirements of local sustainable development.

The use of the SAM was particularly important for identifying both local authority and stakeholders perspectives on the nature of their LA21 processes and on what had been achieved. A view shared by many local authority officers and members (and summed up in the following response) was that:

---

<sup>112</sup> See Evans, B & Theobald, K (2001b) *Evaluation of the LASALA Self –Assessment Method*, ICLEI, Freiburg. Available at <http://www.iclei.org/europe/lasala>



**'The results of this exercise may help me to promote the benefits of genuinely empowering the community'**

Clearly this showed a recognition by many local authorities that both the process of self-evaluation, and the results from this would be useful in the future in developing more effective participatory processes.

Even for those local authorities that were only at the beginning of a process of moving towards local sustainable development, the evaluative workshop was perceived as beneficial in highlighting the complex process of incorporating sustainability thinking into all areas of policymaking:

**'It did help to focus thinking on some areas which we had not covered in detail, and also reinforced the thinking that there is still much to do before Sustainable Development becomes mainstream in policy making'.**

It was noted in a number of cases that the conducting of a workshop for Exercise 2 had assisted stakeholders in evaluating their local authority's approach to Local Agenda 21. It was regarded as beneficial to use Exercise 2 as a starting point, to encourage debate between stakeholders and the local authority about the development of an LA21 process, and more broadly about the concept of sustainable development. The method of self-assessment was also perceived as appropriate to link with other consultation mechanisms within local authorities.

Participating in the self-assessment offered by LASALA helped in many cases to bring LA21 back to prominence within a local authority policy process, and to assist local authorities in reflecting on their central role in moving towards local sustainable development. One comment sums up this achievement:

**I have been able to convince the council to sign the Aalborg Charter, introduce a policy on Fair Trade and commence work on an Environmental System so even a questionnaire has far reaching consequences'.**

There was a concern by some respondents that in general local authorities would tend to 'talk up' their achievements regarding LA21. However, as noted in the LASALA Evaluation Report (Evans and Theobald, 2001), it was equally common for LA21 officers to 'talk down' their achievements and to be self-critical. The central point of the LASALA method, was that it was self-evaluation *for* local authorities, *by* local authorities. Clearly it was important for local authorities to present an 'honest' picture of their LA21 process, as this would be valuable not only for the LASALA project findings, but for their own assessment of the progress and process of LA21.

### **Comments on Exercise 2**

As noted above, Exercise 2 was based on the 13 Aalborg Commitments, and one issue raised by those taking part in Exercise 2 (both local authorities and other stakeholders) was the difficulty in understanding the complex language used in these Commitments. This is clearly an issue that needs to be taken into account, both in terms of the responses to the questions, but also for the possible revision of the self-assessment method for future use by local authorities across Europe. Nevertheless, the Aalborg Commitments provide a useful framework for assessing the level of political commitment of authorities to local sustainable development in their municipalities.

As noted above, the conducting of a workshop for Exercise 2 was intended to encourage local authorities to involve a range of stakeholders from different sectors of the local community. Responses in the feedback questionnaire did however point to the difficulty of involving some sectors, particularly in terms of the time and resource constraints for individuals to attend. However, a further obstacle to attendance was a lack of understanding both on the part of local authority representatives and other stakeholders, of their relevance to, and contribution to, the LA21 process. Local authorities were very aware of the importance of involving both organisations and individual citizens in the decision-making processes required to implement LA21 principles. However it was widely acknowledged that there are a number of excluded groups that are difficult to access. Those individuals taking part in the workshops for Exercise 2 tended to be those identified by local authorities as being the most active in the LA21 Forums. A further barrier to involvement was a lack of contact with stakeholder groups. This lack of contact was suggested by one respondent to be indicative of an institutional barrier, in terms of communication between local authorities and the local community.



These comments can only serve to emphasise the relevance and importance of Exercise 2, and the participation of stakeholder groups, in evaluating a local authority's Local Agenda 21 process.

### ***Benefits of using the internet for distance training***

The utilisation of the LASALA distance-training tool, through the LASALA website, has represented a new and innovative approach to evaluating LA21, and has indeed been recognised by participating local authorities as an effective approach to assessing both the progress and process of their local sustainable development programmes. The feedback from local authorities, through the feedback questionnaire, and from informal discussions between participating local authorities and LASALA regional co-ordinators, emphasised the need for a continuation and expansion of a web-based approach to regular self-assessment of LA21 and more broadly local sustainable development policies. Moreover, the central support mechanism of ICLEI was clearly identified in this feedback as important to the self-evaluation approach.

### **Linking self-assessment with policy development**

It is appropriate at this stage to consider, based on LASALA findings, the benefits of self-assessment as a tool for evaluating both the process and progress of policymaking for local sustainable development. As noted earlier in this paper, the policy process and policy outcomes are inextricably linked, and evaluation of public policy should reflect this link. The key issues arising from the LASALA self-assessment method, and relevant both from the perspective of participant local authorities, and for future research, are as follows:

- Overall, there still appears to be an almost complete absence of formal evaluation mechanisms in local authorities, whether in relation to specific performance evaluation of the achievements of LA21 itself, or more broadly, across policy structures. Policymaking tends to be short-term in nature and many LA21 programmes do not have a long-term policy programme (although the overall 'vision' is often longer).
- Many LASALA respondents have been prepared to be critical and reflective of their LA21 process and of the administrative system in which this operates, although it is acknowledged by the researchers that the LA21 co-ordinators view does not necessarily represent that of other officers and members. A number of respondents have commented on the value of evaluating a process, and on having the tool of LASALA to do this.
- Self-evaluation has exposed or confirmed the complexities and problems for local authorities in attempting to involve a range of stakeholder groups in policymaking. Moreover, there is a concern amongst some stakeholders (and this is recognised by LA21 co-ordinators) that they may merely be legitimising decisions, rather than possessing the power to challenge or change decisions.
- The evaluation has nevertheless also highlighted areas where local 'partnerships' for sustainability are being developed, and in some cases are proving beneficial for policymaking on sustainable development. The reporting of this through the LASALA findings, in particular through the identification of 'good practice' examples (see Joas et al 2001) and through benchmarking reports produced for each participating local authority has been valuable in informing policy and practice in local authorities across Europe.

### **The Future Use of Self-assessment**

According to many local authority respondents, LASALA has been an extremely useful exercise in terms of setting the agenda for the development of LA21 processes, and in terms of highlighting the need for progress in a range of areas. It has also highlighted the importance of self-assessment of local sustainability processes on a regular basis as an integral part of policy development. The self-assessment mechanism has assisted local authorities in being open and self-critical in their evaluation, and in communicating their findings to other local authorities – a sharing of information and of 'what works'. This networking of experiences could be particularly important in developing and improving policies for sustainable development, and is an aspect of self-assessment that requires support and investment by the appropriate local government organisations. This is already being addressed at a European level through the support from network organisations such as ICLEI and the European Sustainable Cities and Towns Campaign.



## Conclusions

The results from the self-assessment exercise have fully justified the project's emphasis upon co-operative and reflexive working with local authorities in the self-assessment of LA21 and local sustainability processes. The extremely high quality of the responses to both of the Exercises indicates the validity of the self-assessment approach and of the supportive and reflexive approach adopted by the project team.

The LASALA experience suggests that self-assessment by local authorities of their LA21 processes is extremely beneficial and can provide an opportunity for critical reflection on current and past policies relating to local sustainable development. LASALA has provided a broad analysis of the process of LA21 across a geographically diverse area, and has permitted an exploration of the complex contexts within which local authorities respond to the demands of sustainable development. Self-assessment is an effective tool with which local authorities have been able to consider the wider governance issues that are integral to sustainable development, such as the nature of decision-making within local government, and the development of partnerships with different sectors of civil society.

LASALA has shown that local authorities recognise the importance of participating in a self-assessment exercise if this is effectively organised and supported, both within their own organisations, and through networks such as ICLEI and the Campaign. It has also highlighted the role that information technology can play in assisting local authorities with self-evaluation, through a distance training technique. Moreover, LASALA has led to an extensive database on the process and progress of LA21 across Europe, to which it may be possible for all European local authorities to contribute through undertaking the self-assessment exercise.

## References

- Evans, B & Theobald, K (2001a) *Accelerating Local Sustainability: Evaluating European Local Agenda 21 Processes. Report of the LASALA Project Team*, ICLEI, Freiburg
- Evans, B & Theobald, K (2001b) *Evaluation of the LASALA Self –Assessment Method*, ICLEI, Freiburg. Available at <http://www.iclei.org/europe/lasala>
- International Council for Local Environmental Initiatives 1998 *Evaluation of Agenda 21: Local Agenda 21 (1997/98)* (ICLEI European Secretariat: Freiburg, Germany). Published 2000.
- International Council for Local Environmental Initiatives 2002 *Second Local Agenda 21 Survey Report* (ICLEI European Secretariat, Freiburg, Germany). (The Survey was conducted in collaboration with the UN Secretariat for World Sustainable Development, and UN Development Programme (Capacity 21)).
- Joas M (ed.) 2000 *Local Agenda 21 – Models and Effects: an analysis of LA21 activities in Finland and the Baltic Sea Region*. Abo Akademi: Turku, Finland.
- Joas, M, Gronholm, B & Matar, T (2001) *Identification of Good LA21 Processes*. Report of the LASALA Project Team, ICLEI, Freiburg
- Lafferty W M, and Eckerberg K. (eds) 1998 *From the Earth Summit to Local Agenda 21*. Earthscan: London.
- Lafferty, W.M. (2001) (eds) *Sustainable Communities in Europe*, Earthscan, London
- Local Government Management Board 1995 *Local Agenda 21 Survey 1994/95*. LGMB: London.
- Local Government Management Board 1996 *Local Agenda 21 Survey*. 1996 LGMB: London.
- Local Government Management Board 1997 *Local Agenda 21 UK Review 1992-1997*. LGMB: London.
- Local Government Management Board 1998/Brunel University *Local Agenda 21 Survey*. LGMB: London.
- Tuxworth B. 1996 *From Environment to Sustainability: surveys and analysis of Local Agenda 21 process development in UK local authorities*. *Local Environment* 1, 3: 277-297.



# EASY-ECO

## EvAluation of Sustainability

### European COncferences

#### OVERVIEW

[EASY-ECO-1](#)
[COUNTRY SURVEYS](#)
[EASY-ECO-2](#)
[NETWORK](#)
[news](#)

#### The Problem

Evaluation of Sustainability of projects, programmes and policies is gaining more and more importance. On the one hand, the legitimization of public spending is becoming more important, on the other hand „Sustainability“ is gradually turning from a model into a political framework conditions. But the involved disciplines (social, economic and environmental sciences) look at „Sustainability“ in different ways. There is still no generally accepted set of indicators and methods as well as standards used. And training possibilities for evaluators of sustainability is limited within Europe. Due to the severe lack of transparency in the European market it is rather difficult for those who commission evaluations to judge the quality both of evaluations and evaluators.

#### The aims of EASY-ECO

The aim of EASY-ECO, supported by the European Commission, is to give an overview of all methods for the Evaluation of Sustainability currently used within Europe. Based on that EASY-ECO aims to increase the quality of the Evaluation of Sustainability substantially.

Within EASY-ECO young researchers have the opportunity to gain expert knowledge in their home country by carrying out a country survey on the state of the art of Evaluation of Sustainability, supported by high level scientists. The scientific results will finally be discussed with commissioning agents, evaluators and the wider public. Thus

- an increase of market transparency
- an international and interdisciplinary exchange of experience
- the professionalisation of commissioning agents
- the institutionalisation of evaluations shall be gained.

The Conference Series consists of a European Workshop EASY-ECO 1, May 23 - May 25, 2002, Vienna, and a European Conference EASY-ECO 2, taking place in spring 2003 in Vienna.

#### Problemstellung

Evaluierungen der Nachhaltigkeit von Projekten, Programmen und Politiken gewinnen zunehmend an Bedeutung. Auf der einen Seite wird die Legitimation öffentlicher Aufwendungen immer wichtiger, auf der anderen Seite wird Nachhaltigkeit vom Leitbild zur politischen Rahmenbedingung. Der Begriff Nachhaltigkeit wird jedoch von den befassten Wissenschaftsdisziplinen (Sozial-, Wirtschafts- und Umweltwissenschaften) unterschiedlich interpretiert. Für Evaluationen von Nachhaltigkeit gibt es noch keine allgemein anerkannten Indikatoren und Methoden, wenig verbreitete Standards und innerhalb von Europa und eklatanten Mangel an Ausbildungsmöglichkeiten. Darüber hinaus sind besonders für Auftraggeber die Qualität von Evaluationen schwer bewertbar, und die Auswahl von geeigneten Evaluatoren schwierig, da sich der Europäische Markt als sehr intransparent erweist.

#### Ziele von EASY-ECO

Ziel der durch die Europäische Kommission unterstützten Konferenzserie EASY-ECO ist es, einen Überblick über die in Europa derzeit eingesetzten Methoden zu Evaluation von Nachhaltiger Entwicklung zu schaffen. Darauf aufbauend soll die Konferenzserie europaweit einen maßgeblichen Beitrag zur Verbesserung der Qualität von Evaluationen Nachhaltiger Entwicklung leisten.

Zum einen wird Nachwuchsforschern im Rahmen von EASY-ECO die Möglichkeit geboten, sich unter intensiver fachlicher Betreuung durch hochrangige Wissenschaftler im Rahmen von Länderstudien weiterzubilden. Zum anderen werden die wissenschaftlichen Erkenntnisse mit Evaluatoren und Auftraggebern von Evaluationen diskutiert und der breiten Öffentlichkeit zugänglich gemacht. Damit sollen

- die Markttransparenz gesteigert ,
- internationaler und interdisziplinärer Erfahrungsaustausch gefördert,
- die Professionalisierung der Auftraggeber erhöht,
- die Institutionalisierung von Evaluationen gefördert werden.

Die Konferenzserie besteht aus einem European Workshop EASY-ECO 1, der von 23. - 25 Mai 2002 in Wien stattgefunden hat und einer European Conference EASY-ECO 2 im Frühjahr 2003 ebenfalls in Wien. Da es sich um eine internationale Konferenzserie mit Englisch als Konferenzsprache handelt, finden Sie alle weiteren Informationen auf Englisch.





# EASY-ECO

## EvAluation of Sustainability

### European COnferences

[OVERVIEW](#)
[EASY-ECO-1](#)
[COUNTRY SURVEYS](#)
[EASY-ECO-2](#)
[NETWORK](#)
**programme**
[photos](#)

## Programme and Papers of EASY-ECO 1 Vienna, May 23 - 25, 2002

### May 23, 2002

10:00 – 12:00		<b>Welcome Addresses</b>
	<b>Uwe Schubert</b>	project coordinator, Vienna University of Economics and Business Administration
	<b>Zé Alves-Pereira</b>	<a href="#">European Commission</a>
	<b>Elisabeth Freytag</b>	Austrian Federal Ministry of Agriculture Forestry Environment and Water Management
	<b>Gunther Maier</b>	Vice-Rector for Scientific Affairs, Vienna University of Economics and Business Administration
	<b>Ursula Kopp</b>	Administrative Announcements, Vienna University of Economics and Business Administration
13:00 – 13:30		<b>Questions of Young Researchers</b>
13:30 – 15:30		<b>MARKET DEVELOPMENT</b>
	<b>Uwe Schubert, A</b>	<a href="#">Evaluation of Sustainability – new challenges for evaluators and methods of evaluation</a>
	<b>André Martinuzzi, A</b>	<a href="#">The market for Evaluation of Sustainability – Trends and Institutionalisation</a>
	<b>Tomasz Zylicz, PL</b>	<a href="#">The Role of Science in Sustainability Policies</a>
	<b>Ernst Max Nielsen, DK</b>	Sustainability in CEE - <a href="#">presentation</a> - <a href="#">notes</a>
16:00 - 17:00		<b>STANDARDS AND METAANALYSIS</b>
	<b>Thomas Widmer, CH</b>	<a href="#">Evaluation Standards: Assuring Quality in Evaluation</a>
	<b>Astrid Kuffner, A</b>	<a href="#">Evaluation of Sustainability? – Lessons learned from a Survey of Evaluation Reports</a>
17:00 - 17:30		<b>Preparation of WS I and II</b>
18:30 - 20:00		Special Guided Tour through the Inner City of Vienna

### May 24, 2002

9:00 – 11:30		<b>EVALUATION</b>
	<b>Jon Lovett, UK</b>	<a href="#">Ecological Sustainability: A Policy Oxymoron?</a>
	<b>William Lafferty, NO</b>	<a href="#">Political Evaluation: Premises, Approaches and Methods</a>
	<b>Wolfgang Meyer, D</b>	<a href="#">Sociological Evaluation - Methods and Limits</a>

	<b>Clive Spash, UK</b>	<a href="#">Valuing Society and Environment: Economic Methods and Limits</a>
	<b>Jürgen Freimann, D</b>	<a href="#">Evaluating the Impacts of Corporate Environmental Management Systems. A Comparison between EMAS and ISO 14.001</a>
12:00 - 13:15		<b>SUSTAINABILITY</b>
	<b>Edwin Zaccai, B</b>	<a href="#">From Conceptions of Sustainability to Indicators</a>
	<b>Thiemo Eser, D</b>	Evaluation and the Appropriate Policy Level under the Demands of Sustainability
	<b>Markus Langer/Aloisia Schön, A</b>	<a href="#">An Integrated Referential Framework for Sustainable Development</a>
14:30 - 18:30		<b>Workshop 1 - 3 parallel sessions</b>
19:00		Dinner at a typical Viennese Heurigen

**May 25, 2002**

---

9:30 – 12:30		<b>FIELDS OF APPLICATION</b>
	<b>Maria-Angeles Diez, E</b>	<a href="#">Current Approaches used in the Evaluation of Regional Policies: a Critical Review</a>
	<b>Alain Thierstein, CH</b>	<a href="#">An Evaluation Framework for the Quality Management of Projects on Sustainable Development</a>
	<b>Diane Warburton, UK</b>	<a href="#">Evaluating Participation in Sustainable Development</a>
	<b>Clive George, UK</b>	<a href="#">Sustainability Impact Assessment - (presentation)</a>
	<b>Kate Theobald, UK</b>	<a href="#">LASALA: Self-assessment of Local Agenda 21 for Social Authorities</a>
13:30 - 15:00		<b>Workshop 2 - 3 parallel sessions</b>
15:30 - 16:30		<b>Presentation of Conference Results</b>



# EASY-ECO

## EvAluation of Sustainability

### European COnferences

[OVERVIEW](#)

[EASY-ECO-1](#)

[COUNTRY SURVEYS](#)

[EASY-ECO-2](#)

[NETWORK](#)

[programme](#)

[photos](#)

Photos taken at EASY-ECO 1 May 23 until May 25, 2002 in Vienna.

Speakers

Participants

Workshops

Heurigen





# EASY-ECO

## EvAluation of Sustainability

### European COnferences

[OVERVIEW](#)
[EASY-ECO-1](#)
**COUNTRY SURVEYS**
[EASY-ECO-2](#)
[NETWORK](#)
[forum](#)
[training material](#)
[guideline](#)
[information material](#)

**COUNTRY SURVEYS** on the **STATE OF THE ART OF EVALUATION OF SUSTAINABILITY** will be carried out and presented at EASY-ECO 2 by

Country	Name	City	Contact
<b>AUSTRIA</b>	<b><u>LANGER Markus</u></b>	Vienna	<a href="mailto:markus.langer@sustainablemanagement.com">markus.langer@sustainablemanagement.com</a>
	<b><u>SCHÖN Aloisia</u></b>	Vienna	<a href="mailto:aloesia.schoen@wu-wien.ac.at">aloesia.schoen@wu-wien.ac.at</a>
<b>BELGIUM</b>	<b><u>BAULER Tom</u></b>	Brussels	<a href="mailto:tbauler@ulb.ac.be">tbauler@ulb.ac.be</a>
	<b><u>DAMIÃO Tito</u></b>	Brussels	<a href="mailto:titodamiao@hotmail.com">titodamiao@hotmail.com</a>
<b>CZECH REPUBLIC</b>	<b><u>CAPEK Ondrej</u></b>	Prague 5	<a href="mailto:ocapek@vse.cz">ocapek@vse.cz</a>
	<b><u>SCASNÝ Milan</u></b>	Prague 5	<a href="mailto:milan.scasny@czp.cuni.cz">milan.scasny@czp.cuni.cz</a>
<b>ESTONIA</b>	<b><u>KOPPEL Katrin</u></b>	Tallinn	<a href="mailto:e01kok01@student.ceu.hu">e01kok01@student.ceu.hu</a>
<b>GERMANY</b>	<b><u>GESSNER Christian</u></b>	Witten	<a href="mailto:christian.gessner@uni-wh.de">christian.gessner@uni-wh.de</a>
	<b><u>HASANAGAS Nikolaos</u></b>	Göttingen	<a href="mailto:nhasana@gwdg.de">nhasana@gwdg.de</a>
<b>GREECE</b>	<b><u>LOUPASIS Stylianos</u></b>	Athens	<a href="mailto:loupasis@hotmail.com">loupasis@hotmail.com</a>
<b>HUNGARY</b>	<b><u>BARSI Orsolya</u></b>	Budapest	<a href="mailto:e01kok01@student.ceu.hu">e01kok01@student.ceu.hu</a>
	<b><u>GRNYA Ildiko</u></b>	Budapest	<a href="mailto:e01gri01@student.ceu.hu">e01gri01@student.ceu.hu</a>
	<b><u>ZSENI Anikó</u></b>	Győr	<a href="mailto:zseniani@sze.hu">zseniani@sze.hu</a>
<b>LATVIA</b>	<b><u>AIZSALNIEKS Janis</u></b>	Riga	<a href="mailto:jansf@mail.lv">jansf@mail.lv</a>
<b>PORTUGAL</b>	<b><u>ABREU Isabel</u></b>	Lisbon	<a href="mailto:isabreu@dgsaude.min-saude.pt">isabreu@dgsaude.min-saude.pt</a>
<b>SPAIN</b>	<b><u>IZQUIERDO Beatriz</u></b>	Bilbao	<a href="mailto:b_izquierdo@teletel.es">b_izquierdo@teletel.es</a>
	<b><u>RAMOS-MARTIN Jesus</u></b>	Barcelona	<a href="mailto:jesusramosmartin@yahoo.es">jesusramosmartin@yahoo.es</a>
<b>UNITED KINGDOM</b>	<b><u>GRACE Matt</u></b>	Peterborough	<a href="mailto:matt.grace@scottwilson.com">matt.grace@scottwilson.com</a>
<b>USA</b>	<b><u>STRUHKAMP Gerlinde</u></b>	St. Paul, MN	<a href="mailto:Gerlinde.Struhkamp@gmx.de">Gerlinde.Struhkamp@gmx.de</a>

All training material and guidelines necessary for these surveys will be available on this site within June.

Intermediate reports and the final surveys will be issued here.



# EASY-ECO

## EvAluation of Sustainability

### European COnferences

OVERVIEW

EASY-ECO-1

COUNTRY SURVEYS

EASY-ECO-2

NETWORK

[forum](#)[training material](#)**guideline**[information material](#)

#### WORKING PHASES FOR THE COUNTRY SURVEYS ON THE EVALUATION OF SUSTAINABILITY

PHASE 1:	<b>Orientation Phase</b>	Milestone 1: September 5, 2002	<a href="#">GUIDELINE A</a>
PHASE 2:	<b>Collection and Screening of Evaluation Reports</b>	Milestone 2: September 30, 2002	<a href="#">GUIDELINE B</a>
PHASE 3:	<b>Analysis Phase</b>	Milestone 3: December 1, 2003	GUIDELINE C
PHASE 4:	<b>Submission of Country Survey</b>	Milestone 4: February 1, 2003	PAPER
PHASE 5:	<b>Paper Presentation</b>	Milestone 5: March/April 2003	FINAL PAPER

PHASE 1:	<b>Orientation Phase</b>
<b>GOAL</b>	Identify the "hot spots" and institutional structure of Evaluation of Sustainability in your country.
<b>CONTENT</b>	<ul style="list-style-type: none"> <li>• EASY-ECO 1: papers, presentations and workshops</li> <li>• press release</li> <li>• expert interviews (25 - 40)</li> <li>• develop a COUNTRY PROFILE together with the other Young Researchers in your country</li> <li>• start of collection of evaluation reports</li> </ul>
<b>YOUR BENEFIT</b>	<ul style="list-style-type: none"> <li>• you will find other experts and relevant evaluation reports easily</li> <li>• you can make yourself known</li> <li>• get an overview of what is happening within your country</li> <li>• COUNTRY PROFILE = basis for your paper</li> <li>• you can compare your country profile with other countries</li> </ul>
<b>TOOLS</b>	<ul style="list-style-type: none"> <li>• structure of the COUNTRY PROFILE as pdf = <a href="#">GuidelineA</a></li> </ul>
<b>DELIVERABLES</b>	one COUNTRY PROFILE per country of 5 - 8 pages
<b>TIME NEEDED</b>	5 - 10 days estimate per person
<b>MILESTONE 1</b>	<b>September 5, 2002:</b> Please send your COUNTRY PROFILE to <a href="mailto:easy@wu-wien.ac.at">easy@wu-wien.ac.at</a> by then.

PHASE 2:	<b>Collection and Screening of Evaluation Reports</b>
<b>GOAL</b>	Collect 30 - 50 evaluation reports (per person) in your country and do a first screening.
<b>CONTENT</b>	<ul style="list-style-type: none"> <li>• collection of 30 - 50 evaluation reports</li> <li>• first screening of the evaluation reports</li> <li>• present evaluation reports on <a href="http://www.sustainability.at">www.sustainability.at</a></li> </ul>

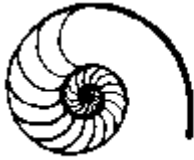
<b>YOUR BENEFIT</b>	<ul style="list-style-type: none"> <li>• you can offer evaluators to present their evaluation reports on the website of a European Network for Evaluation of Sustainability (<a href="http://www.sustainability.at">www.sustainability.at</a>)</li> <li>• you will receive feedback from from the EASY-ECO Team on your draft before you start the analysis</li> </ul>
<b>TOOLS</b>	<ul style="list-style-type: none"> <li>• DATA SHEET for evaluation reports as xls</li> <li>• <b>GuidelineB</b>: Definition of the Scope of the country surveys ("what is in and what is out?") as pdf</li> </ul>
<b>DELIVERABLES</b>	<ul style="list-style-type: none"> <li>• a DATA SHEET of evaluation reports per person</li> <li>• the conference management team gives you feedback on the country profiles and the topics you chose within 3 weeks after delivery</li> </ul>
<b>TIME NEEDED</b>	5 - 10 days estimate
<b>MILESTONE 2</b>	<b>September 30, 2002:</b> Please send your DATA SHEET to <a href="mailto:easy@wu-wien.ac.at">easy@wu-wien.ac.at</a> by then.

<b>PHASE 3:</b>	<b>Analysis Phase</b>
<b>GOAL</b>	A DRAFT PAPER of your country survey (one per person).
<b>CONTENT</b>	<ul style="list-style-type: none"> <li>• reading of the evaluation reports</li> <li>• analyse the evaluation reports according to the RESEARCH QUESTIONS</li> <li>• develop a draft paper</li> </ul>
<b>YOUR BENEFIT</b>	<ul style="list-style-type: none"> <li>• you get feedback on your draft paper before you go into detail</li> <li>• you can exchange your draft paper with other Young Researchers and get inputs</li> </ul>
<b>TOOLS</b>	<ul style="list-style-type: none"> <li>• RESEARCH QUESTIONS are available on request from <a href="mailto:easy-@wu-wien.ac.at">easy-@wu-wien.ac.at</a> = <b>GuidelineC</b></li> </ul>
<b>DELIVERABLES</b>	<ul style="list-style-type: none"> <li>• one draft country survey per person (3 - 5 pages)</li> <li>• you will get feedback on your selection of evaluation reports (only) if we think that changes are necessary within 2 weeks after delivery</li> </ul>
<b>TIME NEEDED</b>	10 - 20 days estimate
<b>MILESTONE 3</b>	<b>December 1, 2002:</b> Please send your DRAFT COUNTRY SURVEY to <a href="mailto:easy@wu-wien.ac.at">easy@wu-wien.ac.at</a> by then.

<b>PHASE 4:</b>	<b>Writing of the COUNTRY SURVEY PAPER</b>
<b>GOAL</b>	Submission of your country surveys to the EASY-ECO team and feedback by the scientific committee.
<b>CONTENT</b>	<ul style="list-style-type: none"> <li>• country surveys</li> </ul>
<b>YOUR BENEFIT</b>	<ul style="list-style-type: none"> <li>• you get feedback on your paper by the scientific committee</li> </ul>
<b>TOOLS</b>	<ul style="list-style-type: none"> <li>• use the forum for your questions to other Young Researchers and Keynotes</li> </ul>
<b>DELIVERABLES</b>	<ul style="list-style-type: none"> <li>• you will receive feedback on your draft paper within 3 weeks after delivery</li> <li>• your paper of 2000 - 6000 words = COUNTRY SURVEY</li> </ul>
<b>TIME NEEDED</b>	5 - 15 days estimate
<b>MILESTONE 4</b>	<b>February 1, 2003:</b> Please send your COUNTRY SURVEY to <a href="mailto:easy@wu-wien.ac.at">easy@wu-wien.ac.at</a> by then.

<b>PHASE 5:</b>	<b>Presentation at EASY-ECO 2</b>
<b>GOAL</b>	Presentation of your COUNTRY SURVEY and the European Synthesis at EASY-ECO 2.
<b>CONTENT</b>	<ul style="list-style-type: none"> <li>• final COUNTRY SURVEY paper</li> <li>• feedback by scientific committee</li> <li>• presentation at EASY-ECO 2</li> <li>• EASY-ECO Proceedings</li> </ul>
<b>YOUR BENEFIT</b>	

	<ul style="list-style-type: none"> <li>• you get feedback on your paper before the presentation by the scientific committee, within 6 weeks after delivery</li> <li>• you can present your paper in front of a large audience of scientists, commissioning agents, evaluators</li> <li>• you receive EU funding for travel, accommodation and conference fee</li> </ul>
<b>TOOLS</b>	<ul style="list-style-type: none"> <li>• use the forum for your questions to other Young Researchers and Keynotes</li> </ul>
<b>DELIVERABLES</b>	<ul style="list-style-type: none"> <li>• final COUNTRY SURVEY paper of 2000 - 6000 words</li> <li>• 15 minute presentation at EASY-ECO 2</li> <li>• your paper will be printed in the EASY-ECO Proceedings</li> </ul>
<b>TIME NEEDED</b>	1 - 3 days estimate
<b>MILESTONE 5</b>	<b>March/April 2003:</b> presentation of your paper at EASY-ECO 2



# **EASY-ECO** **EvAluation of SustainabilitY** **European COncferences**

[OVERVIEW](#)

[EASY-ECO-1](#)

[COUNTRY SURVEYS](#)

[EASY-ECO-2](#)

[NETWORK](#)

[forum](#)

**[training material](#)**

[guideline](#)

[information material](#)

You can download the training material from here.

[Glossary of Terms \(pdf\)](#)

[Literature \(pdf\)](#)

[Links \(pdf\)](#)





# EASY-ECO

## EvAluation of SustainabilitY

### European COnferences

---

OVERVIEW

EASY-ECO-1

COUNTRY SURVEYS

EASY-ECO-2

**NETWORK**

In the year 2000 The University of Economics and Business Administration in Vienna initiated an **international research network** designed to explore the field of evaluating Sustainable Development.

For this purpose, projects are planned in the areas of

- basic research
- application of such research
- exchange of knowledge between science, evaluation practice and the authorities commissioning projects on the evaluation of sustainable development.

#### Our Research Programme

This paper outlines the main research questions which will be tackled in this research network over the next few years. We will define research topics to be investigated by the network on the basis of the following four problem areas:

- Problem area 1:** The market for evaluating Sustainable Development lacks transparency, while demand is rising.
- Problem area 2:** There is no established set of methods (toolkit) for the evaluation of Sustainable Development, but a multitude of disciplines and approaches.
- Problem area 3:** There is no common understanding of the operationalisation of Sustainable Development to be used in evaluation projects
- Problem area 4:** The evaluation of Sustainable Development is part of a larger management system and presents a chance for organized learning.

A multitude of research questions arise from these problem areas, which can be summarized by the following aims:

- Development of an interdisciplinary framework to systemise present know-how available from the various disciplines dealing with the evaluation of Sustainability.
- Basic research to further develop methods for the evaluation of Sustainable Development with a view to creating a modular toolkit.
- Theoretical discussion of the paradigms of Sustainable Development and their consequences for the evaluation of Sustainable Development.
- Critical reflection on current evaluation practice through meta analysis, comparative methods analysis, country analysis and dialogues between scientists and practitioners.
- Creating a network integrating scientific research, evaluation practice and those commissioning evaluations to establish a common understanding of the evaluation of Sustainable Development
- Further development of the evaluation of Sustainable Development into an integral part of a sustainable management system which shows implementation success on different levels and contributes to a transfer of learning effects.



# EASY-ECO

## EvAluation of SustainabilitY

### European COnferences

[OVERVIEW](#)[EASY-ECO-1](#)[COUNTRY SURVEYS](#)[EASY-ECO-2](#)[NETWORK](#)

EASY-ECO 2 will take place in Vienna in spring 2003.

**Main topics:**

- Is an interdisciplinary approach to the "Evaluation of Sustainability" necessary and possible?
- Challenges for the disciplines involved in the "Evaluation of Sustainability"
- Evaluations of project-, programme- and policy level
  
- Presentation of the European country surveys
- Conclusions to be drawn from the country surveys for each country and for the European Commission

The EASY-ECO 2 European Conference will focus on a broad discussion of the „Evaluation of Sustainable Development“ amongst scientists, practitioners, people commissioning evaluations and the interested public. The number of participants will be approximately 120.

Details about the date, programme and call for papers will be issued on this website.

If you want to keep informed about the EASY-ECO 2, please send a short email to [easy@wu-wien.ac.at](mailto:easy@wu-wien.ac.at)