UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN-YVELINES

U.F.R. des Sciences Sociales et des Humanités

Environmental Policy Evaluation in the Service of Sustainable Development: Influence of the OECD Environmental Performance Reviews from the Perspective of Institutional Economics

Thèse pour le doctorat en Sciences Economiques

Présentée et soutenue publiquement par

Markku LEHTONEN

Le mardi 29 Novembre 2005

Directeur de thèse : **Martin O'Connor**, Professeur, C3ED, Université de Versailles Saint-Quentin en-Yvelines.

Jury : **Beat Bürgenmeier**, Professeur, Département d'économie politique, Université de Genève, rapporteur.

Robert Delorme, Professeur, C3ED, Université de Versailles Saint-Quentin-en-Yvelines.

Janne Hukkinen, Professeur, Laboratory of Environmental Protection, Helsinki University of Technology, rapporteur.

Jacques Theys, Directeur, Centre de Prospective et de Veille Scientifique, Ministère de l'Équipement, des Logements et des Transports.

L'Université de Versailles Saint-Quentin-en-Yvelines n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

The University of Versailles Saint-Quentin-en-Yvelines does not approve or reject any of the opinions expressed in this thesis; these opinions should be considered as full responsibility of the authors.

Abstract

This thesis examines what consequences, through which pathways, and under which circumstances, does country-level environmental policy evaluation, carried out by an intergovernmental organisation, have for institutional change towards sustainable development. The philosophical and theoretical frameworks of institutional economics are proposed as the lens through which to assess experiences from evaluation research. This is done in order to elaborate a conceptual framework for analysing the approach, the impact and the factors conditioning the influence of environmental policy evaluation in the context of sustainable development. The OECD Environmental Performance Review programme is examined through an in-depth qualitative case study approach. It is found that in spite of engendering relatively modest interest among the public, the reviews increase the legitimacy of environmental policies and authorities, and strengthen a performance-oriented environmental policy emphasising evaluation, monitoring and cost-effectiveness. These impacts operate through four parallel pathways, representing the purposes of learning and accountability. The type and intensity of influence depends on the capacity of the reviews to mobilise key change agents and to be sensitive to the context, as well as on the credibility and political weight of the OECD in the reviewed country. It is concluded that notwithstanding the distance of the prevailing OECD conception of sustainability from the institutionalist principles adopted as the benchmark for the appraisal in this thesis, the reviews contribute meaningfully to sustainable development by redressing asymmetries of power, thereby improving some of the conditions for deliberative democracy.

Keywords: environmental performance, OECD, evaluation, institutional economics

Résumé

Cette thèse vise à étudier les conséquences en terme de changement institutionnel pour un développement durable de l'évaluation des politiques environnementales d'un pays, meneés par une organisation intergouvernementale. Les fondements philosophiques et théoriques de l'économie institutionnelle fournissent le cadre retenu pour construire une grille de lecture permettant d'analyser l'approche théorique, l'impact et les facteurs conditionnant l'impact de l'évaluation des politiques environnementales dans l'optique de développement durable. Les examens des performances environnementales de l'OCDE sont ensuite étudiés moyennant une approche d'étude de cas qualitative. Malgré l'attention relativement faible accordée par le public à ces examens, leur intérêt réside dans une plus grande légitimité des politiques et des autorités environnementales, et dans le renforcement des politiques environnementales basées sur les principes de contrôle, suivi, évaluation et coût-efficacité. Ces impacts opèrent à travers de quatre chemins parallèles, représentant les objectifs de reddition des comptes et de l'apprentissage. La forme et l'intensité de l'influence de ces impacts dépendent de la capacité des examens à mobiliser les acteurs clés, de leur sensibilité au contexte, ainsi que de la crédibilité et le poids politique de l'OCDE auprès des acteurs des pays étudiés. Nonobstant leur conception du développement durable relativement éloignée des principes institutionnalistes, les examens contribuent au développement durable en réduisant les déséquilibres de pouvoir et en améliorant ainsi certaines des conditions pour une démocratie délibérative.

Mots-clés : évaluation, politiques environnementales, développement durable, institutionnalisme, OCDE

Acknowledgements

It is practically impossible to give an exhaustive list of persons and institutions that have in one way or another helped me get through this work – my humblest apologies for all those who feel being excluded without good reason!

To start with, I am thankful to my thesis supervisor, professor Martin O'Connor, who has guided me through the project, while nevertheless allowing me a great deal of freedom to follow my own path. The always as inspiring comments from professors Janne Hukkinen and Joachim Spangenberg were particularly helpful at the final stages of my work – even though they further postponed the beginning of my summer holidays. Professor Maurice Baslé gave few, but extremely valuable hints concerning literature on learning and participation in evaluation. The contribution of professor Jan Keppler was indispensable in helping me get started with the thesis and better focus my research.

My nine years of participation in the OECD working party responsible for the Environmental Performance Reviews have been full of not only highly stimulating discussions in which I believe I learned a great deal. I am highly indebted to all of the past and present members of the WPEP group for some of the most rewarding and enjoyable moments of my professional career. Without the support from the OECD Environment Directorate I could never have been able to carry out my research. My thanks go in the first place to the head of the Environmental Performance and Information Division, Christian Avérous, but I hereby wish to collectively thank all of my former colleagues and friends at the Division.

I highly appreciated the willingness of all of the interviewed persons to sacrifice their valuable time for answering my questions. I am particularly grateful to Klaas-Jan Moning and Pedro Liberato: without their help I could never have been able to carry out the interviews in the Netherlands and in Portugal.

The determined guidance and supervision of my former chief at the Ministry of the Environment, Heikki Sisula, this thesis would never had even started. Without Heikki I would probably never have become involved in OECD work. The innumerable discussions we have had over the years greatly helped me in better defining my research question and understanding the role of the OECD in environmental politics. By no means underestimating the importance of this contribution, immensely more precious has been the support and friendship from Heikki – he has been my 'tutor', my 'boss', my colleague, but above all a friend over the almost ten years that we have known each other.

The friendships that I have developed during my five years in Paris, especially at the three houses where I lived at the 'cite internationale universitaire' – Avicenne, Maison du Brésil and Maison des étudiants suédois – certainly did not always accelerate the finalisation of my thesis. Nevertheless, those friendships were what made my life in Paris so enjoyable. Our small, but highly qualified Finnish community in Paris, in particular Outi Honkatukia and Annukka Lipponen, has constituted a haven of relaxation, by providing me the possibility to escape the ever intriguing and charming, but sometimes so complicated French society and language. A special thanks goes to Outi, who with her habitual determination saved me from falling into despair faced with what I perceived as insurmountable computer problems. Among the numerous members of the C3ED who have helped me throughout the years at least one merits a special mention: the patience and kindness of Isabelle Coulon was indispensable in creating the material conditions necessary for the finalisation of the thesis.

My sons, Jaakko and Tuomo, have certainly had to suffer from the fact that 'papa' was most of the time not only far away, but especially during the past few months, even when physically present, too often mentally somewhere else. For the absentee 'papa' you have nevertheless provided an irreplaceable source of joy and energy.

Even a 'thésard' needs to eat and survive: without the grants from the Academy of Finland and 'Helsingin Sanomain 100-vuotissäätiö' I would still be only dreaming of one of these days writing an academic dissertation.

Finally, if my work has any intellectual value, this is to a great extent because of the inspiration that I got during my first years as an undergraduate student from the late Kauko Hahtola – former professor of Environmental Economics and Land Use Planning at the University of Helsinki. Had Kauko only known how much the totally unexpected e-mail from him and our subsequent brief period of e-mail correspondence just a couple of months before he passed away meant to me…

Paris, 7 September 2005

Markku Lehtonen

TABLE OF CONTENTS

List of Tables

List of Figures

List of Boxes

PART I

INSTITUTIONALIST ENVIRONMENTAL PERFORMANCE EVALUATION: THEORY AND INFLUENCE

The first part of the thesis consists of four chapters. The introductory chapter presents the background, defines the research question, explains the research approach, and provides a brief chapter outline. Chapter two goes on by outlining more in detail the philosophical and theoretical premises underlying the study, drawing mainly on institutional economics. Chapter three builds upon this basis in order to develop a set of criteria for an 'institutionalist' conception of sustainable development, seen as the ultimate objective of environmental performance evaluation. The first section of chapter four completes the criteria for an 'institutionalist' evaluation of environmental performance pursuant to sustainable development, by drawing upon lessons from evaluation research. The second and third sections of the chapter focus on the influence of evaluation, first developing a typology of evaluation influence, and then identifying the range of factors that potentially condition the influence of evaluations.

Chapter 1

Introduction

Can an international organisation fundamentally committed to the promotion of economic growth and free trade contribute to sustainable development by evaluating environmental policies of its member countries? If the organisation succeeds in transmitting its policy messages, would the impact of the evaluations be in harmony with what is generally perceived as sustainable development? An intuitive answer to the question might be negative: the objectives of economic growth, environmental protection and social cohesion are frequently still seen as antagonistic and mutually irreconcilable. And yet, the very concept of sustainable development had its origin in the belief that progress, modernity, and economic growth are essential preconditions for environmentally sound development. A same type of ambiguity concerns evaluation of public policies. On the one hand, evaluation has been seen as simply a continuation of the 'modern project', leading to the perpetuation of the prevailing 'bads' of modernity, with its illusory trust in the capacity of the humankind to ultimately 'get things under control'. On the other hand, it has been proposed as an instrument for strengthening democracy, the 'reflexivity of modernity' and 'ecological modernisation'.

Regardless of the viewpoint on these matters, one could at least expect evaluations to have an impact – positive or negative – on public policies. However, it is precisely this assumption that has been called into question by scholars studying the use of evaluation and research in policymaking: it is not uncommon to argue that at best, evaluations are exercises aimed at little but legitimising policies and decisions that have been taken already, and that frequently they are simply useless, producing no impacts whatsoever beyond the waste of (often public) money. If we, after all, adopt a more optimistic standpoint, and believe that evaluations could play a role in shaping policies, what

are the possible mechanisms of influence? Who uses evaluations? Can an evaluation have an impact even if nobody actively 'uses' it? And, further, why might an evaluation be influential in one country, under specific circumstances, while lacking influence in another setting? Is it possible to draw any generalisable conclusions concerning the influence of evaluations or does it all simply depend on the context? These are the kind of questions this thesis seeks to answer.

1.1 Evaluation as a tool for promoting institutional change towards sustainable development?

Evaluation of public policies has over the past decades become a popular catchword, increasingly presented as the solution for the problems of inefficiency of public policies, and as a way of bringing much-needed rationality to decision-making. As it has become increasingly difficult for the state to exercise its functions as a mere regulator in the new governance structures characterised by a greater role of the market and non-state actors, evaluation has been suggested as a way for the state to gain back some of its lost power. While evaluation research and practice have developed mostly in the fields of social and education policy, they are being applied more and more frequently to solving the problems of environment and sustainability, the traditional *ex ante* evaluations therefore increasingly being complemented by *ex post* evaluations.

The origin of the rising popularity of evaluation can be traced back to two mutually contrasting perspectives. First, evaluation can be seen as a response to a number of dominant trends in public policy, such as increasing decentralisation, citizens' demands for greater transparency of policymaking, and the pressures to cut public spending, which have strengthened the pressures upon the public sector to demonstrate *accountability* to external sponsors and other stakeholders. The clearest manifestation of such evolutions have been the rise of performance measurement, in the framework of the so-called New Public Management, which has expanded to all fields of public policy. (E.g. Greene 1999, 162; Uusikylä 1999; Thoenig 2000; van der Knaap 2000; Pitarelli and Monnier 2000; Feller 2002, 438.)

Secondly, along with this efficiency-oriented, rationalistic trend, evaluation has been advocated as a response to increasing complexity, interconnectedness, pervasive uncertainty, potential irreversibilities, as well as plurality of values and points of view inherent in today's policy problems, particularly in those relating to sustainable development. This has stimulated search for

new policy instruments and approaches, better suited to the problems at hand, and more likely to foster institutional change towards sustainable development. Therefore, environmental policies have over the past decade or so been characterised by a 'paradigm shift' (e.g. Driessen and Glasbergen 2002) or 'regulatory reform' (Sairinen 2000), entailing a progressive shift of emphasis from the top-down approach based exclusively on 'command-and-control' instruments – highly successful in reducing some of the 'simple' point-source pollution problems in most developed countries – towards a new approach, involving the application of a wide range of instruments – regulatory, economic, and information-based ones – as 'policy-packages'. Stakeholders have become closely involved in the design of policy instruments, and the authorities increasingly adopt a 'facilitating' role instead of acting as mere regulators. (E.g. Schubert et al. 2000; Sairinen 2000; Weale 1992; Hajer 1995; Mol and Spaargaren 2000; Joas 2001). From this point of view, the wider use of evaluations and indicators can be seen as a characteristic of the 'reflexivity' of modernisation (Giddens 1990) and a way of strengthening 'deliberative democracy' (e.g. Dryzek 2000; O'Connor et al. 2001/2002) through inclusive, participatory policymaking, which should ultimately contribute to sustainability through *learning* (e.g. van der Knaap 1995; Baron 1999; Siebenhühner 2001). This would be achieved through in-depth evaluations aimed primarily at explaining the reasons behind the policy outcomes.

For both of these perspectives, the *raison d'être* of evaluations is their utilisation in public policymaking. Indeed, without the promise of utilisation, there would hardly be any reason for policymakers to commission and sponsors to finance evaluations. However, most empirical studies have shown that direct, instrumental use of evaluation results in decision-making is rather an exception than a rule, often limited to the technical, operational level, while various indirect uses, often seen in terms of 'enlightenment', are much more common (Leviton and Hughes 1981; Albæk 1995, 10-12; Lampinen 1992, 30-37; Muller and Surel 1998, 110; Pollit 1998; Weiss 1980; 1987; 1998; 1999; Hanberger 2001a, 58). This has led to disappointed views concerning the potential of evaluations to provoke policy change. Yet, it has also contributed to a more realistic picture of policymaking and evaluation use, as well as brought attention to new elements and pathways through which evaluations influence policies.

There is an obvious tension between the two perspectives to evaluation use and they are often seen as irreconcilable. At least in its simplest version the 'accountability perspective' can be seen as a reincarnation of the traditional, rationalist-technocratic model of policymaking, which expects evaluation results to be used by policymakers in a direct, instrumental manner (e.g. Davies 1999;

Greene 1999). If this is the case, one can legitimately ask whether evaluation is in line with key principles of sustainable development. However, providing accountability and enhancing learning can be seen as essential elements in the endeavour to promote 'social betterment' – institutional change towards sustainable development – through evaluations. While it is highly desirable to monitor and continuously seek to improve policy processes, one also needs to know whether policies are in fact delivering the expected results. The challenge is therefore not to choose between the two, but to look for complementarity by clearly defining the roles of the two approaches (e.g. Biott and Cook 2000; Greene 1999; Davies 1999; Blalock 1999; Wimbush and Watson 2000; Perrin 1998; 1999; 2002a; 2002b; Bernstein 1999; Bukkems and de Groot 2002). Combining the two objectives within a single evaluation framework, however, has been an idea largely absent from debates on evaluation. This thesis aims to help fill this gap by studying the possibilities and limitations involved in combining the functions of accountability and learning in international environmental policy evaluation having the ambition to promote sustainable development.

Evaluation research has traditionally concerned evaluations carried out by consultants at sub-national levels, focusing on specific programmes or policies, whereas the evaluation of the entire range of a country's policies in a specific sector has received less attention. Furthermore, while there is an abundant literature on the role and the responsibilities of the evaluator,[1] the specific challenges faced by institutional evaluators – international organisations in particular – have been subject to less scrutiny. An international organisation is not like any other evaluator, but brings to the picture a much broader range of concerns, related for instance to the overall role and prestige of the organisation in international governance, to its internal dynamics, and to its overall 'policy doctrine'. Given the rising interest in evaluation as a new policy instrument, the increasing international interdependence, and the global character of the most pressing environmental and sustainability problems, international organisations are likely to gain an increasingly prominent role in conducting country studies and evaluations.

In the final analysis, evaluation influence is about power: in which manner do evaluations change power relations between different actors in society, 'empowering' some actors and 'disempowering' others, supporting or legitimising certain ideas and ways of thought, while discrediting others? The role of power is probably easier to grasp in a framework emphasising accountability – evaluations exert pressure on evaluated entities by increasing transparency and thus exposing them to outside criticism. Yet, power is also present in collective learning: what gets learned, by whom and under

which conditions, crucially depends on the relations of power among the actors involved. Starting from the assumption that 'words have power', one is lead to examine the ways and degrees to which an international organisation exercises its power through 'idea games' (Marcussen 2001). In such 'games', the prestige and reputation of the organisation combine with its power to define the terms of the discussion and create analytical frameworks, thereby affecting the 'mental models' (e.g. Denzau and North 1994; Hukkinen 1999) of key players in the political arena.

An evaluation's impact on sustainable development therefore does not consist merely of its impact on policy effectiveness and efficiency on the one hand, and on learning concerning the concrete policies being evaluated on the other. Any evaluation of environmental policies carries with it specific – explicit or implicit – philosophical and theoretical premises, including a particular understanding of the character of sustainable development. It is plausible to assume that the more prestigious the organisation that puts forward a particular world-view or approach to sustainable development, the greater is its influence on the conceptualisations and frameworks of thought of the actors involved. If the evaluator is a highly esteemed international organisation, having economic development as its most fundamental objective, and its views on economy and economics have influenced policies in the industrialised countries since the beginning of the 1960s, it becomes primordial to analyse the conception of sustainable development underlying the evaluation approach. An essential question therefore is to what extent the evaluation serves to consolidate the dominant perspective of modern mainstream economics (Boven 2003), and the dominant premises of modernism, with all of their shortcomings and 'illusions' (Norgaard 1994, 62-63), as opposed to an approach taking institutions seriously, therefore being better adapted to the complex problems of sustainable development, as argued further on in this thesis.

1.2 Research task

The main research question addressed in this thesis is *what consequences, through which pathways, and under which circumstances, does country-level environmental performance evaluation, carried out by an intergovernmental organisation, have for institutional change towards sustainable development*.

---

[1] This issue ill be addressed more in detail in section 4.4.1.

This institutional change is seen in a broad 'coevolutionary' framework, consisting of subsystems – values, knowledge, environment, technology, and organisation – that evolve together in continuous mutual interaction (Norgaard 1994). The OECD Environmental Performance Reviews (EPRs) have been chosen as the case study, as probably the oldest attempt to systematically evaluate environmental policies of the industrialised countries in a comparative perspective.

The main research question can be further divided into three sub-questions:

1. In which ways have the EPRs influenced policies in the reviewed countries?
2. Which factors condition this influence?
3. Does the influence of the EPRs in fact promote sustainable development, as defined in the coevolutionary perspective of institutional economics?

The question concerning the types or pathways of influence starts from the assumption that before asking whether an evaluation enhances sustainable development, one has to first establish whether it has any influence in the first place, and if so, in which manner. This question relates to the policy impacts of the EPRs: how have the conclusions and recommendations been received, how has the process of preparing the review influenced actors in the reviewed country, and which indirect impacts do the reviews have?

The second question – the factors conditioning the influence of reviews – stems from the observation that the EPRs are received differently in function of the reviewed country and the general institutional setting. How to explain the sometimes considerable differences in the reception of seemingly similar reviews in different countries? Which factors in the national policy style render some countries more susceptible to influence? How important, in the end, is the review design – the quality and the character of the reports, the pertinence of the evaluation criteria, the relevance of the recommendations, etc.? How does the role of the evaluator, in this case an intergovernmental organisation, affect the outcomes? Would a more participatory review design better enhance institutional change towards sustainable development or would it compromise the credibility of the organisation and thereby also the capacity of the reviews to exert their influence?

These considerations lead to the third question, namely that of the kind of change that the EPRs promote. Intuitively, one might expect the premises underpinning OECD work to be rather alien to the 'institutionalist' principles of sustainable development, given the organisation's mandate of

enhancing economic growth and free trade, and its reputation as an advocate of neo-liberal pro-globalisation policies.[2] If this were the case, and if the basic approach of the reviews runs against the principles of sustainable development, it might be better not to carry out the reviews at all. The last question therefore deals with two aspects. First, it relates to the extent to which the message transmitted through the OECD EPRs is in line with the fundamental premises of institutional economics concerning issues such as the view of progress, modernity, democracy, the definition of sustainable development (including the interaction between its different dimensions, levels and scales), as well as the methodological and policy choices. Second, it requires a broader perspective on the entire EPR process: does the process as a whole contain elements that bring it in line with the principles of sustainable development this thesis is based upon? On balance, taking into account the impacts of the EPRs throughout the whole review process, what can we say about the impacts on sustainable development?

The first part of the thesis sets up the theoretical framework for analysing the EPRs in the light of the three research questions listed above. The second part, on the one hand, proceeds to applying the theoretical framework to the analysis of the EPR programme, and on the other, elaborates general hypotheses on evaluation influence to allow their further testing in possible future research in the context of international environmental policy evaluations.

1.3 Ontological and epistemological basis of the thesis

The ontological and epistemological bases of this thesis tend towards pragmatic realism, defined by Bazzoli (1999, 192) as a point of view that abandons the ontological realism associated with positivism, but at the same time rejects the constructivist subjectivism of certain post-modernist authors. Contrary to the positivist view, it recognises mental representation as an active process of social construction that is conditioned by the historically contingent context, but unlike the 'extreme version' of constructivism, it does not assume that reality would be a mere mental construction. Reality 'out there', independent of our perceptions, is more than likely to exist, but any attempt to grasp that reality is mediated through culturally and historically contingent human interpretation. In short, perception, whilst unavoidably theory-bound, has a real object (Hodgson 1988, 36; Yearley 2002). Individuals always think and act within a historically defined context, which is at that very

---

[2] See e.g. "L'OCDE est un piston de la mondialisation." Libération 7 août 2001.

moment 'objective', outside of their control, yet shaped by and shaping individuals (Sarbin and Kitsuse 1994, 12).

This thesis starts from the assumption that scientific objectivity in the sense of value-neutrality is impossible, because data are always perceived through theoretical interpretation, which is unavoidably influenced by values.[3] This does not amount to saying that evidence has no importance or that empirical work would have no significance or value. Instead, it means that facts and values are epistemologically different, that facts do not 'speak for themselves'; 'ought' does not automatically spring from 'is'.

Objectivity should be understood as maximum transparency with regard to the underlying assumptions and choices, as well as a conscious attempt to conceptually separate the factual statements from normative prescriptions. Universally valid rules of conduct for achieving scientific objectivity cannot be defined, but they must be outlined separately in each individual situation. Scientific inquiry must not be exclusively concerned with the attainment of the objective 'Truth', which is indeed unattainable, but also with the consequences of the activity for the betterment of the human condition.

In the words of Lukes (1982, 305): *"…in cross-cultural studies, and indeed elsewhere, 'the student must feel he or she is answerable to and in uncoercive contact with the culture or people being studied'. The data must be as systematically gathered as possible, relative to the explanatory purpose at hand. And the interpreter must be as reflexive as possible, maximally aware of his interpretive situation, without supposing that he can escape it."* And further: *"…perhaps the only objectivity that is possible in much of social enquiry: not 'perspective-neutrality', but rather accounts that are not merely theory- but also perspective-relative, yet constrained by evidence that is as systematic and as reliable as possible, and relatable to other perspective-relative accounts. In applying them, one may explain from some perspective what could not be explained from no perspective."*

This thesis relies on the perception that the role of social sciences resembles that of Aristotelian *'phronesis'*, contributing not so much to explanatory and predictive theory – as do the natural

---

[3] In Myrdal's (1978, 778-779) words: "Valuations are always with us. Disinterested research there has never been and can never be. […] Our valuations determine our approaches to a problem, the definition of our concepts, the choice of models, the selection of observations, the presentations of our conclusions – in fact the whole pursuit of a study from beginning to end."

sciences – but rather to "the reflective analysis and discussion of values and interests, which is the prerequisite for an enlightened political, economic, and cultural development in any society". Social science can be reduced neither to mere analytical, scientific knowledge (*episteme*), nor to technical knowledge or know-how (*techne*). (Flyvbjerg 2001, 2-3.)

1.4 Research method and the phases of work

In this thesis, I shall examine the OECD Environmental Performance Reviews so as to provide insights into the potential of environmental policy evaluation to contribute to sustainable development. Beyond the desire to explain the 'hows' and 'whys' relating to the EPR influence. I seek to enhance understanding of the influence of the reviews on sustainable development and develop hypotheses on the influence of environmental policy evaluations for further testing. The case study is analysed through a theoretical framework combining two strands of thought: institutional economics provides the philosophical and theoretical basis of the thesis, while insights from evaluation research are integrated within this general framework, and used in particular for analysing the influence of the EPRs. The research method is pluralistic, relying mainly on a qualitative case study approach, combining different types and sources of information, empirical material and theoretical work, in an attempt to provide a varied picture of the dynamics involved in the reviews and their consequences.

The major part of the material for the research was gathered through participant observation between May 1996 and May 2005, the period in which I participated as a Finnish delegate in the OECD Working Party for Environmental Performance (WPEP). This work also included participation in workshops preparing the 'second cycle' of EPRs, and a survey concerning the experiences from the 'first cycle' of EPRs that I carried out together with the OECD secretariat (OECD 1997b).[4] This survey helped identify the key points of importance for the review programme as perceived by the 'insiders' in the programme – points that came to characterise the discussions on the future of the reviews over the subsequent years in the WPEP group.

---

[4] The survey consisted of a questionnaire sent to the country delegates of the OECD Group on Environmental Performance (from 1998 on WPEP), who were asked to (i) answer the questionnaire themselves, and (ii) sent the questionnaire to some of their colleagues that have experience of the review programme. The questions aimed at discovering the degree of (dis)satisfaction of the participants with the EPRs, their opinions on the different features of the review programme, and suggestions for further improvement of the programme. Responses were received from 38 individuals.

Further sources of first-hand experience concerning the practical conduct of the reviews were my contribution as one of the two main organisers to the Finnish EPR in 1996-1997, participation as a country expert on teams reviewing the environmental performance of Mexico (1997-1998; writing a chapter on energy-environment integration) and Russia (1998-1999; industry-environment integration), and as an OECD consultant on the review of Sweden (2003-2004; general management of the review, 'first author' for the chapter on nature conservation, second author for 'implementation of policies').[5] The notes from the WPEP meetings and the EPR reviews constituted an essential part of the research material.

The second source of information consists of the various *OECD policy documents*, notably those relating to the EPR programme and the organisation's work in the area of sustainable development. I retained considerable freedom to decide which parts of the reports to concentrate on, in an attempt to provide a reliable picture of their message in the light of the research question. Such an approach obviously involves a fair amount of subjectivity, but the understanding of the key issues of the EPRs gained through the participation in the work of the WPEP facilitated the interpretation of the documents and the selection of the most relevant parts for analysis.

The third source of information, *interviews* with individuals involved in the EPR programme, represented a move forward from the two first phases of 'descriptive' and 'focused' observation typical for participant observation, towards 'selective' observation (Flick 2002, 140). The first group of interviewed individuals consisted of officials at the OECD secretariat, and the WPEP delegates of Canada, France, Hungary, Japan, Mexico, the Netherlands, Portugal, and Slovakia. The intention in the thesis was to analyse the EPR programme as a whole instead of individual reviews, and therefore most of the material relates to the totality of the programme. However, in order to gain better understanding of the cross-country variation of the EPRs' influence on policies, three countries – France, the Netherlands, and Portugal – were chosen for more in-depth analysis. Interviews were carried out among a number of mainly governmental stakeholders involved in the reviews of these countries.[6] To select the interviewees in the Netherlands and Portugal, I contacted the countries' WPEP delegates, explained the general research approach and asked the delegates to recommend potential interviewees from different stakeholder groups (different ministries involved

---

[5] For each chapter of the reviews, the OECD secretariat has the 'first author' that can be a country expert, a consultant, or a secretariat official, and whose responsibility is to prepare a first draft of the chapter and lead the discussion in the relevant meeting with the reviewed country's authorities during the review mission. The 'second author' is a permanent OECD secretariat official or a consultant, whose duty is to produce the final version of the chapter, on the basis of the work carried out by the first author.

in the reviews, NGOs, industry, and research community). In France, I directly took contact with persons having participated in the country's Environmental Performance Review meeting in June 2005. The ideal solution, in order to maximise the representativeness of the countries selected for more in-depth analysis, would have involved interviews in both European and non-European, federal and non-federal, old and new OECD countries, but resource limitations prevented non-European countries from being included in the sample. Moreover, the choice was limited by the desire to study countries, in which an EPR had been carried out not too long ago, to ensure interviewees would have the experiences still fresh in their memory. Of course, this choice had the disadvantage of preventing an analysis of the long-term impact of the reviews. Such an analysis would, in any case, have been beyond the reach of the chosen research approach, as it would have required analysis of policy developments over a longer period of time. I therefore concentrated on examining the perceptions concerning the relatively short-term impacts of the reviews.

Finally, interviews with persons involved in the Economic Surveys, notably delegates from Finland, the Netherlands and Portugal as well as from the OECD secretariat, allowed me to compare the influence of the Economic Surveys with those of the EPRs, gauge the importance of country-specific conditions on the influence of the reviews, and to obtain a broader perspective to the role of the OECD and its peer reviews. Beyond the issues relating to the organisation of the reviews, the topics addressed in these interviews concerned their influence in general and the sustainable development section in particular, the relevance and the tone of the policy recommendations, as well as general attitudes towards the OECD, the Economic Surveys, and the opinions coming from the outside of the country.

The literature on evaluation research and international relations provided insights into the influence of evaluation and the factors potentially affecting such influence. It also helped identify further issues of potential relevance, such as the purposes of evaluation, factors affecting the use of evaluations, and the different types of evaluation use documented in the literature. I started with a range as broad as possible of types of influence and conditioning factors, and focused on the most relevant ones only as the research process advanced. To allow the reader to judge the validity of my choices concerning the most relevant factors, the theoretical part of the thesis presents the broad panoply of types of influence (section 4.3) and factors conditioning the influence (section 4.4). By moving back and forth between the literature and my empirical background material (the survey on first cycle of reviews and my experience from the WPEP work), I could better specify the themes

---

[6] The numbrer of persons interviewed was eight in France, seven in the Netherlands, and nine in Portugal (see annex II).

and questions to be addressed in the interviews. The collection and the interpretation of data on the one hand, and the selection of empirical material, on the other, were therefore closely linked, which made it possible to continuously reflect upon the adequacy of the methods and theories in relation to the subject and the data under study (see Flick 2002, 43). The research approach can be illustrated through figure 1.1, describing the way in which, first, the selection of the case studies and the collection and the interpretation of material are influenced by the researcher's preliminary assumptions. These interpretations then further influence the sampling, collection, and interpretation of material in further cases. The process is completed through comparisons between the cases, leading to building of theories and hypotheses concerning the phenomena being studied. Such a strategy is close to the 'grounded theory' research, based on an attempt to approach the research subject with an open attitude, allowing to 'ground' the theories in the context, instead of trying to fit the phenomenon being studied into a predefined theoretical framework.[7]

Figure 1.1. Qualitative case study research approach
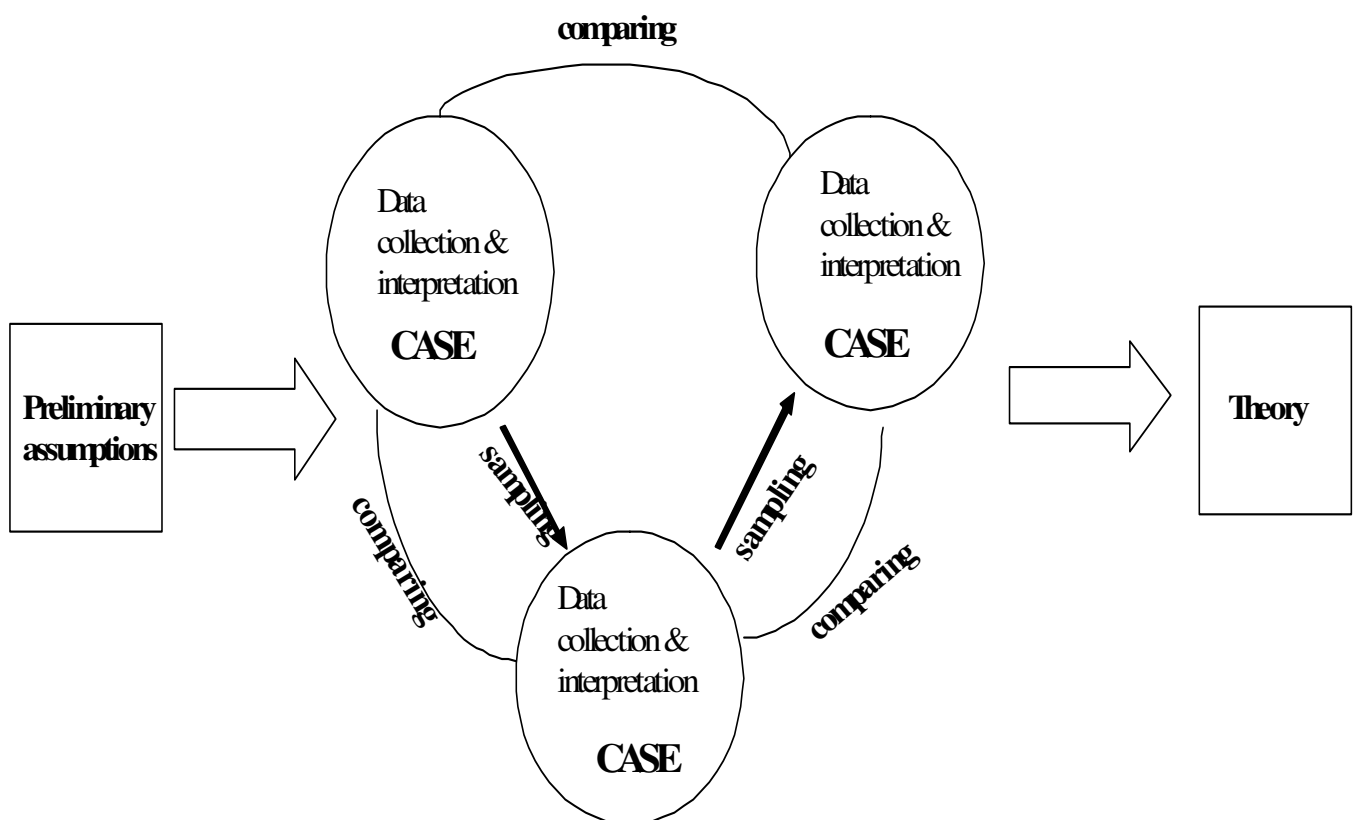
(Adapted from Flick 2000, 44).

---

[7] Flick (2000, 41) describes the openness inherent in the 'grounded theory' approach by explaining that "the theoretical structuring of the issue under study is postponed until the structuration of the issue under study by the persons being

In the semi-structured interviews, an interview guide (annex II) was used to keep the discussion 'on track', to ensure a focus on the specific issues expected to be the most relevant, and to make sure that the entire range of potentially relevant themes were dealt with. I nevertheless let the interviewee concentrate on those themes that she felt were the most pertinent ones. The interviews proceeded from unstructured general questions towards more structured ones. The list of questions was not exhaustive; new, relevant issues were taken on board as they were brought up by the interviewees, and were further integrated into the subsequent interviews. Likewise, some of the questions that I thought relevant in the beginning turned out to be of minor relevance and were dropped in the later interviews. In the three 'mini-case study' countries, the first interview was carried out with the country's WPEP delegate, in order to obtain a general picture of the flow of events and an understanding of the contextual issues, which allowed me limit the range of themes to be discussed with the other interviewees. The questions were adapted according to the allegiance of the interviewed individuals: not all of the questions posed to the authorities were of relevance when discussing with researchers or NGOs, for instance. When contacting the selected interviewees for the first time, by telephone or through e-mail, I explained the general research problem and the contributions expected from the interviewee. In many cases, I also supplied a list of the major themes of the interview to the interviewees in advance. This was done not only in order to help the interviewees prepare for the interview, but also to minimise the risk of the interviewees being carried away with telling stories totally unrelated with the research question at hand.

The material was analysed both from the '*fact perspective*' and '*sample perspective*' characteristic to qualitative research (e.g. Alasuutari 1999). In the former, the individuals are seen as 'informants', the objective being to find out 'what actually happened' in the reviewed country. This perspective was dominant in seeking to answer the question of what actually was the influence of the reviews in the reviewed countries (chapter 6). Of course, the descriptions of different individuals of the same event varied depending on the persons' position in the process, but rarely directly contradicted each other. Therefore, I was able to reconstruct the chain of events, identify the main topics concerning the reviews in the three case study countries, and the different types of influence from the EPRs. The interviews also gave an idea of the significance of the reviews in the different contexts and for different groups and individuals. The analysis of the interviews at this stage was a two-way process of, on the one hand, checking the pertinence of the categories identified – on the basis of the background experience and through literature research – as potential pathways and types of influence, and on the other hand, completing the typology of review influence by insights from the

studied has emerged."

interviews. The typology therefore functioned as a tool for extracting from the interview and other material statements that testified to the operation of certain types of influence. By combining insights from the literature, from the WPEP work, and the interviews, I could then construct the 'ideal types' of review influence presented in section 7.1.

The interviews were also examined from the 'sample perspective', based on a social constructionist approach, which considers that it is ultimately impossible to get information of 'reality as such', because all accounts of that reality are inevitably social constructions. This implies considering the interviews not as reflections or descriptions, but rather as 'samples' of the socially constructed reality. From this point of view, the reliability and sincerity of the interviewees' descriptions of the course of events is not interesting, but instead, their accounts allow the researcher to understand the underlying interpretative framework and context within which the interviewees operate. Such a 'sample' may represent well or poorly the studied phenomenon, but it cannot be 'true' or 'false'. This perspective was particularly present in the identification of the factors that condition review influence. The first task was to find out the interviewees' opinions concerning the review design – an issue on which I already had a fairly good understanding on the basis of previous participation in WPEP work and studying the background material. One aspect analysed in the interview material at this stage were the views of the interviewees on the largely dichotomous choices concerning the review design that had been identified throughout the work within the WPEP (section 7.1, table 7.1). Another task was to identify the gaps that existed between the interviewees' perception of the way in which the EPRs influence policies at present, and the way they would like the reviews to operate. In constructing the different 'repertoires' operational in the reviews (section 7.2), I tried to understand which were the key elements structuring the thoughts of the inteviewees concerning the EPRs. The aim was not to identify a single dominant repertoire for each interviewee, but instead to chart the whole range of the repertoires, recognising that an individual often operates within several, even mutually contradicting repertoires.

The identification of the country-specific factors that shape review influence combined both the 'fact' and 'sample' perspectives. The former consisted of posing questions on what actually happened, and which were the contextual conditions in the country, whereas the 'sample' perspective was more prominent with respect to the questions relating to the political style of the country, and attitudes towards the OECD, for instance. On the one hand, the intention was to find out 'how the country operates' – what are the prevailing socio-cultural, economic, and political circumstances – and on the other, to obtain information on the attitudes of the interviewees

themselves, in order to assess the meaning and significance of such attitudes for the review programme.

The interviews were mainly analysed through a hermeneutical, understanding approach, involving constant movement between an attempt to grasp the overall 'message' of the interviewees, taking into account the context within which they operated, and an analysis of the specific passages in the interview material.

1.5 Case study research, validity, and generalisation

In case study research, the questions of validity, control, and reliability take a different meaning than in quantitative research, for instance. Here, the frame of reference and the precise choice of the research object – i.e. the research strategy – instead of the number of cases, becomes decisive. For instance, the people to be studied are selected in function with their relevance to the research topic rather than their representativeness (Flick 2000, 41). In order to ensure the validity of my inferences concerning the accounts of the interviewees, I tried, as much as possible, to 'triangulate', by relying on multiple sources of information (in particular, my notes from the WPEP meetings, the survey concerning the experiences from the first cycle, and OECD memoranda on review missions and review release events). The interviews with the Finnish, French and the Portuguese informants were carried out in the interviewees' mother tongue, so as to avoid difficulties interviewees might face in expressing themselves in a foreign language. The rest of the interviews were conducted in English.

As one measure for ensuring the validity of inferences, a memorandum was written (in the language used in the discussion) after each interview on the basis of the discussion and, with few exceptions, sent to the interviewee for correction and approval. Direct quotations from these memoranda are used, mainly in chapters 6 and 7, to illustrate and substantiate arguments put forward in the text. For the Finnish, French and Portuguese interviews, I translated the quotations into English. To protect the anonymity of the interviewees, reference is made only to the group represented by the interviewee. The reference group mentioned in the text may differ for the same interviewee depending on the context in which the quotation appears; e.g. 'a WPEP delegate' and 'a civil servant from the Ministry of the Environment' may be the same person. The reference group was chosen case-by-case with a view to protecting anonymity and in function of its relevance to the main argument.

In order to verify the reliability of the 'ideal types' and the inferences concerning the relationships between the different factors, I on the one hand sought to constantly re-examine the key factors and types of influence that I had identified, and on the other, went back to my material, and sought for any deviations from the general patterns. For instance, where an actor situated in the 'political' repertoire favoured a review design with 'soft' and diplomatic recommendations and an administration-oriented review setting, I sought for other factors – e.g. those relating to country-specific context – that could explain such a deviation.

Case studies are sometimes regarded as interesting only in so far as the case in question can be demonstrated to be representative of a typical case or the event affects many lives. However, an individual phenomenon can also be significant if its analysis provides new insights into the life of society. The commonly held view that case study can be valuable at most at the preliminary stage of an investigation is based on a number of misunderstandings or oversimplifications about the nature of the case study as a research method. For instance, generalisation is indeed possible on the basis of a single case, but formal generalisation is not the only source of scientific development. Instead, the 'power of the good example' can play an important role as well. Furthermore, it is true that summarising case studies is often difficult, but these problems are more often due to the properties of the reality studied than to the case study as a method, and it may sometimes not be desirable at all to summarise and generalise case studies. (Flyvbjerg 2001, 66-87.)

This thesis has the characteristics of a case study approach at different levels. On one level, the EPR programme can be seen as a case representing environmental policy evaluation in general. On the other hand, the reviews of different countries can be seen as cases representing the EPR programme. Therefore, also the question of representativeness of the cases appears at these two levels.

It is clear that the EPR programme cannot be considered as a representative case of all possible environmental policy evaluations. However, large part of the interest of the case stems precisely from the unique character of the reviews: they are one of the few, if not the only, systematic attempts at comparative evaluations of industrialised countries' environmental policies, and understanding the dynamics of the programme can help design similar reviews in different contexts. This is so not because the experience could be directly transferred to another context, but because an analysis of the specific case of the EPRs can help develop more general theories and hypotheses

of the interaction between the context, mechanism and outcomes of evaluations (Pawson and Tilley 1997).

The second level of generalisation – from individual reviews to the level of the EPR programme as a whole – entails a slightly different set of problems. In this thesis, an attempt was made, indeed, to use the different country cases as representative examples of how the reviews usually operate. However, such generalisations cannot be based on statistical inferences (e.g. about the number of occasions where a review has been used in a certain manner by a certain type of stakeholder), but rather on the theory-based conclusions concerning the ways in which reviews may operate, under specific conditions. A single, unique event of evaluation use, for example, may provide a theoretically interesting and relevant observation, if it helps understand the dynamics of the reviews. As opposed to quantitative research, in qualitative case study research the observations deviating from the general pattern are often the most fruitful sources of new insights and provide an opportunity for falsification of the researcher's preconceived notions (Flyvbjerg 2001, 81-84).

Finally, the EPRs can be seen as a 'case' representing OECD work and the role of an international organisation in international governance. Even though analysing OECD work in general is not a part of the main research question, and this perspective is therefore not pursued very far, the thesis does provide insights into the role of an international organisation in environmental governance, and thus lays ground for future research on the subject.

1.6 Thesis outline

The first part of the thesis lays out the theoretical framework for analysing the case study – the OECD Environmental Performance Review programme. The presentation starts, in chapter two, by outlining the philosophical and theoretical premises of this thesis, essentially based on institutional economics. In particular, it provides the basis for analysing institutional change in a coevolutionary framework.

The third chapter goes on asking how one can recognise institutional change towards sustainable development, taken here as the ultimate objective of environmental policy evaluation. The chapter looks first at the key issues relating to sustainable development; the relations between sustainable development, progress, and modernity; the limits set by biophysical and ecological systems to the

scale and character of human activities; and the dimensions of sustainable development and their relations with each other, with a special attention to the social dimension of sustainability. A key question is how to measure sustainable development if one accepts the view – advocated in this study – that sustainable development is a combination of substantive and procedural issues, questions of 'survival' and those of democracy. The chapter ends with a list of principles for an 'institutionalist' perspective on sustainable development.

The objective of institutional change towards sustainability having been defined in the third chapter, the fourth chapter addresses three topics. First, lessons will be sought for in evaluation research in order to define what an 'institutionalist' evaluation should look like – an evaluation that fosters change towards sustainable development. The second question relates to the pathways of influence from evaluations. Still drawing upon experience from evaluation research – notably the use and influence of evaluation – a typology of the types of review influence is developed, so as to provide a basis for the analysis of the types of impact from the OECD EPRs. Third, and to complete the theoretical part of the thesis, literatures on evaluation, political science and international relations help to identify the factors potentially conditioning the influence of evaluations on policy. Such factors include the role of an international organisation as evaluator; the involved actors with their expectations, interpretative frames and policy networks; the general policy context; and the evaluation design.

The second part of the thesis starts with a presentation of the case study – the OECD Environmental Performance Review (EPR) Programme. In order to situate the programme in its context, chapter five first looks at the role of the OECD as an international organisation, then gives an overview of the recent OECD efforts in the field of sustainable development, and briefly presents the Environment Directorate before describing the structure and logic of the EPR programme. Finally, to situate the programme within the range of the OECD peer reviews, the chapter briefly describes the OECD's 'flagship', the Economic Survey programme, notably the experiences from including a short section on sustainable development in a full cycle of economic reviews in 2001-2004.

Chapter six applies the conceptual framework developed in section 4.3 to analyse the ways in which the EPRs have influenced the reviewed countries. It looks at impacts at the individual and collective level; stemming from the review process as well as from the publication of the review report; and examines the impacts of the collective argumentative processes on policies, shared understandings,

legitimacy of actors and policies, networking, and agenda-setting. The chapter concludes by identifying the main types of influence from the EPRs.

Having identified the main types of influence, the study moves on to examining the factors behind the variations in influence of the EPRs between different countries and contexts. It first looks at the different aspects of review design, and links these findings with the insights from the previous chapter concerning the types of review influence, thereby identifying four main pathways through which the reviews can influence policies. The attention is then turned to the actors with their varying 'repertoires' and the networks within which they operate. In particular, the chapter presents the 'economist' and 'environmentalist' 'epistemic communities' as the two key networks of relevance for the EPRs, and looks at the role of the OECD as an evaluator operating within both of these communities. Further distinguishing a 'bureaucratic' and a 'political' repertoire allows us to see the pathways of influence in the context of the different repertoires. Finally, the chapter examines the importance of the country-specific factors within which the reviews are carried out, and tentative hypotheses are put forward concerning the interplay between the pathways of influence and the country-specific factors.

The direct influence of the EPRs – like that of many evaluations – is often rather disappointing at least to the more optimistic participants. Moreover, an evaluation may well be influential, yet the influence may go to a direction undesirable from the perspective of the premises laid out in this thesis. Chapter eight therefore examines to what extent the message implicit in the EPRs and the way in which the reviews are carried out are in line with the principles of 'institutionalist' evaluation. Four questions are addressed, relating to both the substantive and procedural – 'survival' and democracy-related aspects of sustainable development.

Chapter nine draws together the lessons of the study, summing up the pathways of influence, the factors conditioning the influence, and the contribution of the reviews to sustainable development. Despite their largely 'non-institutional' overall ethos, the EPRs do have capacity to foster sustainable development, mainly through 'empowering' weaker actors and thereby improving the conditions for deliberative democracy and by enhancing environmental policy integration. More ambivalent from the perspective of sustainable development is the reviews' role in advocating more control through rational, market-oriented, results- and performance-based environmental policies – in the spirit of modernisation. Ultimately, the choices concerning the design of the EPRs hinge on

the role of the OECD as an organisation, and the power relations between the different 'epistemic communities' operating within and influencing the organisation.

Chapter 2

Philosophical and theoretical premises: institutional economics

This chapter briefly presents the argument for adopting institutional economics as the basis when seeking to define an 'ideal model' of environmental policy evaluation pursuant to sustainable development. The assumptions underlying institutionalism will then function as our guide throughout the entire study. The key elements of such a perspective are the conception of institutions, the normative character of institutional economics, the pervasiveness of power and conflict in all social phenomena, deliberative democracy as the model for an 'ideal society', and methodological pluralism as the guide in scientific work.

2.1 Natural vs. social scientific approach: from dichotomies to coevolution

Complexity has always been a pervasive characteristic of the social world, and integrating it into the scientific approaches and models has for a long time already been a major challenge. However, it can be argued that complexity has been increasing, in particular in issues of sustainable development, which tend to consist of multiple interrelated problems, involve different disciplines, various hierarchical, temporal and spatial scales, as well as a large number of stakeholders (Rotmans and Asselt 2001, 43). Sources of increasing complexity include increasing geographical interdependence, acceleration of technological development and the processes of social change, as well as the constant increase in knowledge concerning the interactions between social, economic, and ecological processes (Rotmans 2001, 29-30).

Social studies adopt different analogies, based on different *metaphors*,[8] to get to grips with complexity. Simplistic, yet illustrative *dichotomies* have been proposed to distinguish between general approaches to scientific inquiry, such as that between the 'natural science approach' and the 'human science approach' (Leskinen 1994, 12), between analogies based on the Newtonian mechanics on the one hand, and coevolution, on the other (e.g. Norgaard 1994, 63-64; Passet 1996; Maréchal 2000), or between Laplacian and dialogical approaches to the nature and purposes of scientific knowledge (O'Connor 1999; Kallis 2001, 2). Such dichotomies reflect the 'paradigm wars' or 'Science Wars' (see e.g. Flyvbjerg 2001, 1-5) waged within the social sciences between the positivistic research approaches and its alternatives, (e.g. hermeneutical, phenomenological, realistic, post-modern or 'post-positivist' approaches). Evaluation has been no different from other social sciences in this respect, as we shall see later on in section 4.2. The mechanistic analogies typical of the 'natural science approach' see social phenomena to consist of interactions between stable, unchanging parts, governed by causal regularities, whose outcomes are reversible. Once the regularities of a particular system are known, the reactions of the system are predictable, and thus controllable through planned interventions. (Norgaard 1994, 63-64; Passet 1996; Maréchal 2000). Positivism thus adheres to the possibility of objective interpretations of reality through observation, scientific method and 'laws' of cause and effect (Khakee 2003, 341).

The main criticisms against this type of approach concern the incapacity of the linear models to address uncertainty and complexity as inherent elements of both societal and natural phenomena. In the latter, uncertainty stems from the irreversible, synergistic and cumulative effects. The *complexity of the social world* is at least as important: motivations behind human behaviour cannot be reduced to maximisation and economic rationality, but random elements and human intentionality are constant sources of uncertainty in society. (e.g. Hahtola 1990; Leskinen 1994, 13-14; O'Connor 1999; Söderbaum 2000, 7-14; Hinterberger et al. 2000; Khakee 2003, 341.)

This study suggests that a coevolutionary framework[9] provides a fruitful alternative to seeing societal phenomena in terms of mechanistic analogies. More specifically, coevolution, as a useful

---

[8] Hodgson (2002, 263) points out that "[M]etaphor is different from, and prior to, any analogy. Metaphors make connections between different domains of discourse. Analogies involve more detailed formal similarities using the broad connections established by metaphor. By doing this vital work, and by helping to form analogies, the influence of metaphor is neither superficial nor merely preliminary."

[9] Swaney (1985; 1987) was the first to introduce the concept in institutional environmental economics, but the notion has later on been adopted by a number of others, notably Norgaard (1994), as a foundation for the treatment of environmental problems (van den Bergh and Gowdy 2000, 42; Petit 2002, 195-196). In economics, the co-evolutionary

basic analogy, would imply the abandonment of the idea of one-dimensional causation. The coevolutionary view emanates from the merging of ecology and evolutionary biology, consisting of the "assumption that 'evolutionary' processes involve ongoing or periodic novelty and creativity, thus generating and maintaining a variety of institutions, rules, commodities and technologies" (Hodgson 1999, 131). In contrast with notions of adaptation within a passive environment, all sense of directionality is lost in a coevolutionary process, as species do not evolve towards a pre-determined physical niche but rather in response to changes in each other (Norgaard 1994, 83).

An evolutionary perspective – lying at the basis of the concept of coevolution – is particularly well suited to analysing questions of sustainable development, since it allows one to address issues such as diversity, risk minimising, path-dependence and lock-in[10] in the context of uncertainty, unpredictability, and complex systems – issues largely ignored by mainstream economic theory (Rammel and van den Bergh 2003). The coevolutionary view gives great importance to history, instead of searching for universally valid, ahistorical regularities (Söderbaum 2001, 186-187). The idea of open systems allows for emergent properties, i.e. phenomena, whose explanation at one tier cannot be reduced entirely to phenomena at lower levels[11] (Hodgson 1988, 10, 57; 1999, 139-141).

This coevolutionary framework will in the present study provide the general frame of reference within which the influence of environmental policy evaluations and institutional change will be analysed. The framework strongly draws upon institutional economic theory, the basic principles of which we shall describe in the following.

---

system consists of a continuous interaction between values, organisation, technology, environment, and knowledge – interaction whose outcome is always unpredictable (Norgaard 1994).

[10] Path dependence involves one form or another of the idea that 'history matters': where we are today is a result of what has happened in the past. A minimal form of path dependence is present whenever there is an element of persistence or durability in a decision. Stronger versions of path dependence refer to decisions that lead to economically inefficient decisions as economic agents adjust their decisions to prevailing institutions. (Liebowitz and Margolis 1998.) Lock-in takes place when an undesired situation cannot be changed without significant costs – e.g. when a client cannot change vendor without incurring significant costs, be they real or perceived, or when a technical solution cannot be replaced by another one that is more efficient, effective, or otherwise considered 'better' by the society, because the old solution has gained a dominant position in the market and the prevailing institutional structure has been adapted accordingly.

[11] Notwithstanding the numerous adjustments made over the years to the basic assumptions of modern mainstream economics, including the introduction of concepts such as 'bounded rationality', the theoretical framework and analytical tools developed within this school of thought rely on the idea that all explanations can be reduced to the rationally acting, utility-maximising individual having fixed preferences and possessing perfect information. The individual hence simply reacts in a programmed and optimising fashion to the economic environment. Reductionism in Marxism implies explaining all societal phenomena by reference to the 'economic superstructure', and 'iron laws' of history, giving it a strongly deterministic tone. (Hodgson 1988, 10, 57.) However, at least some of the current-day Marxists deny that Marx's writings would necessarily lead to a deterministic vision of society (e.g. Burkett 1999, 2).

2.2 Foundations of institutional economics

Economic scholars from different theoretical orientations have increasingly recognised the importance of institutions in shaping the evolution of economic phenomena (see e.g. Hodgson 1988; Bazzoli 1999; Williamson 2000). This development can be seen as a response to critique from within and from outside the modern mainstream paradigm[12] against the neglect of institutions as an object of analysis. Among the economists claiming an institutionalist title, the underlying assumptions concerning history, human rationality, the character of the social world, and the precise definition and role of institutions vary from one school to another. Public Choice (e.g. Buchanan and Tullock 1962; Tullock et al. 2000) and New Institutional Economics (e.g. Olson 1965; Alchian and Demsetz 1972; Williamson 1985; North 1990; 1996a; 1996b; Eggertsson 1991; Williamson and Winter 1991) explicitly endorse the modern mainstream theory, but modify and extend it to allow it to come to grips with a range of issues previously beyond its reach. By and large, they do not share the premises adopted in this study. By contrast, what shall in the following be referred to as 'institutional economics' or 'institutionalism' is the school of thought that has its intellectual origins in the 'old' American early institutional economics of Veblen, Commons and Ayres, whose thinking was clearly at odds with a number of the modern mainstream assumptions.[13]

Institutionalism does not constitute a uniform 'school' of economics that would share a common conceptual framework providing a solid basis for an axiomatic and formalised set of theories. Rather, what unites institutionalists is their approach to economics as a science. At the most general level, institutionalism can be characterised as holistic, systemic, and evolutionary. At a more concrete level, it attaches a particular role to power and conflicts in economic processes. (Wilber and Harrison 1978, 71.) To a large extent, institutionalism defines itself in opposition to the

---

[12] 'Modern mainstream economics' is in this thesis adopted instead of 'neoclassical economics' as the term describing the core of economics that is distinct from a variety of heterodox schools, including institutionalism. The terminological preference for 'modern mainstream' instead of 'neoclassical' stems from the fact that the modern mainstream approach can be distinguished from older approaches that also used the neoclassical label, and because modern mainstream economists have explored a wide range of assumptions and approaches that are not in line with the older neoclassical approaches in use until the 1930s (Boven 2003, 26-27).

[13] This study draws on the thinking of institutional economists such as Geoffrey Hodgson (1988; 1999; 2002), Daniel Bromley (1991), Allan Schmid (1978), Warren Samuels (1971), Marc Tool (1979) (see also Samuels and Schmid 1981). In the field of environmental economics, one can also mention the 'Nordic school' of institutionalism, exemplified by the works of Peter Söderbaum (2000; 2001), Kauko Hahtola (1990), Antti Leskinen (1994) as well as Hans Opschoor and Jan van der Straaten (1993). Many scholars identifying themselves as 'ecological economists' are likewise close to institutionalism and the boundaries between the two are sometimes fluid (see e.g. Söderbaum 2000, 18-25). Some institutionalists have been more inspired by Keynes and Marx than the Veblenian heritage (e.g. Hodgson

dominant modern mainstream economics,[14] notably by rejecting its methodological individualism, rational conception of action, reductionism, and exclusive reliance on one-directional causal explanations. It contests the idea of economics as a 'positive', value-free science, and considers that standard economics' failure to recognise its value commitments and the refusal to take into account power relations amount to an apologia for the status quo and the vested interests (Klein 1987, 1348).

Beyond the fundamental metaphor of coevolution, the characteristics of institutionalism of most relevance for the present study are the assumptions of man as a 'cultural animal', a particular conception of institutions, the omnipresence of power and conflict, and the importance given to values and ethical considerations.

2.2.1 View of man: what motivates human action?

One of the most fundamental issues defining a research approach concerns its view of man, notably the motivations behind human behaviour. This study rejects the so-called 'rationalist conception of action' (Hindess 1977), which interprets all human action as being governed by rational calculations aiming towards maximisation[15] of individual utility. The rationalist conception of action has been extensively criticised for its methodological inconsistency and neglect of evidence from psychology and other social and political sciences. (e.g. Koestler 1967, 238; Giddens 1984, xxiii; Hodgson 1988, 112-115; Bromley 1991; Finnemore and Sikkink 1998, 888; March and Olsen 1998; Vatn 2002).

1988, 24). For a review of institutionalism's roots, its present state, and a comparison between the 'old' and the 'new' versions of institutionalism see Rutherford (1996; 2001).

[14] In general terms, such an approach assumes rational, utility-maximising individuals with stable preferences; focuses on equilibrium based on mechanistic analogies; adheres to reductionism and monism; sees markets as the sole mechanism of coordinating the isolated acts of individuals; hardly pays attention to history, information problems and transaction costs; and perceives welfare as the sum of individual utilities (Williamson 1985; Pearce and Turner 1990, 10; Norgaard 1994, 65-66; Boyer and Saillard 1995, 9; Faucheux and Noël 1995, 28-39; Hodgson 1999, 29, 131; Maréchal 2000, 51-52; Söderbaum 2000, 8-9). It sees sustainable development essentially as a problem of integrating the environment with the economy, treats environmental assets like any other commodity, and relies on monetary valuation to determine the 'real' price of the environment. It applies discounting in order to compare present costs and benefits with those occurring in the future. These methods would enable an economist to identify the appropriate (optimal) level of environmental protection, and the policy-maker to use taxes, subsidies or create markets for environmental goods so as to translate these imputed prices into real-life prices. (see e.g. Mebratu 1998; Bartelmus 2000.)

[15] 'Satisficing' has been proposed as an alternative, yet even satisficing fails as an exhaustive explanation of human behaviour. Deeper explanations are needed to interpret phenomena such as the 'paradox of voting', i.e. to explain why people vote, even though the rewards from the outcome of a single vote are in most cases likely to be extremely small (Hodgson 1988, 92-93).

Institutionalists consider people as 'cultural animals', whose behaviour cannot be explained solely in deterministic terms. Rational calculation obviously plays a role in guiding human behaviour, but it is only one among many motivations. This conception can be seen as the most distinctive aspect of institutional economics, and is often used as the fundamental demarcating line between institutionalism and New Institutional Economics (Dequech 2002).[16] The essence of this postulate is that there is no single logic of action, but different context-dependent motivations and forces guide human behaviour. These forces include social, socio-psychological, cultural and biological factors, in particular ideology and ethics. Human action is being guided as much by custom, habit, routine and creative role-play as by instrumental calculation (Hodgson 1988, 80; Heap 1994).[17] Preferences are not stable, given, exogenous variables, but are instead being constantly formed in a complex process of social interaction. Preference functions are also affected by experience and means often affect ends.[18] (Hodgson 1988, 97.)

2.2.2 Institutions

The above discussion already pointed at the importance of institutions in shaping human behaviour. Along with the conception of man, the specific conception of institutions is what most clearly distinguishes institutionalists from other schools of economics. Different scholars have defined institutions in slightly different ways.[19] This thesis adheres to the conception taking into account the multiple purposes of institutions, as opposed to the New Institutional Economics and many 'common sense' perceptions that see institutions simply as constraints on the freedom of human

---

[16] However, Dequech (2002) points out that while this way of contrasting institutionalism from the New Institutional Economics still continues to be very useful in most cases, some authors associated with the latter are beginning to recognise the deeper roles of institutions in shaping individual behaviour.

[17] In a similar vein, Spash (2000) has suggested that individuals often act as 'moderate deontologists', having ethical and rights-based convictions about desirable behaviour, which are nevertheless moderated by considerations relating to the consequences of actions. For instance, an individual may consider harming an innocent human or non-human being unacceptable up to a certain threshold defined in terms of the good expected to result from the act.

[18] This renders doubtful the commonplace claim that economics only deals with finding the best means to achieve given ends.

[19] Thorstein Veblen (1919, 239) defined institutions as "settled habits of thought common to the generality of men". In the institutionalist writings, institutions are variously understood as "regular patterns of behaviour, which embody values and shared habits of thought and which enable social action" (Kallis 2001, 6), "the rules, i.e. predefined patterns of conduct, that the members of a social group have generally accepted" (Hukkinen 1999, 16-17), or: "an institution is a regularity of behaviour or a rule that is generally accepted by members of a social group, that specifies behaviour in specific situations, and that is either self-policed or policed by external authority" (Rutherford 1996, 182). The definition given by March and Olsen (1998, 948) is very similar: "a relatively stable collection of practices and rules defining appropriate behavior for specific groups of actors in specific situations." Söderbaum (2000, 22; 2001, 181) refers to 'institutional arrangements' as a broad category including both formal and informal rules as well as organisation and power relationships. Further, Hodgson (1988, 10) defines a social institution as "a social organization which, through the operation of tradition, custom or legal constraint, tends to create durable and routinized patterns of behaviour". In general terms, institutions can be either formal or informal; they embody values, and regulate behaviour through habits and norms.

agent – as 'rules of the game'. Instead, institutions determine 'the way the game is played'. Institutions are what enable humans to act, they are constitutive of the individual and provide social identity as well as meaning to human action (Vatn 2002, 7). Institutions are the substance, rather than merely the boundaries, of social life (Hodgson 1988, 134). Economic change is seen as a process of institutional change, and institutions are understood as endogenous to the open economic system.

Institutions can be seen to exercise three main functions. Firstly, in the *restrictive function* institutions constrain economic behaviour. The *cognitive function* refers, firstly, to the information that institutions provide to the individual, including the indication of the likely action of others, and secondly, to the 'deeper cognitive function' of influencing the very perception that people have of reality, i.e. the way people select, organise, and interpret information. This conception is manifest for instance in the approach highlighting the importance of 'culturally shared mental models' in learning processes (e.g. Denzau and North 1994; Hukkinen 1999). Finally, institutions have a *'motivational'* or *'teleological'* function, influencing the ends that people pursue. This runs directly against the modern mainstream assumption of fixed preferences. (Dequech 2002, 566.)[20]

Human ability to learn from experience is a central feature of the behavioural assumptions in this study. Therefore, human action is neither simply determined by outside circumstances or institutions beyond the influence of the individual, nor is man completely free from institutional influences to pursue his aspirations. Just as institutions have their constraining, cognitive and motivational/teleological functions, concerning not only the means to achieve given ends, but also the ends themselves, so does learning apply to all these levels, affecting both the means and the goals that individuals set for themselves. Learning does not necessarily lead to optimising, maximising behaviour, but may involve, for instance, modifications in goal structures. Furthermore, and most importantly for the purposes of this study, learning takes place not only at the individual, but also at the institutional and organisational levels. (e.g. Hahtola 1990; Hodgson 2002.) Language, concepts, and prevailing ideas play a determinative role in shaping these learning processes. This perspective considers that 'words do count', and endorses in particular the critique against the dominance of the kind of economic rationality, which tends to subordinate all human activities to market relations (see e.g. Harribey 1997, 34; Nelson 2003). Hukkinen (1999, 20) emphasises the

---

[20] However, Dequech (2002, 569) argues that recent writings of Douglass North, a prominent writer in the New Institutional Economics tradition, demonstrate a willingness to admit that institutions "may channel unchanging incentives or, more radically, that purposes or preferences are themselves changed by institutions".

tension between formal and informal institutions as the drivers of change, noting the considerable resilience of the latter as opposed to the relative facility with which formal institutions (laws and regulations, for instance) can be modified. This thesis will address the question of the extent to which evaluations are capable of affecting not only the formal but also the informal institutions – the routines, habits, customs, and frameworks of thought. Hence, learning takes a central place as a vector of institutional change.

2.2.3 Values, objectivity, truth, and ethics

The impossibility of drawing a sharp distinction between facts and values was already evoked in chapter 1. This perception of social science as an essentially value-laden endeavour is inherent also in institutional economics. Classical economics, as represented by the ideas of scholars such as Adam Smith, Thomas Malthus, John Stuart Mill, and Alfred Marshall, was a discipline designed to "make the world a better place […] providing a higher standard of living, most particularly for those who had least resources" (Goodwin 2003, 1). This ethical character of economics was weakened in the 20th century along with the 'neoclassical revolution', aspiring to construct economics as a 'value-neutral' science, concerned exclusively with the most efficient means of attaining given ends, i.e. studying the optimal allocation of scarce resources, while remaining neutral with regard to the objectives of the society. (Maréchal 2000, 128-129; Goodwin 2003.)

However, for instance, the criteria for objectivity postulated by modern mainstream economics[21] are ultimately founded on a common consent concerning the subject matter and analytical instruments of economics – a consent that is necessarily a value-laden agreement (Deblonde 2000).[22] The value-laden character of economics, and any scientific endeavour for that matter, begs the question of which ethical values and societal objectives research should be based on, what values are to guide scientific inquiry, and what kind of institutional change the society should aim at. The relevant question is not how to separate the positive from the normative, since any act is normative by nature, and the 'evidence' provided by facts always depends on the frame of interpretation, but

---

[21] Deblonde (2000) has listed these as 1) the empirical testability of the data flowing from the assumption of individual preferences; 2) the logical consistency of the axiomatic system based on the rational conception of action; 3) the empirical relevance of the assumptions; and 4) the mathematical exactness of calculations conferred by the use of money as a common unit of measurement.

[22] As will be seen later on, in section 4.4.1, the issue of objectivity is one that continues to arouse debates also within the evaluation community, notably in debates concerning the role of the evaluator ranging from a view seeing evaluator as a neutral analyst and observer, to those considering the evaluator's role to be an advocate of a particular position or group in society.

rather "whether one is aware of the normative nature of acts and whether the normative effects of acts are beneficial or detrimental to individuals and their communities" (Stevenson 2002, 265).[23]

Institutional economics therefore includes among the subjects of research the way in which the culture of a society, including the language and concepts utilised to describe and explain social phenomena, can enable the manifestation and perpetuation of adverse mental states and harmful habits. Such an approach runs directly against the perception cherished by modern mainstream economists of economics as a neutral, value-free science, and also stresses the importance of language and symbolic aspects in shaping societal phenomena. (Stevenson 2002.)

2.2.4 Power

Central to the institutionalist view adopted here is that economic and political processes are seen as intimately linked and that power and conflict are pervasive aspects in all social interaction. Power is a multidimensional phenomenon that cannot be reduced to the special case of monopoly power – according to Rotschild (1994) the only form of power fully addressed in modern mainstream economics. Institutions can be seen as the medium and embodiment of power. On the one hand, they constrain action, but on the other, they have an enabling role, without which meaningful action would be impossible. In the words of Foucault (1982), power "is not a renunciation of freedom, a transference of rights, the power of each and all delegated to a few". Indeed, power is *not* the same as violence because the opposite pole of violence "can only be passivity" (ibid.). Hence, Westergaard and Resler (1976, 142-146) note that "there is power inherent in anonymous social mechanisms and assumptions – in 'social institutions' – not just in individuals or groups […] Power is to be found more in uneventful routine than in conscious and active exercise of will." Power play between interest groups consisting of rationally behaving individuals seeking to maximise their utility certainly reveals a part of the role of power in society, but in its exclusive focus on the 'negative' forms of power, it fails to grasp the whole breadth and meaning of the concept.

Flyvbjerg (2001, 116-118) argues that many strands of power theory in social and political science are united in their concern with power in terms of possession, sovereignty, and control – something

---

[23] Tool (1993; cited in Stevenson 2002, 267) described the *normative objectives of institutional economics* as including: 1. Providing for the continuity and efficiency of the provisioning process. 2. Non-discriminatory participation in the social process. 3. Expansion in the growth of warranted knowledge. 4. Full development and use of individual capabilities and skills. 5. Assurance of coevolutionary development of biotic and social communities. 6. Enhancement of democratic self-governance.

to be possessed, won, held or lost – in short, power as entity. By contrast, the Foucauldian view is to see power as force relations, seen in terms of its concrete application in strategies and tactics. In the latter perspective, power is exercised rather than possessed, and a scholar should therefore ask how power is exercised, instead of asking where the power resides. Both perspectives are legitimate and necessary, but the latter, relational, view of power is commonly underestimated. Flyvbjerg (2001, 131-132) therefore proposes a synthesis, seeing power as characterised by six features:

1. Power is seen as productive and positive and not only as restrictive and negative.
2. Power is viewed as a dense net of omnipresent relations and not only as localised in 'centres' and institutions, or as an entity one can 'possess'.
3. The concept of power is seen as ultradynamic; power is not only something one appropriates, but also something one reappropriates and exercises in a constant back-and-forth movement in relations of strength, tactics, and strategies.
4. Knowledge and power, truth and power, rationality and power are analytically inseparable from each other; power produces knowledge, and knowledge produces power.
5. The central question is how power is exercised, and not only who have power, and why they have it; the focus is on process in addition to structure.
6. Power is studied with a point of departure in small questions, 'flat and empirical', not only, nor primarily, with a point of departure in 'big questions'.

Another way to look at the issue is to distinguish between three dimensions of power. At its simplest, power means possessing something considered of value in the society (e.g. money, knowledge, education, occupation). This form can be called 'resource power'. The second dimension, 'structural power', stems from a central position in the networks and hierarchies of power, knowing the right people, etc. This form is explicitly relational, as opposed to resource power, which is tied to individual characteristics and possessions.

Finally, the most intractable form, and the one that is of particular interest here, is symbolic power, which resides in concepts, language, dominant 'mental models', prestige and reputation. Discourses transfer and produce power. They can reinforce power, but they can also subvert and conceal it. In the same way, secrecy and silence can mask power, but can also weaken its grip. Likewise, power and knowledge directly imply one another. (Flyvbjerg 2001, 124-125.) An example is the power of language and rhetoric (e.g. McCloskey 1994; Keppler 1998; Luks 1998) in shaping the thinking and action of individuals, and its capacity to direct attention either towards or away from the nature of

the human condition. Söderbaum (2001, 182), for instance, emphasises the importance of studying patterns or "schemes of interpretation" – their origin and change, competition between different groups advocating differing schemes of interpretation. From this perspective, the critique from the Marxist and socialist economists against the encroachment of market logic into all aspects of human life is relevant: Marx criticized "vulgar economics" for objectifying commodities, and commodifying subjects, thus tending to commodify the whole society, including human and emotional relationships (Honkanen 1999, 99-100). In the same line of reasoning, Spash (1999, 428) affirms that "free market systems educate individuals to act as selfish hedonists and create self-perpetuating power structures which reinforce inequity".

The institutionalists argue that much of the language of modern mainstream economics abstracts from the human condition, by talking about labour as an input; utility as means of material gratification arising from axiomatically ordered immutable preferences; marginal rates of utility, production, substitution, and transformation; allocative and technical efficiency; etc. (Stevenson 2002, 274.). Harribey (1997, 34) argues that the dominance of economic rationality leads, on the one hand, to uniformisation of lifestyles and collective worldviews, and on the other, to the exclusion of poor countries and of the poor within the developed countries. While these consequences of assumptions contained in economic theory may seem irrelevant to an economist adopting an 'value-free' stand, from the institutionalist perspective they are seen as prime examples of the motivational or teleological function of institutions: through hidden power play and the power of language and concepts economics as a science reinforces a certain world-view and prompts people to pursue certain goals (see also Boven 2003).

Symbolic power is highly relevant for the study of policy evaluation. Indeed, the 'power' of policy evaluations and international organisations partly resides in their potential to provide knowledge, prestige and legitimacy to certain groups or individuals (e.g. Yee 1996; Finnemore and Sikkink 1998). Evaluations can strengthen or weaken the position of actors in different networks, but their power also stems from their influence on the language, concepts, and reputation of actors.

2.2.5 Deliberative democracy

Participation and democracy were already evoked above as among the normative objectives of institutional economics. The institutionalist conception of an 'ideal society' is ultimately based on the 'instrumental value principle', which recommends society to adopt what provides for "the

continuity of human life and the noninvidious recreation of community through the instrumental use of knowledge" (Tool 1979, 293). This principle is widely interpreted as fostering pluralism and deliberativeness, thereby making "democratic participation the central element of social value theory" (Dugger 1995, 195). Nevertheless, this is not alone very helpful in distinguishing 'good' participatory democracy from a 'bad' one. Deliberative democracy is therefore here suggested as a kind of a 'meta-ideology' for institutionalism and an essential element of sustainable development.

Elster (1998b, 8), in reviewing different definitions of deliberative democracy, argues that there is a consensus among scholars "that the notion includes collective decision making with the participation of all who will be affected by the decision or their representatives" (the democratic element) and that it includes decision making by means of arguments offered *by* and *to* participants who are committed to the values of rationality and impartiality" (the deliberative element). This study adopts the definition provided by O'Neill (2001, 6) of deliberative democracy "as a forum through which judgements and preferences are transformed through reasoned dialogue against the picture of democracy as a procedure for aggregating and effectively meeting the given preferences of individuals". Unlike many conventional ideas of democracy, it replaces the idea of given preferences of individuals, to be 'revealed by an expert, by a conception that sees preference formation as one of the key roles of a deliberative, democratic process. Hence, deliberative democracy consists of three core propositions: 1) Some form of communicative reason, rather than bargaining between competing interests should be guiding the political procedure; 2) The essential political act – the giving, weighing, acceptance or rejection of reasons – is a public act, as opposed to the purely private act of voting; 3) It is democratic deliberation, not deliberation without modifier. (Parkinson 2001.)

Deliberative democracy has its origins in Habermas's critical theory, but has since then gained an important position in political science more generally (e.g. Dryzek 1990; 2000; Bohman and Rehg 1997; Elster 1998a; Fishkin and Laslett 2003). The foundation for deliberative democracy is provided by what Habermas (1984) has called the 'ideal speech situation', whose conditions are the following (Flyvbjerg 2001, 91):

1. No party affected by what is being discussed should be excluded from the discourse (the requirement of generality).
2. All participants should have equal possibility to present and criticise validity claims in the process of discourse (autonomy).

3. Participants must be willing and able to empathise with each other's validity clamins (ideal role taking).

4. Existing power differences between participants must be neutralised such that these differences have no effect on the creation of consensus (power neutrality).

5. Participants must openly explain their goals and intentions and in this connection desist from strategic action (transparency).

It is notable that the ideal speech situation requires participants to have communicative competence such as mastery of the rules of formal logic, linguistics and interaction (Palerm 2000, 584). Moreover, Flyvbjerg (2001, 91) argues that, taken together, the five requirements imply a sixth one: unlimited time.

Habermas's model has been criticised on several counts. For instance, Foucault (1982) has argued that a state of communication free of domination is absurd, and that any attempt at a normative ideal failing to recognise its existence, is an embodiment of this very domination, which has thus succeeded in disguising itself. In the same vein, Flyvbjerg (2001, 93) notes that the fundamental political dilemma in Habermas's thinking is that he "describes us the utopia of communicative rationality but not how to get closer to it." Moreover, the theory of communicative action has been accused of denying the differences in plural societies, lacking consideration for social movements, and disregarding situations when no consensus can be reached. (e.g. Webler 1995; Palerm 2000, 585-586.) The Habermasian model can be seen as a futile attempt to seek an 'Archimedes' point', a universally valid ideal against which one could analyse any social situation (e.g. Hahtola 1990; 1997).

Notwithstanding its shortcomings, the Habermasian concepts are here adopted as the basis for an ideal model of deliberative democracy, as they lay out what preconditions must be fulfilled for a decision or a political process to be termed 'democratic'. However, to complement such idealising notions, we must keep in mind the pervasiveness of power in all social interaction, as pointed out in the previous section. Moreover, deliberative democracy is here seen as a complement to, instead of a substitute of representative democracy.

2.2.6 Methodological pluralism

Finally, in the practice of science and policymaking, the philosophical and theoretical premises presented above translate into methodological pluralism, based in particular on the rejection of the idea of the 'unity of science'[24], i.e. the belief that since scholars in the various disciplines are studying different aspects of the same reality, as they advance, accumulating truth and weeding out falsehoods, their understanding must merge into a unified understanding of the whole (O'Connor 1999; Norgaard 2003). Institutionalism rejects such a notion, and argues instead that different disciplines and approaches carry with them fundamentally different ways of looking at the world, and that the pictures they provide are more often than not fundamentally different from each other (Norgaard 1994, 65-66.) As Latouche (1984, 19-24) has noted, "the plurality of incompatible and antagonistic discourses is a situation that cannot be gone beyond." Methodological pluralism assumes that there is "no universally applicable, logically compelling method of theory appraisal" (ibid.).[25] The inherent complexity of coevolutionary systems makes it impossible to grasp the phenomena through a single approach, or a single discipline, and therefore different methodological and conceptual perspectives are needed in order to provide a varied picture of the phenomenon (Norgaard 1989; 1994). The adoption of methodological pluralism can also be seen as a logical consequence of the roots of institutionalism in the American pragmatic philosophy, emphasising the importance of action and its practical consequences in determining the 'truthfulness' of ideas (Gauchotte 1992, 3; Dequech 2002, 565). Such an approach is illustrated in Gunnar Myrdal's (1975, 142) remark that "[i]n reality, there are no 'economic', 'sociological', or 'psychological' problems, but just problems, and they are all complex."

The purpose of a pluralistic approach is to critically analyse the methodological content of alternative theories with the aim of promoting novelty and dialogue. Hence, the opponents of any given theory should adopt an approach of internal critique, that is, try to assess the theory from the perspective of the epistemological premises of their opponents, while the proponents of the theory should attempt to show how their theories can be compared with those of other groups. The advocates of a theory would be motivated to engage in such comparisons, as this would probably be the only means available for them to incite potential adherents to join their ranks. Obviously, a

---

[24] Also referred to as 'monism', or 'Laplacian reconciliation' of knowledge (Norgaard 1994, 65-66; O'Connor 1999).

[25] Methodological pluralism has its origins in the critique against positivism by the 'growth of knowledge' philosophers such as Thomas Kuhn, Imre Lakatos and Joseph Agassi.

consistent application of pluralism leads to the paradox that one must also consider methodologies that are at odds with pluralism itself. (Caldwell 1982, 244-249; Norgaard 1989, 51.)

Methodological pluralism does not imply extreme relativism in the spirit of 'anything goes'. Neither does it require that a wide range of methodological approaches be applied in a single study or even a research programme. Pluralism should rather be understood as an invitation to openness towards alternative views, consciousness about the premises underlying one's own approach, and preparedness to submit the latter to constant critical examination.

2.3 Summary: key characteristics of institutionalism

The preceding presentation has been an attempt to delineate the factors of institutionalism that seem the most relevant for the purposes of this thesis, that is, provide a basis for a definition of an 'ideal model' of evaluation of environmental policy in the pursuit of sustainable development. The following features stand out as the key elements of such an approach (see also Medema et al. 1999, 423-424):

Box 2.1. Key characteristics of institutional economics relevant for evaluation of environmental policy for sustainable development.

- Conception of man as a 'cultural animal', whose behaviour is motivated by a multitude of factors, instead of the assumption of rational actor, maximising his individual utility.
- Rejection of mechanistic analogies in favour of a more holistic and pluridisciplinary approach based on coevolutionary ontology.
- Emphasis on institutions, their two-way interaction with human behaviour, and their multiple functions in shaping human behaviour and life of the society.
- Abandonment of the ideal of value-free science in favour of a conception of social science as a quintessentially value-laden activity.
- Sensitivity to power and conflict as a multidimensional phenomenon, and especially the importance of language, ideology, and symbols in shaping societal phenomena.
- Deliberative democracy as a leading normative ideal.
- Adherence to the principle of methodological pluralism instead of the assumption of a single 'correct' scientific approach.

These principles will serve as a guide in the rest of the study, and a foundation for the next chapters, devoted to the definition of an institutionalist perspective to sustainable development and environmental policy evaluation.

Chapter 3

Institutional change towards sustainable development

The previous chapter laid out the fundamental premises of the present study, and defined institutional change towards sustainable development as a key reference point of environmental policy evaluation. The chapter likewise explained what is meant by institutions in this study, i.e. what needs to change. However, what was not made explicit was the direction of such change: towards which general values and precise objectives should institutional change contribute? This chapter lays the basis for the analysis of the 'substantive' side of our research question by defining what is understood by institutional change towards sustainable development.

The argument pays special attention to conceptualisations of sustainable development in economics. Two reasons can be evoked to justify such choice. First, economic reasoning has played an important role in guiding, or at least rationalising, economic development along an unsustainable path. There are good reasons to suspect that economics, as it evolved within existing institutions, is at the heart of the problem of 'unsustainability'. (Norgaard 1994, 18; Boven 2003.) Second, since the OECD is, after all, an organisation for *economic* cooperation and development, analysing, from the perspective of institutional economic theory the way the organisation conceptualises sustainable development should provide clues as to the compatibility or not of the 'OECD message' with sustainable development.

After an introductory section identifying the key issues relating to sustainable development, we shall look at the uneasy relations between sustainable development, progress, and modernity, and

the question of what attitude should be adopted with regard to the limits imposed by biophysical and ecological life-support systems to the scale and character of human activities. The chapter then looks at the dimensions of sustainable development and their relations with each other, with a special attention to the social dimension of sustainability. Adopting the view that sustainable development consists of a combination of substantive issues relating to the biophysical and ecological 'limits', and procedural matters having to do with the processes of decision-making in the society, we shall address the question of measurement of sustainable development. The chapter concludes with a list of principles for an 'institutionalist' perspective on sustainable development.

3.1 Definitions and key issues of sustainable development

Sustainable development has been defined in a multitude of ways since the term was permanently introduced into the vocabulary of international politics by the so-called Brundtland Commission (WCED 1987).[26] The Commission optimistically declared that "humanity has the ability to make development sustainable – to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs" (ibid., 8). Since then, the 'official' definition has remained essentially unchanged, while numerous actors – from academic scholars to various governmental and non-governmental actors – have tried to interpret and operationalise the concept.[27] As the examples described further down will seek to demonstrate, there are compelling reasons to believe that attempting to operationalise sustainable development in the traditional sense, implying the capacity to predict, steer and control social and natural systems may be a futile

---

[26] The antecedents of the concept can be traced back to the emergence of international environmental policy, particularly the first major international conference on the environment, the UN Conference on the Human Environment in Stockholm 1972. Without explicitly mentioning the concept, the final declaration of the conference contains a number of the principles of what was later to become known as sustainable development, e.g. future and present generations, integrated management of resources, sustainable resource use, quality of life, environmental responsibility, international cooperation, transfer of technologies, demographic aspects, participation, and the priority given to developing countries. The term used at the time was 'ecodevelopment' (Sachs 1980), which however can be seen to represent a greater rupture with the conventional view of development than the more consensual concept of sustainable development. In particular, ecodevelopment stressed more the need to find solutions that would be in harmony with local contexts, ecosystems and cultural conditions. The International Union for Conservation of Nature and Natural Resources (IUCN), in its World Conservation Strategy (1980), was the first to apply the term sustainable development at an international level. The 'developmentalist' perspective of sustainable development, in turn, was partly inspired by the UN General Assembly's declaration concerning the establishment of a 'New International Economic Order', in 1974. For a review of the emergence of the concept of sustainable development in international politics, see Zaccaï (2002, 117-180).

[27] Within the modern mainstream environmental economics, sustainable development has been defined as non-declining human welfare over time – therefore, a development path that makes people better off today but makes people tomorrow have a lower 'standard of living' is not 'sustainable' (Pearce et al., 1989), or simply "economic development that lasts" (Pearce 1998, 69). For Solow (1994, 25), sustainable path is one "that allows every future generation the option of being as well off as its predecessors". For a review of the various definitions of sustainable development see Pezzey (1997).

endeavour, owing to the enormous complexities and uncertainties involved. Such operationalisation would require overwhelming efforts of monitoring and calculation, not to mention the difficulties involved in attributing value to different components of welfare so as to enable optimisation or maximisation (Norgaard 1994; Hinterberger et al. 2000). In Norgaard's (1994, 22) words, "[i]t is impossible to define sustainable development in an operational manner in the detail and with the level of control presumed in the logic of modernity."

The impossibility of defining sustainable development in an operational manner does not, however, prevent us from identifying a set of key notions that unite most definitions. Despite some divergences of view on details, an overall consensus seems to prevail over the key elements of sustainable development. Among the most commonly cited ones are the consideration of intra- and intergenerational equity; long-term global vision; concern for environmental constraints; and integration of the various components of development. Irreversibility of ecological change, uncertainty and complexity, as well as the processes of technological change are often mentioned as additional key aspects. (Faucheux 2000a; Zaccaï 2002, 39.) Recently, 'decoupling' of environmental pressures from economic growth has been established as one of the main objectives at the EU and OECD level (e.g. OECD 2001e; EEA 2003). Zaccaï (1999) has identified nine aspects brought to the fore by the notion of sustainable development:

1. *Environmental problems* as a major cause of the actual development crisis.
2. *Planetary limits* in the face of the growth of population and human activities.
3. Need for a *multisectoral* (environmental, economic and social) and *multidimensional* (from local to global) approach to dealing with the problems.
4. *Need to integrate* the questions of environmental protection with those relating to overall development.
5. Importance of *technological progress* in enhancing sustainable development.
6. Compatibility of sustainable development with *economic liberalism*.
7. *Intergenerational equity*.
8. Need for changes in *consciousness and ethics*.
9. Need for *participation* by all sectors at all levels.

This description tries to capture the essence of the 'consensus view' on sustainable development, the aspects on which the participants in the debate tend to agree. The first two points refer to the question of the planetary limits dictated by the biophysical environment to the scale and nature of

human activities. The third point has to do with the relationships between the different dimensions, levels and disciplinary perspectives of sustainable development. Point four lies at the basis of the principle of policy integration. Points five and six affirm the compatibility of sustainable development with the long-standing ideals of Enlightenment, progress and modernity. The three last aspects are commonly dealt with under the heading 'social dimension of sustainable development'. The following section looks at how these issues are dealt with in institutional economics.

Let us start describing an alternative view on sustainable development by addressing the question of the relationship of sustainable development with progress and modernity, including the question of the 'limits' placed by the biophysical and ecological systems on the scale and nature of the human activity. The analysis then moves on to questions of integration – between different sectors of policymaking, different dimensions and scales of sustainability. Finally, we shall look at how the concept of sustainable development could be operationalised through indicators.

3.2 Progress, modernity, economic growth, and sustainable development

The disagreements between different approaches to sustainable development reflect differing convictions, beliefs and assumptions about the nature of development itself, more particularly the notions of progress and modernity. As the nine-point list of Zaccaï (1999) showed above, sustainable development is by essence a reformist concept implying the assumption that capitalism can indeed be made sustainable. This is, however, also one of the most contested aspects of the concept. Critique against this view has come in particular from the Marxist-oriented analysts and the 'anti-developmentalist' or 'post-developmentalist' authors (e.g. Latouche 1989; Rist 2001; 2001; Harribey 1997, 1998; Benton 2002). Thus, Benton (2002, 252) considers that the Club of Rome's 'Limits to Growth' (Meadows et al. 1972) played a crucial role in bringing about the concept of sustainable development, by launching the search "for a strategy that could address the ecological crisis, but within the framework of liberal democratic capitalism, and without calling into question its central cultural dynamic of consumerism". Further, according to Benton, "a key strategic problem was that some crucial environmental assets were located in Third World countries, and that some global ecological 'life support systems' were potentially threatened by the hitherto authoritative models of economic development in Third World countries" (ibid.). Likewise, while Touraine (1999) considers sustainable development as the third stage of a resolutely

modernised and democratised capitalism,[28] Lascoumes (2001, cited in Theys 2002b, 1) sees the concept as nothing but old ideas in a new packaging, a simple trick of marketing.

These contrasting views are likewise reflected in the opinions and policy prescriptions concerning the roles of economic growth and free trade in promoting sustainable development. In view of our discussion on the 'symbolic' form of power in the previous chapter, entailing the idea that 'language counts', the notions on progress and modernity merit a central place in the present analysis. This section examines views on progress and modernity as to their links with sustainable development. A brief introduction will be given to the concepts of reflexive modernisation and ecological modernisation, the latter being considered today by many as the leading theory describing the change in environmental governance since the late 1980s (e.g. Joas 2001, 227).

3.2.1 Modernity and its discontents

Modernity – the modes of the Western social life that have developed during the past 200 years, first in Europe, later all over the world (Wallerstein 1974; Giddens 1990, 1) – has on the one hand brought about growth of material production, cultural reproduction, health and welfare, and a democratic institutionalisation of power, while on the other, having created a host of new problems, seen as overwhelming and unavoidable by the opponents of the modernisation idea (Seippel 2002, 197-198).

Post-modernity, 'high modernity', and 'reflexive modernisation' (Giddens 1990; Beck 1992; Beck et al. 1994) are among the concepts that have been proposed to retain the good sides of modernisation while at the same time mitigating the negative ones, thereby allowing a shift to a development path closer to the ideals of sustainable development. Theories of the reflexivity of modernity do not contain a commonly accepted definition of the concept, but the different views converge around the idea that the institutions and hierarchical structures of the industrial society no longer are capable of responding to the present-day challenges. Reality needs to be reconstructed, yet the elements for such reconstruction process must be found in modernity itself. (Beck et al. 1994.) Lash defines reflexive modernisation as "having as its core assumption the progressive freeing of agency from structure" (ibid., 119). For him, reflexive modernisation is something of a 'third way', which offers an alternative to the 'Enlightenment utopia of metanarratives of radical

---

[28] Touraine distinguishes three stages of economic development: capitalism of accumulation, construction of a welfare state, and sustainable development.

change' and to the postmodern 'dystopic evolutionism' of Michel Foucault (ibid., 128). Giddens, in his idea of a move away from 'simple modernity' (the West since the Enlightenment)[29], foresees a demise of class politics, the emergence of new social movements aiming at the radicalisation of democracy and against surveillance and authoritarianism. The disappearance of the traditional certainties of life leads to a reflexivity of modern social life, meaning that "social practices are constantly examined and reformed in the light of incoming information about those very practices, thus constitutively altering their character" (Giddens 1990, 38).

The relevance of the idea of reflexive modernisation for this study stems on the one hand from its position as an 'umbrella theory' for ecological modernisation, and on the other from the fact that policy evaluation has been proposed as one of the tools favouring a transformation towards a genuinely transparent, reflexive society, or "cognitive democracy" (Joas 2001, 226, 249; Theys 2002a).

3.2.2 Ecological modernisation

The concept 'ecological modernisation' was first introduced by German social scientists such as Joseph Huber (1982) and Martin Jänicke (1985; 1988; 2000). While the dominating belief among the 'early environmentalists' since the late 1960s was that modernisation would inevitably lead to ecological crisis, ecological modernisation theory asserts that modernisation not only is in harmony with environmental improvement, but that in fact modernisation – and most notably capitalism – is a prerequisite for progress in environmental conditions (e.g. Buttel 2000, 61).[30] While ecological modernisation cannot be equated with sustainable development (see e.g. Langhelle 1998), it shares many of its basic ideas, in particular that of the compatibility of environmental protection with economic growth, and can be considered as a central way of conceptualising the environmental 'pillar' of sustainability.

---

[29] Giddens (1990, 55-63) defines simple modernity as consisting of four "institutional dimensions": a political/administrative power that is typically a liberal, representative democracy; an economic order overwhelmingly capitalist in form, with the communist regimes as a temporary variant; a relation to nature defined by modern science and industrial technology; and a state monopoly on the legitimate use of violence.

[30] The changes of society inherent in ecological modernisation resemble what some scholars have called 'paradigm shift' in environmental policies (Driessen and Glasbergen 2002). Moreover, the discussion on ecological modernisation is closely related to the burgeoning interest in the concept of 'governance', notably in its idea about an increasing role of non-state actors in policymaking. Despite its character as an inherently 'fuzzy' concept, governance can be considered as a heuristic tool helping understand and analyse state action in a profoundly transformed policy environment (Froger 2001, 37). In particular, 'participatory governance' has been suggested as a way of fostering sustainable development (e.g. Froger et al. 2001). In its focus on policymaking structures, no matter how broad categories of actors they are seen to encompass, governance is, however, a more limited concept than ecological modernisation.

Mol and Sonnenfeld (2000, 6-7) have identified five core elements in the ecological modernisation theories developed so far (see also Lundqvist 2000; Joas 2001, 237-246):

1. An increasing and changing role for *science and technology* in the service of system-oriented problem solutions and 'eco-technological' innovations.

2. An increasing role for *economic agents and the market*. Enterprises take environmental values more seriously, seek more eco-efficient forms of production, and replace end-of-pipe solutions by broader 'cleaner production' strategies. The state aims to facilitate the 'greening of growth'. Monetary valuation is used to integrate these concerns into decision-making. 'Win-win' solutions are sought for in the spirit of an idea that 'pollution prevention pays' (to complement the simple 'polluter pays principle').

3. A changing position and role of *social movements*. The environmental NGOs become part of the 'establishment' through professionalisation, specialisation, centralisation, and institutionalisation. New, more loosely organised, more radical activist groups emerge to take place of the old movements as protest groups. (see also Seippel 2002, 213.)

4. A transformation of the *role of the state*. The policy style changes from a centralised, static and hegemonic style to a decentralised, flexible, deliberative and consensual one. Environmental concerns become integrated into all sectors of the state apparatus. Self-regulation is adopted as one of the main instruments of environmental policy, which entails a greater role for the provision of information – for instance through indicators. Non-state agents participate actively in the policy processes at all stages, including implementation. Supranational, international institutions partially take over the tasks of the nation-state. The main task of the state is to establish a regulatory framework favourable to economic efficiency and environmental protection – in short, strategic planning and the promotion of structural change at the macro-economic level. Introducing an ecological tax reform is among the main concrete tasks of the government.[31]

5. An emerging new *belief-system* or *ideology*.

---

[31] Contrary to the rather common perception that the importance of the state in policymaking is diminishing as a result of globalisation processes, ecological modernisation stresses the transformation of the role of the state. The role of the state may even become more important, given the global nature of environmental problems (e.g. Langhelle 1998, 9; Theys 2002a; Joas 2001, 240-247).

3.2.3 Reflexivity and rationality: shortcomings in the theories of reflexive and ecological modernisation

Reflexive modernisation as well as its 'daughter' – ecological modernisation – has faced a number of criticisms. The ideas of reflexive modernisation have been accused of being essentially rooted in substantive rationality, relying on science, expert systems and knowledge, and rational cognition. Reflexivity would therefore mean nothing but 'more of the same' – more risks and more attempts to control, manage, and limit risks by the means characteristic of modernisation. (Mckehnie and Welsh 2002.) Instead of focusing only on such cognitive and institutional elements, reflexivity should embrace the cultural habits and meanings that are being created in social interaction and which produce specific kinds of cultural behaviour (Louhimaa 2002, 146).[32] Likewise, the theories of reflexive modernisation have been accused of adopting a linear-developmental view of historical change; assuming the existence of a single form of global modernity; and overlooking questions of power, as well as the crucial role of the "globalising processes of capital accumulation" and profit-seeking as sources of the problems of modernity (Benton 2002). This latter criticism, in turn, is closely related to the 'anti-' and 'post-developmentalist' critique against sustainable development, referred to above. While this perspective merits attention – economic power certainly plays a key role – it sees power in an overly limited perspective, sidelining its more symbolic aspects, thus tending towards excessive reductionism.

Beyond these Marxist-oriented criticisms, ecological modernisation has been accused of Eurocentrism, incapability of handling the social inequalities between developing and developed countries – created by the externalisation of environmental problems to the developing world – and excessive optimism concerning the possibilities of the market economy to adjust itself to sustainable patterns of production and consumption (Langhelle 1998; Schnaiberg et al. 1999). Finally, ecological modernisation theories have been criticised for failing to create links between the global and the local, putting excessive emphasis on the former. Thereby the theory would forget that the apparently global risks are heavily mediated by local physical and social factors, neglecting

---

[32] This criticism is close to the ecocentric critique, which stresses the impact our culture and its values have on consciousness and action, including political action, criticises the functional differentiation process of modernisation, and supports 'alternative rationalities' along normative and aesthetic lines (e.g. Eckersley 1992). Seippel (2002, 218-220) points out, however, that this view seems deficient in its tendency to overlook the positive sides of modernisation, the negative consequences of a return to a traditional society, material conflicts and the broader political and institutional context or the cultural 'master frames' predominant in modern society – mass media, culture, politics, or economics.

local knowledge in the pursuit of modernist generalisable knowledge (Joas 2001; McKechnie and Welsh 2002; see also Theys 2002b and Hukkinen 2003).

The social constructionist critique leads us to the path of institutionalists, as it contests the thesis that assumes a progressive rationalisation of environmental policy in the direction of modernisation. In short, "ecological modernisation appears to offer a rationalised track along which environmental policy travels, a vision quite at odds with the contingency of constructionist views" (Yearley 2002, 281). Yearley (2002, 282) accepts ecological modernisation as a *normative* prescription for environmental policy, but contests its *descriptive* value. Social learning does not take place in the straightforward manner proposed by advocates of ecological modernisation, that is, through demonstration and group pressure following successful environmental policy improvements by industry and governments. There is no simple directionality in social learning about environmental policies because the 'environmental' is itself a contingent, social construction. Moreover – and this is where ecological modernisation meets sustainable development, and what is critical for ecological modernisation as a normative concept – the 'scientification' of environmental problems does not necessarily yield the best or most correct construal of the problems facing society, and no measures of most environmental policy variables can be incontrovertible. Therefore, rather than being a theory of social development, ecological modernisation can be seen as a socially constructed discourse (ibid., 283), an ideology, "a category of discourse that is a flexible and powerful instrument for criticising the assumptions built into the first-wave of environmental protection" (Weale 1992, 75-79).

3.2.4 Institutionalists and progress

As opposed to the idea put forward by Yearley (2002), who accepts ecological modernisation as a normative description of the evolution of environmental policy towards sustainable development, there are reasons to doubt whether ecological modernisation represents an ideal worthwhile pursuing. The debate among institutional economists on the nature of history and progress illustrates these doubts.

Kallis (2001, 6) has distinguished two classes of institutionalist scholars, in function of their views on history and the role of technology in the process of development. The *'progressivists'* largely draw upon the tradition of Clarence Ayres and consider cumulative technological change inevitably as a form of progress. Critique towards technological development is seen invariably as a kind of

regress, most often attributable to an unfounded resistance against change by certain groups of the society – resistance stemming from superstition, religion or other ideological or emotional reasons. The progressivists therefore share with the modern mainstream economists the idea of history as a linear progress towards the 'better'. This conception is in line with Veblen's belief that humankind indeed was progressing towards instrumental type of behaviour, while the institutional behaviour was correspondingly becoming scarcer.[33] Consequently, DeGregori and Shepherd (1994) assert that "there is one fundamental pathway to sustainability and that is through sustainability of the technological process of continued resource creation".

Mirowski (1987, 1029) criticises this tradition of the "reification of technology as the sole category of legitimate knowledge, begun by Veblen," and "carried to its extreme conclusion by Ayres". He further notes that the central theme in Ayres' work was the tension and dichotomy between 'ceremonial' processes on the one hand, and 'technological' or 'instrumental' processes, on the other, which seems to reduce all social life into non-scientific and scientific endeavour. As a result, "the pale shadow of pragmatism has become – irony of ironies – a Popperian version of science." (ibid., 1029-1030.)

The *'holistic'* institutionalists have a view on history more in tune with the basic premises of the present study. Historical change is seen as contingent on initial institutional conditions and power structures, but at the same time partly unpredictable, because of the underlying complex, inter-dependent cumulative processes. What is to be considered progress cannot be determined independently of the prevailing social values. The history of humankind does not necessarily consist of steady evolution towards the better, but periods of decline and decadence are possible as well. (Kallis 2001, 6.) Given the absence of teleology, 'holistic' institutionalists turn their attention to the processes of change, instead of static conceptions of equilibrium. To take an explicit or static view of the desirable type and organisation of society is an unduly restrictive approach, since there is no ideal timeless state to which one can refer to as an appropriate end. The ideals are constantly changing and the society is always in a process of becoming. This does not, of course, mean that there would not be periods of stability, but that a balanced ontological importance should be given to stability, order, habit and the past on the one hand, and flux, chaos, novelty and the future on the other. (Klein 1987, 1365; van den Bergh and Gowdy 2000, 45; Nelson 2003, 52.)

---

[33] For Veblen, behaviour was instrumental, if its objective was to increase human knowledge. He argued that human history proves progress has taken place and that there was no reason to doubt why such progress would not continue in the future. (Foster 1995.)

Logically enough, the progressivists and holists also differ when it comes to their beliefs concerning the role of technology in the process of development. Norgaard (1994, 54) adopts the holist point of view, seeing the faith in the blessings of technological progress as an extrapolation of past good fortune experienced over the past two centuries in the northern industrialised countries, rather than as a result of a fresh analysis of what might lie ahead for the North or what are the actual conditions in the South. The technological progress required to solve the problems caused by previous technological advances would ultimately be constrained by our ability to develop prerequisite new forms of social organisation. In other words, new technologies arise faster than our organisational capacity to control their social and environmental side-effects increases. (ibid.)

Both holists and progressivists consider technological development as a dynamic part of a coevolutionary interaction between technical, socio-economic and ecological systems, in which policy action does have a crucial role (Norgaard 1994; Faucheux 1997). Yet, while progressivists call for an educational policy for the creation of what the mainstream economics has come to call 'human capital' (DeGregori and Shepherd 1994), Norgaard (1994, 54) again reminds of the limits of knowledge creation arguing that to be able to know and manipulate the environment our societies must constantly increase the level of education – a process that cannot continue indefinitely, because sooner or later the majority of citizens would be on one or another type of schooling, and no one would be doing the necessary 'real work'.

Norgaard (1994, 62-63) sees the roots of the problems of modernity to lie in the "dominant premises" of modernism itself, which are "unconsciously, implicitly, and eclectically invoked in arguments presented in both public and scientific occasions", as well as "in the arguments of academics, capitalists, environmentalists, and politicians with diverse and frequently opposing interests." He does not reject these premises (atomism, mechanism, universalism, objectivism, and monism) as such, but condemns their exclusive domination in public discourse and beliefs to the detriment of other metaphysical and epistemological premises which he deems "more appropriate for understanding the complexities of environmental systems and which are more supportive of cultural pluralism" (ibid., 63.)

Table 3.1. Metaphysical and epistemological premises underpinning modernity, and the proposed coevolutionary alternative

Dominant Premises	Alternate Premises
Atomism: Systems consist of unchanging parts and are simply the sum of their parts.	*Holism*: Parts cannot be understood apart from their wholes and wholes are different from the sum of their parts.
Mechanism: Relationships between parts are fixed, systems move smoothly from one equilibrium to another, and changes are reversible.	Systems might be mechanical, but they might also be deterministic yet not predictable or smooth because they are chaotic or simply very discontinuous. Systems can also be evolutionary.
Universalism: Diverse, complex phenomena are the result of underlying universal principles which are few in number and unchanging over time and space.	*Contextualism*: Phenomena are contingent upon a large number of factors particular to the time and place. Similar phenomena might well occur in different times and places due to widely different factors.
Objectivism: We can stand apart from what we are trying to understand.	*Subjectivism*: Systems cannot be understood apart from us and our activities, our values, and how we have known and hence acted upon systems in the past.
Monism: Our separate individual ways of understanding complex systems are merging into a coherent whole.	*Pluralism*: Complex systems can only be known through alternate patterns of thinking which are necessarily simplifications of reality. Different patterns are inherently incongruent.

(Norgaard 1994, 62).

What the debates on reflexive and ecological modernisation have shown is the irreducible ambiguity inherent in the concepts such as modernity and progress. These debates cannot be decided upon *a priori*, but the desirability of a certain type of progress, modernity, economic growth and technological development will need to be examined empirically, case-by-case. This does not amount simply to an often-made assertion that 'progress, sure has negative sides too, but on the whole, the positive ones dominate.' Average measures of the consequences of modernisation have little relevance if and when the curses and blessings of modernity fall differently upon various individuals and groups of society. Hence the question: sustainability for whom?

Unlike the outright anti-modernist thinkers, this study does not reject the ideas of progress and modernity altogether, but believes in the capacity of modernity to reform itself through reflexive learning. Indeed, human capacity to learn from experience can be considered as the most fundamental characteristic of humanity, and the ultimate foundation for any hopes of sustainable development or 'genuine progress' of humanity (Hahtola 1990). A central function of environmental policy evaluation would be to facilitate such learning to enable, first, an analysis of

the means and the ends of environmental policy – an analysis open in relation to the tenets of modernity – and, second, promote social learning through an informed dialogue on the pros and cons of modernity in the specific policy contexts.

3.3 Limits to growth?

A question intimately linked with that of progress and modernity is whether the biophysical and ecological life-support systems place ultimate limits to economic expansion and whether we should be concerned about such limits. The idea of limits to economic growth was present in the thought of the classical economists – in the ideas of 'dismal science' of Malthus and 'stationary state' of Smith and Ricardo, but was pushed away from the agenda of economics with the 'neoclassical revolution'. The question re-emerged with the recognition of global environmental problems in the late 1960s, with the imminent exhaustion of natural resources foreseen by the Club of Rome (Meadows et al. 1972), with the UN Conference on the Human Environment in Stockholm 1972, and finally, with the first oil crisis in 1973.

The predictions of the Club of Rome proved overly pessimistic, partly because of some of the simplistic assumptions, and partly because of the measures undertaken to curb resource (especially energy) use. In its original form, the 'growth debate' of the early 1970s (see e.g. Kapp 1950; Mishan 1967; 1977; Meadows et al. 1972) has lost a lot of its appeal, but the arguments have become more nuanced, and the contestation of the 'growth ideology' has reappeared in the writings of some 'ecological economists' (e.g. Daly 1992; 1996; Ayres 1998), within the 'anti-globalisation' movement, and among the development critics (e.g. Latouche 1989; Sachs 1992; Rist 2001; Harribey 1997; 1998). The advocates of the 'mainstream' growth ideas, in turn, have sought evidence notably in the 'Environmental Kuznets Curve' hypothesis to prove the inevitably positive impacts of economic growth on the environment (e.g. Stern 2003; Yandle et al. 2004).[34] The idea of 'limits' has, however, become a key element of the concept of sustainable development. The question is not so much one of the existence or not of such limits – simple common sense tells us that our planet has a limited capacity to support human activities – but rather about how to define the limits and how to deal with them. In particular, the issue is about the extent to which the 'limits'

---

[34] It is notable that the growth debate was not exclusively about the environmental impacts of economic growth, but also about its allegedly perverse social effects, an idea that the current 'anti-globalisation' movements have brought to the fore again. Hence, as early as already in 1976, Hirsch brought attention to the social limits to growth, arguing the limits to growth critique to be "strikingly misplaced" in its focus on "distant and uncertain physical limits" at the cost of

are fixed instead of flexible, and to what extent environmental assets and values can be substituted by technology. Likewise, the debate between the 'weak' and the 'strong' sustainability[35] has become one of defining the extent of substitutability: which part of the environmental values, 'natural capital', can be substituted by man-made capital?

From the 'holistic' institutionalist perspective adopted in this study, the essential lesson from the 'growth debate' is rather obvious: economic growth may or may not contribute to human well-being, depending on the context. For example, when the basic needs[36] of a large part of the population are not satisfied economic growth most certainly will improve well-being and help move towards sustainable development. Things get more complicated in an affluent society, where more growth may well mean that its negative consequences increase and predominate over the positive ones. If one brings on board issues of distribution and justice, the matter becomes more complex yet: for example, should we consider economic growth to contribute to well-being in a dictatorship, where a tiny elite reaps the lion's share of the benefits of growth, while the majority of the population does not get poorer, yet remains in misery? The institutionalist answer, hence, calls for attention to the context-specific issues, and the society's values, organisation, distribution, level of technological development, and the ways of measuring growth and well-being, instead of seeking universally valid answers.

The idea adopted here is that the existence of limits to growth must be taken seriously. As Ekins (1994, 295) has noted, further GDP growth will only *not increase* current rates of economically induced environmental destruction if change in composition of GDP, substitution of factor inputs, and increase in resource productivity act in sum to reduce the environmental intensity of GDP. This implies the necessity to maintain a stable resource base; avoid the over-exploitation of renewable resource systems or environmental sink functions; deplete non-renewable resources only to the extent that investment is made in adequate substitutes; and take into account ecosystem resilience, as well as the uncertainties caused by complexity, irreversibility, thresholds, synergistic and

---

the "immediate if less apocalyptic presence of social limits to growth" (Hirsch 1976, 4), deriving from phenomena such as increasing crowding and congestion on the one hand, and the breakdown of individual morality on the other.

[35] In essence, the weak sustainability criterion postulates that limits to growth do not exist, assuming that substitutes for non-renewable resources, for instance, would be found as a result of technological progress stimulated largely by the rising relative prices of resources, stemming from their increasing scarcity. The strong sustainability criterion, in turn, requires that the natural capital stock should not be allowed to decline, and therefore assumes that a part of the natural assets are critical for the functioning of the ecological life-support systems, and therefore cannot be substituted. (Faucheux et al. 1997; O'Connor 2000b; Hanley et al. 2001, 134-136.)

[36] While the definition of what should constitute basic needs is a subject of persistent controversy, even a minimalist definition, limited to the basic necessities of survival, will do: countries whose citizens suffer outright lack of food, shelter, and clean drinking water are numerous enough to make the point.

cumulative effects when making decisions affecting the environment (e.g. Holling 1973; O'Connor 1998b; Harris 2000). However, while there are good reasons to believe that biophysical limits of what could be considered 'sustainable' are being exceeded, these limits are not fixed but they tend to 'coevolve' together with the level of knowledge, technological development, the values and the way our societies are organised. These limits are far from objective in the sense that they always depend on the framework of interpretation. 'Pure' knowledge concerning our physical or socio-cultural environment cannot be obtained. (see e.g. Hodgson 1988.) In this substantive sense of the term, it has proven easier to define what *is not* rather than what *is* sustainable (Theys 2002b). Instead of being simply a matter of limits to what is possible, the question becomes that of our values and preferences: what kind of development do we desire and how do we go about deciding what kind of society we want to live in?

3.4 Dimensions of sustainable development and the question of scale

Even though the definition by the Brundtland Commission does not make such a distinction, sustainable development has later become perceived as a combination of three dimensions or 'pillars'[37], namely the environmental (ecological), economic, and social dimensions.[38] Since the Rio conference in 1992, this tripartite description has constituted the basis for most of the generally accepted definitions of sustainable development in international organisations (e.g. OECD 2001f; Commission of the… 2001), called 'triple bottom line' in the business circles. However, much less consensus reigns over the relations between the dimensions. The perspective endorsed by many institutional players is that of three, mutually interacting dimensions, each with an equal status. For instance Zoeteman (2002, 20) stresses the consensual character of sustainable development, defining it "as a dynamic concept, describing a situation in which economic growth and social development are not reducing environmental capital and in which decisions are taken by consensus, attributing equal value to the economic, social and environmental aspects".

While politically attractive, as it provides authorities and other actors operating in any of the three dimensions each with its own legitimacy and terrain to defend, the three-pillar conception of sustainable development can be criticised for at least three shortcomings. The political critique

---

[37] Other concepts such as factors, components, logics, or elements are sometimes employed to describe the same phenomenon.

[38] Other dimensions have been suggested as well. Institutional dimension has been among the most frequently proposed ones (O'Connor and Spangenberg 2004, 5; Spangenberg 2001), and Sachs (1993) adds to the concept the territorial, political and cultural dimensions.

evokes the risk that such an approach would likely reinforce the status quo, by legitimising the existing goals of the society, each government agency finding its own objectives corroborated in the concept of sustainable development. By assuming all dimensions to have equal weight in decision-making situations, the three-pillar approach tends to evade the crucial questions of what is to be sustained, why, and for whom (see e.g. O'Connor 2000c, 9). Likewise, the emphasis on the consensual character of sustainable development easily leads actors to avoid treating issues that might involve significant conflicts and divergence of views, thus resulting in an agreement around the smallest common denominator. Building consensus does not always genuinely reduce conflict, because the trust and co-operation that arise at the negotiating table neither necessarily translate to other settings nor diffuse throughout the organisations formally represented in a negotiation. Neither do deliberative processes, no matter how successful and well organised they might be, always lead to consensus. Moreover, as argued above, conflict and differences in power are inescapable elements of human interaction and should not, indeed cannot, always be avoided. Procedures that involve members of the public in a dialogue about policy, but that do not expect any consensus to be reached, can yield the benefits expected from a consensus approach. (Coglianese 1999.)

Another, more conceptual critique, based on the perception of symbolic power residing in language and discourse, claims that the three-pillar model perpetuates the 'economism' and 'productivism' characteristic of modern societies. By continuing to distinguish the 'social' from the 'economic', the three-pillar model contributes to strengthening the idea that the economy can be treated as a separate sphere, detached from the social context within which all human activities are embedded. This radical critique considers the three-pillar representation as a false consensus, which reflects fundamental flaws in the relations between human societies and their environment. (Passet 1996; Le Bot 2002.)

A third criticism focuses on the trade-offs and synergies between the dimensions. Each of the three 'pillars' has its own characteristics and logic, which are likely to conflict with each other. The model does not give any guidance on how to arbitrate between the unavoidably conflicting objectives of economic rationality (profitability), social justice and the stability and resilience of ecological systems. For instance, the objectives of improving material well-being and conserving natural ecosystems often conflict with each other. (Harribey 1998, 102-103; Upton 2002.) Moreover, as an all-encompassing concept, the three-pillar approach hardly gives guidance for

distinguishing what is essential for sustainability from what is secondary, and can lead to endless negotiations concerning the distribution of wealth at the global level (Upton 2002).

In the light of the preceding discussion on the environmental 'limits', we must ask whether the environmental 'pillar' should not be given primacy over the two other dimensions. Do biophysical conditions determine the ultimate, fixed limits for sustainable development or are the limits flexible? How much is there room for substitution between natural and man-made assets? Unsurprisingly, different actors tend to see the issue differently. For many, notably in the tradition of ecological economics, the primacy of the environment is a logical necessity, flowing directly from the idea of 'limits' (e.g. Daly 1992; 1996; Passet 1996). From the perspective emphasising the social construction of reality, by contrast, the goals of the society, in the end, always depend on the socially constructed values, norms and beliefs that ultimately define what is to be considered sustainable, and therefore any conception of sustainable development is ultimately determined by the 'social dimension' (e.g. Ballet et al. 2003). Finally, one can likewise postulate the primacy of the economics, given that environmental protection only comes at an economic cost, and is therefore conditioned by economic welfare. This view can also imply the idea that all we need in order to reach sustainability is to 'get the prices right', internalise environmental externalities, in the spirit of 'sustained economic growth'.[39]

From the perspective of institutional economics, the three dimensions can be considered as a useful basis for conceptualising sustainable development, but the relation between them should be reconsidered. One alternative is the so-called bioeconomy model, based on the works of Nicholas Georgescu-Roegen, and advocated by scholars in the school of ecological economics (Passet 1996; see also Maréchal 2000).[40] In this model, the three pillars have been replaced by three concentric circles, the environment circumscribing the social dimension, and the economic sphere constituting the innermost circle. This reflects the idea that economic activities should be in the service of all human beings while at the same time safeguarding the biophysical systems necessary for human existence. The social would thus be in command of the economic, but at the same time subsumed to the ultimate biophysical constraints, nevertheless taking into account what was said above about the

---

[39] In recent years sustainable development has increasingly become understood as sustained economic growth, thus relegating to a secondary role the issues of planetary limits, no matter how relative and changing these might be (Bartelmus 2003, 62). Symptomatic of this 'ethos' is the approach applied in the OECD Economic Surveys to treat sustainable development. In the first phase of the experiment, in 1999-2001, the sections dealing with environmental policy were named "enhancing environmentally sustainable growth".

changeable and socially constructed nature of the 'limits'[41] The clear hierarchy between the circles does not mean that the environment would necessarily always be the most important and relevant dimension. In any particular situation, at any particular point of time in history, social or economic aspects may be the most relevant and meaningful points of departure as long as the operation of socio-economic systems does not enter into conflict with the environmental framework conditions (Hahtola 1990; Norgaard 1994).

3.5 Institutionalist view of the social dimension

The treatment of the social-environmental interface is where the bioeconomy model inspired by institutionalism seems particularly useful. While the environmental dimension has traditionally been considered as the core of sustainable development, the awareness of the importance of the social dimension has increased over the past years – both within the policymaking community, and among academic scholars. However, it has likewise proven the most difficult dimension to operationalise and measure (e.g. OECD 2001c; 2001e) and the one that has received the least attention from the research community (Bürgenmeier 2005). Nevertheless, without the social dimension, there would hardly be a reason to talk about sustainable development – the traditional concepts of environmental economics alone would do (e.g. Theys 2002b).

There are a number of different economic approaches to analysing the social dimension of sustainable development. Traditional welfare economic approaches would lead one to identify a social welfare function, use a theory of optimal development, or look at social issues as externalities. These approaches share the problems of the modern mainstream economics in general, and the tendency to treat the social aspects as mere 'externalities' (see e.g. Bürgenmeier 2005). Approaches such as that of 'capabilities' (e.g. Sen 1985; 1999; 2000; Nussbaum 2000; Ballet et al. 2003; Robeyns 2005) or 'social capital'[42] go some way towards the premises of the present study.

---

[40] We adopt here a restricted definition of bioeconomics, and therefore exclude a number of approaches identified by Vivien (1998) as 'bioeconomic' (e.g. socio-biological and neo-liberal approaches, economics of renewable natural resources, and eco-energetic bioeconomics).

[41] To illustrate the difference between the bioeconomy model and the standard neoclassical environmental economics one could say that the latter attempts to integrate the environment to the economy, whereas the bioeconomy model integrates economy to the environment (Zaccaï 2002).

[42] The roots of the concept of social capital have been traced back to the works of Durkheim and Marx – even Aristotle (Carroll and Stanfield 2003, 397). Douglass North's (1990) writings on informal and formal institutions, Francis Fukuyama's (1995) work on the role of trust in economy, and Peter Evans' (1995) writings on the nature of state-society relations are other intellectual predecessors of the concept. The contemporary use of the term is, however, most often attributed to Bourdieu (1986), Coleman (1988) and Putnam (1993; 2000). In recent years, the World Bank has

The concepts of capabilities and social capital are helpful in calling attention to hitherto relatively neglected issues in economics, such as distributional equity, and enabling a perspective of 'the social' that goes beyond the mere equity in distribution of economic resources. They have brought attention to issues such as the need to look not only at social outcomes, but also the capacities of individuals to take themselves out of disadvantaged positions; importance of the stability of social systems in reducing vulnerability; learning as a key means of coping with change and adaptation; significance of power relations and participation; and the need to apply a 'social precautionary principle' (Robeyns 2000; Adger 2001; Ballet et al. 2003; Chiesura and de Groot 2003; Rammel and van den Bergh 2003). However, the concepts still seem somewhat poorly equipped to handle the procedural aspects of sustainable development, the collective processes in which meanings and identities are created, the factors shaping our consumption patterns, etc. The capability approach also has the major drawback of relying on a very relativistic notion of needs and poverty.[43] The social capital approach, in turn, runs the risk of treating the social and human relations as nothing but capital – a resource for production and consumption. Ultimately, this perspective sees the social dimension and social justice as externalities. This creates an illusion that the traditional economic framework could address sustainable development without having to pose questions about the ultimate objective of economy and economics, integrate other disciplines, or redefine the concept of social justice beyond the mere question of redistribution of revenues.

Experience from some of the empirical and theoretical studies attempting to systematise the environmental-social interface (e.g. Sachs 1999; Coenen et al. 2000; Theys 2002b; Hukkinen 2003) highlights aspects that are crucial for dealing with the social dimension. First, the traditional linear models of cause and effect relationships are poorly adapted to analysing the complex social phenomena, or their interactions with the environmental ones. What is contested here is not the existence of causality as an ontological phenomenon, but its value as an epistemological framework: causal relations may exist, but they can be intractable, often work in both directions, and may be even useless to trace. Secondly, while intuitively appealing, the idea of treating the environmental-social interface in abstraction from the economic one, and clearly separating the

been one of the most prominent advocates of the social capital approach (e.g. Grootaert 1998; Woolcock and Narayan 2000; Woolcock 2001).

[43] The emphasis on the freedom of each individual to lead a life that she considers worthy of living could imply that, in a wealthy society, owning a car, for instance, would be considered a basic, essential functioning. A major challenge is therefore to integrate into the capability approach a perspective that takes ecological sustainability seriously. (See e.g. "Greed Audit website, article "Sen and Sustainability: Poverty, Consumer Pressure, and the Future of the Planet" http://www.greenaudit.org/sen_and_sustainability.htm; consulted 5 May 2005).

social from the environmental is problematic. The 'social considerations', for example, may refer to poverty or unemployment, both of which in turn have a strong economic component. Conversely, no economy can function without social actors. In the field of the environment, the calculus of damages is omnipresent as well as the social perception of problems. (Zaccaï 2002, 103.) This speaks in favour of the 'whole development' approach adopted by Sachs (1999), and a conception that sees the three spheres of sustainability as embedded one in another.[44] Thirdly, the integration of the various dimensions of sustainable development works best at the local level – tied to the concrete living conditions of individuals in their daily environments.[45]

The social should not, therefore, be treated in abstraction from the other dimensions of sustainability, but should instead be seen as a pervasive aspect of all issues relating to sustainability. Institutions, meaning, cultural identity, procedural issues, participation and democracy are at the heart of the social 'dimension'. Equity should be seen in a context broader than that of mere distribution of economic costs and benefits, embracing the idea that aspects such as human dignity or cultural identity cannot be reduced to simple monetary calculations. (e.g. O'Connor 2000c; Bürgenmeier 2005)[46]

3.6 Interfaces and integration

In the nature of sustainable development is the integration of different scales of decision-making, between different disciplines and different sectors. Three issues are of concern here. First, the call for better integration of policies in different sectors of decision-making has become almost a political mantra, accepted by practically the entire policy community, including the EU and the OECD. The OECD has described policy integration[47] as one of the major challenges of sustainable development (e.g. OECD 2001e, 47-56; 2002a, 11-24). Integration of *environmental concerns* into

---

[44] By contrast, Daly (1996) criticises any attempt to extend the concept of sustainable development to areas not related to the environment, and argues that a distinct concept should be developed in order to take into account the 'ethicosocial' limits to growth.

[45] For example, Theys (2002b) sees local territories as the level at which the questions of socially sustainable development become concrete, where the interactions between the different dimensions are most explicit, where participation and dialogue, allocation of responsibilities, and analysis of interactions are the most feasible.

[46] A possible way of integrating the social with the other dimensions of sustainable development is provided by the 'tetrahedral organisation of knowledge' (O'Connor 2000c; O'Connor and Spangenberg 2004), in which the social is built into the model through the concepts of purpose, significance, and governance, instead of being a separate dimension in a three-pillar conception of sustainability.

[47] Underdal (1980, 162; quoted by Mickwitz and Kivimaa 2004, 4), defines a perfectly integrated policy as "one where all significant consequences of policy decisions are recognised as decision premises, where policy options are evaluated on the basis of their effects on some aggregate measure of utility, and where the different policy elements are consistent with each other".

all areas of policymaking[48] has in turn been adopted by many international institutions (e.g. WCED 1987, 313-314; United Nations 1994; Commission of the... 2001) as a central objective aimed at improving environmental policies and fostering sustainable development.

While the general thrust to improve policy coherence across sectors can hardly be called into question, integration might have its downside as well and certainly is not an answer to all problems of 'sustainability policies'. For instance, contrary to the mainstream argument, Hukkinen (1999) criticises the notion of integration arguing that the present problems stem from excess of policy integration in the context of large asymmetries of power. The ever closer cooperation between the actors responsible for promoting the short-term economic interests and environmental authorities supposedly working for long-term environmental sustainability leads to long-term considerations systematically being overridden by the short-term economic interests. In the absence of an institution whose sole responsibility would be to promote long-term sustainability, the short-term economic interests prevail in the end. A practical example would be a situation in which environmental authorities refrain from bringing into the policy arena a solution they consider the most appropriate from the perspective of sustainable development, because proposing such a measure would be 'unrealistic', given the likely opposition on the part of powerful economic interests. While the principle of integration is certainly valid and useful at a general level, the need and the desirability of integration must be seen in the light of the context, paying particular attention to relations of power that might distort genuine attempts at policy integration.

Second, rather than defining the individual dimensions as such, the real challenge sustainable development is the integration across its different dimensions. Although the multidimensional character of sustainable development is usually mentioned, in actual practice, the interactions between the dimensions, notably the trade-offs between alternative conflicting goals, have been poorly analysed. The interaction between the social and environmental dimensions is probably the least developed notably when it comes to measurement of sustainable development (OECD 2001f, 63; Lavoux 2003). Any conceptualisation of sustainable development should therefore help see the relations between different dimensions of sustainability, in particular those between the social and the environmental.

---

[48] Lafferty and Hovden (2002, 15) have defined environmental policy integration as "the incorporation of environmental objectives into all stages of policymaking in non-environmental policy sectors, with a specific recognition of this goal as a guiding principle for the planning and execution of policy; accompanied by an attempt to aggregate presumed environmental consequences into an overall evaluation of policy, and a commitment to minimise contradictions between environmental and sectoral policies by giving principled priority to the former over the latter."

The third issue of integration has to do with the relationships between the geographical and administrative scales. Seen in the perspective of planetary 'limits', sustainable development is by definition a global concept, ultimately concerned with the survival of humanity, if not of the whole planet. On the other hand, sustainability debate has right from the beginning been marked by concerns for local level implementation and citizen participation. The challenge is to build bridges between the various scales and levels – local and global, sectoral and spatial – situating local level phenomena into a broader context, and identifying the roles of actors and institutions at different levels. As pointed out by Theys (2002b), such 'bridge-building' is precisely what seems to be missing as the discourse on sustainable development tends to become polarised between two levels. On the one side are the geographers, planners, landowners, etc., mainly interested in integrating environmental concerns into local development, as well as into infrastructure and spatial planning. On the opposite side one finds the economists, large enterprises, consumer organisations, international NGOs and diplomats, more occupied with ethical consumption, precautionary principle, eco-taxes, emissions trading, WTO, etc. Theys (2002b, 8) argues that while the questions of equity are indeed addressed in the latter, more economic and global, 'non-territorial' approaches, they are posed exclusively in terms of revenue or access to development, overlooking the growing environmental inequities.

The key question for Theys (2002b, 11-12) is "our collective capacity to modernise the public action" so as to arbitrate between the present and the future, and between the different dimensions of sustainable development. His suggestion for a response to the challenge serves as the guide for our approach to the problems of integration: he argues that the local level should function as a 'motor' of sustainable development, and be in charge of integrated policies of sustainability promoting eco-efficiency, synergy between the dimensions of sustainability, etc. responding to the question 'how to live together'? The function of the state and the intergovernmental actors, in turn, would be to take care of the intolerable environmental or social risks ('how to survive together?'). This would consist of two tasks: firstly, developing a flexible framework of common principles to guide local level action (such as the precautionary principle), and secondly, the identification of manifestly unsustainable situations – be they environmentally, socially or economically unsustainable. This suggestion is similar to that of Hinterberger et al. (2000) who advocate two starting points for environmental policies in the pursuit of sustainable development. The first one consists of a search for a guiding principle to operationalise the 'heuristic rule' of sustainable development, i.e. an imperative to seek consensus between actors about certain values (e.g. risk

preferences and intergenerational justice), and the identification of system interdependencies. The second one would establish ecological 'guardrails' that define unsustainable paths of development.

3.7 The procedural side of sustainable development: post-normal science, procedural rationality, and deliberative institutions

While the question of 'limits' is crucial for sustainable development, the above discussion has a number of times alluded to the importance of the procedural aspects of sustainability, notably the ways to bring about a transition towards a more sustainable society. Sustainability is not only a question of survival, but it is intimately linked with the perceptions of an 'ideal society' (e.g. Hahtola 1990), the desired type of democracy, and the role of citizen participation in policymaking. Deliberative democracy was in the previous section described as the model of ideal society and an objective of institutional economics. Two further concepts – post-normal science and procedural rationality – can be brought to bear in support of the idea of creating deliberative institutions when dealing with issues of sustainability.

Today's environmental and sustainable development problems differ from traditional scientific problems in that they are global in scale, have long-term impacts, and require urgent decisions to be made on the basis of inadequate and uncertain knowledge. In the issues of science relating to the protection of health and the environment, typically "facts are uncertain, values in dispute, stakes high, and decisions urgent" (Funtowicz and Ravetz 2003, 1). This turns upside down the traditional distinction between 'hard', 'objective' facts, and 'soft', 'subjective' value-judgements, since hard policy decisions need to be taken where the only available scientific inputs are irremediably 'soft'. In such circumstances, 'truth' as the goal of science becomes a distraction, and a more relevant guiding principle is quality, "understood as a contextual property of scientific information". (ibid.) The need for post-normal science springs from the recognition that we are no longer, and never really were, the 'masters and possessors of Nature' as Descartes imagined, but that the present problems constitute complex systems, involving novelty and emergence, technology being embedded in its societal and natural contexts, and the nature itself being shaped by its interactions with humanity. There is no 'invisible hand' to lead the society towards better solutions, and

therefore no substitute for morality in the good conduct of our affairs. (Funtowicz and Ravetz 1990; 2003.)[49]

Under conditions of uncertainty and plurality of values, means-ends rationality is incapable of alone responding to the needs of decision-making. Since there is neither consensus over the ends to be pursued, nor certainty concerning the cause-effect relationships between the means and the ends, and since decision-making involves arbitration between often equally legitimate claims and interest of various groups of the society, the perceived legitimacy of the process becomes crucial. This calls attention to the 'procedural rationality', concerned with the legitimacy of the planning and decision making process as a whole. The 'procedural' hence refers "to the collective processes of action and decision-making, to the qualitative and social institutional dimensions that cannot be expressed in input-output terms" (O'Connor 1998b, 38). The process of decision-making may indeed be as important as the decision itself for public acceptance of policy decisions (Gowdy 2003).[50]

Since the definition of the limits imposed by the biophysical environment on the scale of the economy is an unavoidably social process, the irreducible complexity, uncertainty, and plurality of values call attention to the processes through which meanings are being attached to the notion of sustainability, and which would help resolve conflicts over the objectives relating to sustainability. Even if we do not have the knowledge and the capacity required to quantify the conditions needed for development to be sustainable, we can nevertheless define requirements for the processes conducive to sustainability. In the same way as for the aspects relating to sustainability as an 'end-state', the criteria for procedures favouring sustainability must rely on general principles, which must be operationalised separately for each and every case and context. However, the focus on the procedural should by no means imply the neglect of the substantive aspects. Even deliberative democracy – as a concept essentially based on procedural considerations – should have its

---

[49] Funtowicz and Ravetz (1990; 2003) describe their approach by reference to three problem-solving strategies. When both system uncertainties and decision stakes are low, the 'normal', applied science is relevant, and scientific expertise fully effective. In a case of medium level of uncertainty or moderate decision stakes, the application of routine techniques is not enough, but skill, judgement, creativity, and sometimes even courage are required. A typical case is professional consultancy, the work of a surgeon or a senior engineer as examples. Finally, when risks cannot be quantified, or when possible damage is irreversible, we are in the realm of 'post-normal science', out of the range of competence of traditional sorts of expertise and traditional problem-solving methodologies. It is to be noted that post-normal science is seen to apply even in cases of rather low system uncertainties, if the decision stakes are very high (e.g. when an institution is seriously threatened by a proposed policy).

[50] The distinctions between instrumental and value-rationality, or the logic of consequences as opposed to the logic of appropriateness (March and Olson 1998) are yet further examples of calls for a stronger emphasis on the decision-making procedures. The former follows Maw Weber's distinction between 'Zweckrational' and 'Wertrational' (often called 'substantive' rationality) action (Flyvbjerg 2001, 53-55).

substantive side through recognising principles such as basic liberty and fair opportunity, as suggested by Gutmann and Thompson (2003).

An integral element of post-normal science is the need to extend the range of participants in decision-making processes, create deliberative institutions[51] for conflict resolution and create 'spaces of negotiation'. Hence, not only persons with an institutional accreditation, but all those with a desire to participate in the resolution of the issue would be involved in the processes of negotiation. Such 'extended peer communities' not merely extend the liberty of individuals, but also the accountability of governments to include the institutions involved in the governance of science and technology. They likewise mobilise local knowledge and act as forums for collective formation of individual preferences. (Funtowicz and Ravetz 1990; 2003; O'Hara 1996; Funtowicz et al. 1999; Gowdy 2003, 5; Munda 2003.)

For instance, the quality of an evaluation framework would not be determined merely by its scientific accuracy – there is usually an inevitable trade-off between accuracy and relevance – but also by its capacity of enhancing dialogue in the spirit of deliberative democracy. O'Connor and Spash 1999, 33) have suggested that the validation of a valuation method depends not only on standards of rigour and internal coherence, but also on external considerations, including its social adequacy. Quantitative measurement of the degree of sustainability as a state is certainly helpful, but only within a framework of an evaluation process aimed at facilitating the process of collective preference formation, conflict resolution and decision-making. This in turn requires a good understanding of the positions of various participants along the dimensions of power.

3.8 Measurement of sustainability

Where the links between environmental policy evaluation and sustainability become concrete is in the attempts to operationalise the term through measurement of sustainable development. A common set of principles, in line with our broad institutionalist perspective, should underlie the approaches for evaluation of environmental policy and measurement of sustainability.

---

[51] Concrete examples of deliberative institutions include citizen juries, focus groups, and consensus conferences. Regardless of the name given to such institutions, they have one important element in common: they assess the quality of proposals, including their scientific element. (Funtowicz and Ravetz 2003, 7.)

There is a relatively broad consensus among economists about the insufficiency of GDP as a measure of human well-being (e.g. Ekins and Jakobs 1995; Hanley et al. 2001, 122; OECD 2001e, 19-21).[52] The modern mainstream answer to the question of 'limits' has been to define non-declining utility and consumption per capita as the key criteria for a sustainable economy (e.g. Hanley et al. 2001, 134-136), and apply the so-called 'capitals approach', which analyses sustainability from the perspective of changes in the stocks of different forms of capital (e.g. Hartwick 1977; Solow 1986; Pearce et al. 1989; Victor 1991; Ekins and Simon 1999; Faucheux 2000b; Zaccaï 2002, 241-248).[53] Despite its origin in economics, the 'capitals approach' has been accepted as a 'common sense approach' by the academic community more generally (e.g. Farrell and Hart 1998; Harris 2000, 5-6; Spangenberg 2001; Ballet et al. 2003).

Numerous systems of indicators, at different geographical scales and levels of decision-making, for different purposes, have been proposed as a means of operationalising sustainable development and better capturing the different components affecting human well-being. Modern mainstream economics' indicators of sustainability, relying on the concept of 'weak sustainability', include 'net savings' taking natural capital depreciation into account, or 'environmentally corrected net national product' ('green NNP'). Environmental Space, Ecological Footprint, Total Material Requirement, Material Input Per Service Unit (MIPS), Human Development Index, and Index of Sustainable Economic Welfare are but a few of the other proposed single indicators (for reviews and experiences see e.g. O'Connor et al. 2001; OECD 2001e; 2001h; 2004e). Yet others have striven to define 'safe minimum standards', so as to deal with uncertainty and irreversibilities (Ciriacy-Wantrup 1952; Bishop 1978), or have defined rules aimed at limiting the scale of human activities to forms and levels that are within the long-term carrying capacity of ecosystems at relevant scales (Daly 1996). A less ambitious and more common approach is to put together sets of 'sustainability indicators' covering different dimensions of sustainability (for a review, see e.g. OECD 2002d). An example is to rely on 'key environmental functions' to define 'critical natural capital' (see e.g. de Groot 1992; Faucheux 2000b; Chiesura and de Groot 2003.)

The 'weak sustainability indicators' have been criticised extensively, and do not seem to provide a solid basis for sustainable development, notably because of the existence of irreversibilities,

---

[52] Ekins and Jacobs (1995) also contest the ability of GDP to measure production, since it excludes a great majority of non-monetary economic production and treats 'defensive expenditure', for instance to mitigate the damage caused to the environment, as any other economic production.

[53] The 'capitals approach' often identifies four types of capital: manufactured, natural, social, and human, the latter sometimes being included in social capital.

thresholds, and cumulative and synergistic effects, which imply that some of the natural capital is non-substitutable (see e.g. Daly 1996; Faucheux et al. 1997; Victor et al. 1998; Gowdy and McDaniel 1999; Faucheux 2000b). Likewise, it has been argued that these indicators are deficient since they only cover a very small number of categories of marketed natural resources and environmental degradation, and use, as basis of calculation, market prices that are probably systematically 'wrong' from the perspective of inter-temporal opportunity costs (Faucheux et al. 1997).

A number of problems plague the modern mainstream economics approaches to measuring sustainable development. Monetary valuation faces insurmountable problems stemming from the complexity of environmental problems, manifested notably by the irreducible uncertainties, the high decision stakes and long risk/impact time-horizons, impossibility of interpersonal utility comparisons (Bürgenmeier 1999), ethical convictions and the problems of distributional justice (O'Connor 2000b; Söderbaum 2000), as well as more technical difficulties associated with the measurement techniques (Bartelmus 2003, 72). Changing wealth distribution inevitably changes cost and demand functions, and thereby affects what is defined as efficient. Market prices are not unique, 'natural' measures of worth or value, but are instead the result of public choices over property rights. Questions of distribution and allocation are intimately linked with each other, which makes it impossible to focus exclusively on a 'positive' analysis of the latter.[54] The standard economic reference to Pareto efficiency as the leading principle is not applicable when deep disagreements prevail concerning the goals of the society. One can argue that many, if not most, environmental policy problems, have to do more with (re)distribution of costs and benefits than with efficiency gains. Notions of justice, honour, or fair play must be brought to bear in order to decide in situations of several contrasting Pareto efficient policy options. Uncertainty therefore concerns not only ecological processes, but also social norms and their evolution. (e.g. Swaney 1987; Bromley 1991; O'Connor and Muir 1995; Munda 1997; Martinez-Alier et al. 1998; O'Connor and Spash 1999; Turner 1999; Faucheux 2000b; Söderbaum 2000; Goodwin 2003.)

Internalisation of external costs, or *'getting the prices right',* i.e. valuing all goods and services at their correct prices, including their environmental and social costs is likely to help moving towards sustainable development, but it is nevertheless based on the idea of an efficient use of resources, with all the associated difficulties of defining efficiency. Rather than being conceived in terms of an

---

[54] The issue in intergenerational welfare comparison concerns the "choice of what generation owns the resources and the implementation of trustees to establish reservation prices, perhaps infinite in some cases" (Schmid 1994, 107).

ideal of Pareto efficiency in resource allocation, internalisation should be seen in a broad sense, referring to political processes and institutions for expressing and resolving or accepting conflicts over environmental concerns (O'Connor and Spash 1999; see also Harris 2000, 19-20).

Three issues stand out as crucial for the design of indicators. First, indicators must address both the substantive and the procedural side of sustainable development as defined above. Second, they need to strike a balance between the decision-makers' desire to have aggregated information upon which to build their decisions on the one hand, and the need to take into account the context-specificity and respect the richness and diversity of local-level information on the other. Third, an appropriate indicator system must be able to build links between different levels, scales and dimensions.

As for the procedural question, the coevolutionary perspective reminds us that policies should be evaluated in a dynamic framework that looks not only at the 'results' of policy instruments, but also at their broader consequences, in particular their impacts on the diversity and freedom of choice that present decisions leave for the future. Stability of social systems thus becomes an important evaluation criterion. (Rammel and van den Bergh 2003.). This is an additional reason for developing process indicators, which have, however, received little attention so far (Zoeteman 2002; Hukkinen 2003).

The issue of aggregation versus context-specificity is crucial for institutional economics. On the one hand, aggregation is often defended on the grounds that it enables decision-makers as well as the general public to have an easily accessible measure of sustainable development. Decision-makers are thought to need or want concise information, because they hardly have time for extensive information search. The opponents of aggregation, by contrast, underline that one cannot meaningfully press impacts occurring on different dimensions of sustainable development into a one-dimensional unit of measurement.[55] Taking the context seriously also implies that the ultimate objective of any indicator (set) is its capacity to hold decision makers accountable, generate debate, enhance understanding, and provide legitimacy for measures leading towards sustainable

---

[55] Hukkinen (2001), for instance, has criticised the concept of eco-efficiency for relying on an erroneous notion of universal applicability, and for running against the cognitive and institutional bases of sustainable human-nature interaction by decoupling collective local organisation from local ecosystem management. Instead of being promoted as a universally valid policy prescription, eco-efficiency and its indicators must be applied in the right context and at the appropriate scale. They should ensure that individual actors perceive their everyday activities to be materially grounded in ecosystem services, and help build local level governance systems for ecosystem management. Moreover, there is in fact evidence showing that stakeholders do not actually favour a single indicator, because for real learning, understanding and engagement, there is a need to work with a well-balanced range of indicators. However, indicators should not be too numerous either, because this would reduce transparency. (Faucheux et al. 2002a.)

development (e.g. Schmid 1994, 108). Furthermore, the assumption that all members of society should accept the 'optimal' solutions calculated by experts has proven illusory, as well as the belief in non-experts' capacity to understand the complicated language of most aggregative methods.

Our institutionalist perspective relies on the idea of disaggregative decision-making methods,[56] based on the principle of 'weak comparability' of values (Martinez-Alier et al. 1998), and on the premise that to solve complex problems involving social conflicts, impacts should be assessed and presented each in their characteristic form.

This approach does not reject the idea of attributing economic value to the environment, but contests the reductionism involved in the belief that *all types of values* could be captured through a single measure. The aim is not to seek the 'optimal' solution, but to allow better informed choices on the basis of clear and well-argued description in such a way that all parties to the planning and decision-making process can form their own opinion and settle down to the job at hand. Furthermore, the objective is to facilitate communication and foster understanding across different value and belief systems or 'legitimacy orders'. If externalities are seen not so much as 'market failures', but as 'cost-shifting successes', conflicts may in fact help sustainability, and the movement for environmental justice can be seen as an asset for sustainability. (e.g. Funtowicz and Ravetz 1994; Leskinen 1994, 37-42; O'Neill 1996; 1997; Martinez-Alier et al. 1998; Söderbaum 1999; 2000; 2001; O'Connor 2000a.)

When it comes to the *links between levels, scales and dimensions*, the essential question is that of relevance. O'Connor and Spangenberg (2004, 3) note that "the question of the 'generic versus site specific' character of indicators has to be resolved with reference to the several different purposes of evaluation, their distinctive communication contexts and their respective organisational scales." The pertinence of the indicators would be evaluated in relation to their scientific quality and the policymakers' needs. The balance between the 'generic' and other (e.g. local) legitimate stakeholder interests must be sought through deliberation, in an attempt to attain 'representative diversity' across production sites, stakeholder types, and performance issues. (ibid.)

---

[56] Such as participatory multi-criteria decision aid (e.g. Funtowicz and Ravetz 1994; Martinez-Alier et al. 1998; O'Connor 1998a; 2000a; Froger and Oberti 2002) situational analysis (e.g. Hahtola 1990; Leskinen 1994; Söderbaum 1999; 2000; 2001), or participatory integrated assessment (Hisschemöller et al. 2001; Rotmans 2001; Rotmans and van Asselt 2001).

'Genuine' or 'true' indicators are unlikely ever to be found, because the nature of sustainable development as a contested, open, and multidimensional process with a host of internal contradictions and trade-offs implies that any analytical framework is bound to represent only a temporary agreement, which evolves along with our understanding of sustainability (Simon 2003, 6). Yet, the search for clear definitions and ways of measuring sustainability is by no means useless or irrelevant; on the contrary, such a search constitutes an essential basis for deliberative democracy. It helps, for instance, avoid false advertising, stemming from the tendency to equate sustainability with simply 'a bit more environmentally responsible', but it also encourages different parties to take beneficial actions (Graedel and Klee 2002).

3.9 Summing up: the key criteria of sustainable development as seen by institutionalism

We can now summarise the key characteristics of the 'institutionalist' perspective to sustainable development, which will later on be used for analysing the evaluation approach of the EPRs on the one hand, and the policy message, the recommendations the OECD gives to its member countries in the EPRs on the other.

1. Enhance *critical reflexivity* by not taking premises such as progress, modernity, economic growth and technological development as automatically desirable from the point of view of sustainable development and human well-being, but evaluating their value in the context.

2. Take into account the *limits to the scale of human activity* posed by the biophysical life-support mechanisms, and identify evidently unsustainable situations should be identified for instance through concepts such as critical natural capital, ecosystem functions, resilience, as well as irreversible, synergistic and cumulative effects.

3. See sustainable development not only as an end-state to be achieved, but essentially as a *coevolutionary process* of *institutional* transformation. Pay attention to both formal and informal institutions and their multiple functions, as well as to the pervasive role of power and conflict.

4. Facilitate the understanding of the *links* – synergies and trade-offs – *between the different dimensions* of sustainable development, and at different scales, as well as help distinguish essential from the non-essential; take into account the qualitative differences between the dimensions, therefore conceptualising the *'social'* through notions such as values, meaning, identity, and culture, rather than merely through distributional equity.

5. Recognise *complexity*, *multiple rationalities* and explanatory frameworks in the spirit of *methodological pluralism*, instead of relying exclusively on a single explanatory and methodological framework grounded in the ideas of, for example, efficiency or maximisation of individual utility. Base policy recommendations on a careful examination of the context, instead of universally valid 'axioms'.

6. Ground *measurement of sustainable development* in the idea of incommensurability of values, instead of seeking a single measure of performance or well-being. Pay equal attention to qualitative and quantitative elements. Seek appropriate level and degree of aggregation. Adopt a broad conception of internalisation, referring to the entire political process, instead of advocating internalisation simply through a strict Pareto efficiency framework.

Chapter 4

Evaluation approach and influence

To find out how evaluation could enhance sustainable development, it is not enough to know what is meant with the term 'sustainable development' – the final objective of evaluation – as defined in the previous chapter, but one also needs to be aware of the mechanisms through which policy evaluation could help work towards such an ultimate goal. This 'procedural' question is the subject of the first part of this chapter: how should environmental policies be evaluated in order to enhance institutional change towards sustainable development? To answer this question, we look for inspiration in evaluation research. Evaluation is first defined in relation to its sister disciplines, notably performance measurement and *ex ante* assessments, and international peer review is located within the field of evaluations. The second section of the chapter then looks at the experience from the main evaluation approaches that have emerged during the short history of evaluation research in order to elaborate general criteria for an evaluation approach that would be in line with the basic premises of the present study.

Most people involved in evaluation would probably consider that the value of an evaluation is ultimately measured by the degree to which it is utilised[57] in policy-making. Even though evaluators no longer share the great optimism of the profession's pioneers that tended to see evaluation and social science in general as capable of 'bringing rationality' to decision-making and enabling planned, controlled social change, the promise of utilisation is still undoubtedly the main reason why decision-makers continue to commission evaluation studies. If evaluations were not expected to bring about any effects, why should governments, organisations and foundations support evaluation studies in the first place?[58] The third section of the chapter draws upon theories and

---

[57] The concepts 'utilisation' and 'use' are here applied as synonyms.

[58] However, the views on the importance of utilisation differ even within the evaluation research community. While Patton (1997a) considers evaluation worthless unless it is utilised, Chelimsky (1997a; 1997b) warns against holding utilisation as an absolute quality criteria for evaluation. For instance, 'accountability' evaluations can be carried out for their own sake, as a means of administrative guidance and control, without concern for other potential uses of the

experiences concerning evaluation use and evaluation influence, so as to develop a typology that will serve as a basis for analysing the influence of the OECD reviews in the second part of the present study. After a brief survey of the various types of evaluation use identified in the literature, and the relations between the purposes and uses of evaluation, arguments are given to justify the adoption of 'evaluation influence' instead of 'evaluation use' as the unifying concept for the analysis. Learning, dialogue, argumentation, and 'process use' are described as the key mechanisms through which evaluations exert their influence. The perspective is finally extended to the influence of evaluation on democracy.

Since developing a typology of the various kinds of evaluation use and the ways in which evaluations exert their influence does not in itself suffice to explain why this influence varies from one situation to another, the fourth section of the chapter draws on the literatures on evaluation, political science and international relations in order to identify the factors that have the potential to condition the influence of evaluations on policy. Among the key factors figure the role of an international organisation as evaluator; the involved actors with their expectations, interpretative frames and policy networks; the general policy context; and the evaluation design.

---

knowledge. The same goes for evaluations aimed at knowledge-creation: even though one objective may be to arouse debate, open new perspectives and provide new concepts, their 'raison d'être' does not lie solely in their use. Chelimsky's ideas have, in turn, been contested by Helmut Egger, who considers evaluation to have only one purpose – "learning from experience to improve future work". While not disagreeing with this general purpose of evaluation, Chelimsky argues that more differentiated typologies are needed to account for differential rates of use of evaluations with different purposes. (Eggers and Chelimsky 1999.)

4.1 Definitions of evaluation research

Divergences of view characterise the field of evaluation in general – already when it comes to defining what is understood by the term 'evaluation'. The disagreements concern the questions of whether evaluations should focus on assessing the extent to which policies reach their objectives or instead adopt a broader perspective, whether evaluation includes both *ex ante* and *ex post* analysis of policies; whether evaluation requires the use of social science methodologies or whether any analytical approach would do.[59] Common to all definitions are the stress on *valuation*, *systematic data collection and analysis*, and the aim of *informing future decisions* concerning the issue or phenomenon being evaluated. In this sense, evaluation comes close to various types of assessments, policy analysis, performance measurement, and even planning. Rather than attempting to give an exhaustive definition, it may be easier to define evaluation by describing what it is not, and explaining how it differs from these 'sister' disciplines.

4.1.1 The rise in popularity of evaluation research and performance measurement

The origins of evaluation research can be traced back to the 1960s and the attempts to appraise the ambitious social programmes of the 'great society' in the U.S. (Johnson 1964). The exponentially increasing costs of social welfare gave rise to the need to evaluate whether the money invested was used in an efficient manner. Later on, same kind of an evolution took place in other parts of the world, with a number of regional (e.g. European), and national evaluation societies being created in the 1990s. The field of application of evaluations expanded from social, health and educational policies to practically all sectors of public policy, and evaluations have become an accepted activity

---

[59] Different definitions of evaluation include: "the systematic application of social research procedures in assessing the conceptualisation and design, implementation, and utility of social intervention programs" (Rossi and Freeman 1988, 19); "a systematic inquiry into the worth or merit of an object" (Sanders 1994); "policy evaluation research is the objective, systematic, empirical examination of the effects ongoing policies and public programmes have on their targets in terms of the goals they are meant to achieve" (Nachmias 1979, 4); "analytical inquiry based on collecting and analysing evidence, and drawing conclusions and recommendations from this evidence" (Valovirta 2002, 60); "careful *post facto* assessment of the implementation, performance and outcomes of a public policy, which is aimed at playing a role in further decision-making and action" (Vedung (1991, 33); "determining the value or the utility of an issue or a phenomenon, based on information obtained thereupon" (Laukkanen 1998, 32); and "programme evaluations are systematic, analytical assessments, addressing important aspects of a programme and its value, and seeking reliability and usability of findings" (OECD 1999, 13). Patton (1997a, 23) distinguishes between 'programme evaluation' and 'evaluation research', the former referring to "the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/or inform decisions about future programming", whereas the latter stresses the need to apply social science methods, not only systematic data collection methods typical of practical programme evaluation. For Weiss (1999, 469) the "overall *aim of evaluation* is to assist people and organisations to improve their plans, policies and practices on behalf of citizens", while House (1993, ix) understands by evaluation the determination of the value of a programme or policy according to chosen, explicit criteria.

in many departments, states, local authorities and NGOs through methods such as self-evaluation, self-assessment, benchmarking and total quality measurement (Leeuw et al. 1999; Hansson 2000; Pitarelli and Monnier 2000).[60]

The reasons for the increasing use of evaluations can be sought in particular in the changes in public policy environment over the past decades. These include the trends towards decentralisation and 'internationalisation' of public administration; citizens' demands for greater transparency of policymaking; the increasing influence of experts in policy and administration; the declining legitimacy of public administrations, entailing the public's loss of faith in the government's ability to spend their taxes wisely; and the pressures to reduce public spending. (e.g. Greene 1999, 162; Uusikylä 1999; Thoenig 2000; van der Knaap 2000; Pitarelli and Monnier 2000.) In short, there is an increasing pressure on the public sector and non-profit organisations to demonstrate accountability to external sponsors and other stakeholders (Feller 2002, 438).

Pawson and Tilley (1997, xii-xiii) have noted that two directly opposing views can be traced behind the rise of the evaluation movement. On the one hand, from the traditional perspective that emphasises the objectivity and independence of science in relation to politics, where evaluation "purports to offer the universal means with which to measure 'worth' and 'value'." On the other hand, there are the post-modern ideas of the end of certainty that preach the dawning of the age of "mass cynicism and cultural relativism", in which human understanding is perceived as being inevitably locked into the particular discourse surrounding each social group. In this atmosphere, the evaluator should abandon the search for truth and certainty, and instead celebrate the "free, instinctive play of imagination within decision making" (ibid.). We therefore find in the field of evaluation the same juxtapositions presented above, in the previous chapters – those between objectivity and subjectivity; truth-seeking and truth-creation; quantitative and qualitative methods; expert evaluation and participatory approaches; realism and constructivism.

Lately, the 'objectivity' view seems to have gained ground along with the increasing popularity of the New Public Management (NPM) in the public sector with its focus on evaluation and auditing.

---

[60] The OECD created in 1990 the OECD Public Management Committee (PUMA), later renamed as Public Governance Committee, with the mission to "identify and help address the emerging forces and trends which constitute strategic governance challenges; to assist Members and non Members to raise the performance of their public institutions so that they are better equipped to manage those forces; and to focus on key elements of good governance framework" (OECD 2000b). The Committee is supported at the OECD secretariat by the Public Governance and Territorial Development Directorate, which "identifies changing societal and market needs, and helps countries adapt their governmental

It has been argued that NPM reforms have been put in place so as to replace traditional but declining sources of legitimacy, such as community and state (Power 1997, 146). These reforms, increasingly applied since the 1980s, involve trends towards more indirect forms of government steering, such as information-based and economic instruments. The reasons for the relative rise of popularity of performance measurement to the detriment of evaluation can be seen to stem largely from the disuse of evaluations (Davies 1999). Evaluations are often not used, at least not in the way evaluators would like to, because they are considered costly and time-consuming; reports are too long and complex for the majority of potential readers; the often dominant 'formative' (instead of summative)[61] focus of evaluations means that issues of programme relevance are seldom addressed; the information provided by evaluations often lacks timeliness and is seldom considered useful for resource allocation or budgeting decisions. The accounting and auditing profession has spread its influence to an ever broader range of areas, including performance auditing and environmental auditing. (ibid., 153-154.)

4.1.2 Performance measurement and auditing

Auditing has a long and established tradition of serving the parliamentary oversight function in many countries. It focuses on legal compliance and accountability of public organisations, seeking evidence that established rules, regulations, procedures or mandates have been followed. By contrast, evaluation is typically less institutionalised, more participatory, and less independent than audit in relation to stakeholders. Evaluation is also used more as a tool for programme management than for public accountability. (OECD 1999, 16.)

Performance measurement can be defined as a continuous monitoring process, carried out throughout the execution of the programme, with the intention of immediately correcting any deviation from operational objectives, typically addressing the 'what' question, i.e. what have been the results of policies. Evaluation, by contrast, consists of an in-depth study and is specifically conducted at a discrete point in time in the life cycle of a programme, asking the 'why' and 'how' questions (European Commission 1997; Bukkems and de Groot 2002, 152).[62] The purpose of

---

systems and territorial polices. This involves improving government efficiency while protecting and promoting society's longer-term governance values." (http://www.oecd.org/about/0,2337,en_2649_33735_1_1_1_1_1,00.html)

[61] Summative evaluation has as its main objective to judge the worth of a programme or a policy, whereas a formative evaluation aims to promote understanding of the policy processes and mechanisms, thus supposedly leading to policy improvements.

[62] Performance measurement is therefore close to the so-called *'black-box evaluations'*, which have no intention to examine the processes behind the outcomes, i.e. the policy processes and policy implementation. A *white-box*

performance measurement is to establish conformity with given standards of the process and outcomes of financial management (in financial audits) or the evaluation of government outputs and subsequent outcomes for society against given targets (in performance or value-for-money auditing), ideally set by the government itself. Targets are then compared with actual performance. (Bukkems and de Groot 2002.) Performance measurement can thus be seen as a combination of auditing, which has the task of controlling the compliance with the rules by the audited system, and monitoring, which applies indicators to provide information on the progress of policy implementation.

However, the boundaries between auditing and evaluation have begun to blur as the scope of audits has extended to questions or performance, and traditional financial audits have been supplemented with performance audits (OECD 1999, 16). The concepts of performance measurement and performance auditing are, indeed, often used interchangeably. Both can be seen as key elements of *performance management*, described as "the process by which we collect performance information and use it for improvement, by drawing on a range of evaluative processes such as performance measurement and evaluation" (Davies 1999, 150).

In simplified terms, performance measurement (auditing) responds to the expectation of 'truth seeking', the role of the auditor being to reveal the truth by means of objective methods and standards of measurement. Evaluations, in turn, not only question the existence of one single reality, but are also based on the assumption that a sharp distinction between fact and judgement is impossible, since the so-called facts are not random and it is virtually impossible to present the facts without certain prejudices and assessments creeping in. (Gundvaldsen and Karlsen 1999, 464.)

Performance measurement has promised to improve the quality of public policymaking mainly through the following ways (Greene 1999, 162; Martin and Sanderson 1999, 247; Feller, 2002: 438):

- *Free from the preoccupation with complying with rigid regulations*, agencies and implementing authorities can redirect their energies away from following rules and toward creatively solving the problems in the field.

---

evaluation', in turn, places a major emphasis on explaining and understanding the policy processes that have lead to the observed outcomes. This latter type of evaluations are sometimes called implementation studies, given their attempt to identify problems in the policy implementation process. (Leeuw 1995; Kazi and Spurling 2000; Lang 2001, 14.)

- A focus on outcomes provides an opportunity for *collective, shared deliberation* about what constitutes valued outcomes from a given endeavour.
- Emphasis on outcomes can illuminate *whether investments are adequate to achieve expected results*, and thus reveal situations in which implementation personnel are asked to accomplish massive tasks with inadequate tools and resources.
- Performance measurement provides a readily comprehensible way in which an organisation can *communicate its goals and achievements* to external publics.

This is what performance measurement could ideally achieve. However, as we will see further down, performance measurement has its downsides, which often prevent these objectives from being reached, or give rise to other, undesirable side effects.

4.1.3 Ex ante evaluations

As was seen above, some definitions of evaluation include not only *ex post* evaluations, but also *ex ante* assessments such as environmental impact assessment, social impact assessment, technology assessment, strategic environmental assessment, and integrated assessment. In the present study, we shall focus exclusively on *ex post* evaluation. Yet, it is interesting to note that the same debates that have divided the evaluation community and social science scholars more generally, can be identified in the evolution of *ex ante* assessment as well. Environmental Impact Assessment (EIA) originated as a critique towards cost-benefit analysis, and was therefore largely seen as a method of assessment, but has since then become increasingly understood as a socio-political negotiating process aimed at resolving conflicts (e.g. Leskinen 1994, 42; Saarikoski 2000). A simple distinction can be made between perceptions of EIA as an expert-led process and as a forum for rational dialogue. These differences are related to the concepts of rationality underlying the assessments – notably the instrumental and communicative rationality – with the concomitant differences in the relative importance given to assessments as analytic techniques and as a forum of democratic dialogue, respectively (e.g. Hilding-Rydevik 2001; Sager 2001). The field of Integrated Assessment (IA)[63] has for long been divided between on the one hand the 'analysts', or engineers, having a practical 'natural science' approach, developing and applying integrated tools such as models, indicators or scenarios, and on the other, 'social scientists', who stress the character of IA as a

---

[63] Parson (1995) has defined the two distinguishing characteristics of integrated assessment as follows: "(a) that it seeks to provide information of use to some significant decision-maker rather than merely advancing understanding for its own sake; and (b) that it brings together a broader set of areas, methods, styles study, or degrees of certainty, than would typically characterize a study of the same issue within the bounds of a single research discipline."

participatory process (e.g. Parson 1995; Hisschemöller et al. 2001; Rotmans 2001; van Asselt and Rotmans 2001). The need to build bridges and seek complementarity between the two orientations has been recognised,[64] in concepts such as *participatory multi-criteria decision aid*, which tries to combine 'the best of two worlds' – the rigour of quantitative measurement and the capacity to enhance deliberation through active stakeholder participation, and can be seen as an attempt to apply to the practice the principles of procedural rationality and post-normal science evoked in the previous chapter (e.g. Hahtola 1990; Funtowicz and Ravetz 1994; Martinez-Alier et al. 1998; O'Connor 1998a; 2000a; Söderbaum 1999, 2000, 2001; Froger and Oberti 2002.).[65]

Evaluation, performance measurement, and *ex ante* assessments share largely the same fundamental aspirations of improving the rationality of policy making in the context of increasing complexity and plurality of values. The experience from *ex ante* assessments has brought up both the need of and challenges associated with involving stakeholders in the assessment processes – an issue we shall come back to in section 4.4.4.7 further down. More fundamentally, the experience has highlighted the persisting dichotomies and contradictions between the two general approaches, one emphasising scientific detachment, objectivity, rationality, and independent assessment of results, and the other stressing 'understanding', adopting a hermeneutic stand focusing on inclusive, participatory processes in the spirit of 'procedural' rationality. Yet another attempt to reconcile these two strands is international peer review.

4.1.4 International peer reviews: evaluations or performance measurement?

Before drawing direct lessons from evaluation research, performance management, and *ex ante* assessments, one must ask how closely international peer reviews of environmental policies resemble these other types of evaluation, and to what extent they constitute an activity of their own, not comparable with other assessments and evaluations.

Peer review – systematic evaluation by colleagues – has a long-standing tradition in many traditional professional fields such as medicine and science policy. In a peer review process,

---

[64] The same need for reconciliation has been recognised in foresight and futures studies, where the trend seems to lead towards the integration of and recognition of the complementarity between the two approaches (see e.g. Faucheux et al. 2002b). Faucheux et al. (2002b, 14) distinguish between the 'technological' and the 'social' foresight, the former referring to the prospects for technological innovation and the associated constraints and opportunities, whereas the latter attempts to establish a "social contract" between science and technology on the one hand, and society's needs on the other.

[65] For the application of multi-criteria decision aids to evaluation, see Floc'hlay and Plottu (1998).

members of a profession evaluate the activities of other members of the same profession, criteria of judgement being drawn directly from the profession's own quality criteria (Vedung 1991, 64-66). While peer reviews in their traditional fields of application have come under threat by the increasing use of performance measurement (e.g. Hansson 2000), in international organisations the idea of peer reviews seems to be gaining ground, albeit in a different form and on a different level.

The concept of peer review by international organisations has not yet acquired a rigorous definition accepted by all. However, over the years, the expression has assumed a specific meaning in the OECD context, and a tentative definition has been provided as "the systematic examination and assessment of the performance of a State by other States, with the ultimate goal of helping the reviewed State improve its policy making, adopt best practices, and comply with established standards and principles" (OECD 2003j, 9). The examination should be *non-adversarial* and rely on mutual *trust* among the States involved, as well as on their shared *confidence* in the process. With these elements in place, peer review is assumed to create a system of mutual *accountability*. (ibid.)

Peer review differs from performance measurement by the fact that the final outcome is not a binding or legal judgement by a superior body, and it never implies a punitive decision or sanctions. Moreover, peer review is characterised by dialogue and interactive investigation, and usually does not involve formal reporting by the examined state. (ibid., 5.) Compared with evaluations, in turn, peer reviews pay less attention to the causal mechanisms behind the policy outcomes – the main emphasis is on results instead of explanation, even though the explanatory element is usually more present than in performance measurement.

Peer reviews are expected to exert their influence through '*peer pressure*'. Public scrutiny, dialogue with peer countries, comparisons, and in some cases even ranking of countries, put pressure on the domestic public opinion, national administrations and policymakers. Media involvement and the ensuing public scrutiny are essential in enhancing the effectiveness of peer pressure. (OECD 2003j, 10.) A country seldom wants to be seen in a negative light among its peers, and therefore peer pressure may be a powerful tool in promoting compliance, notably in the 'laggard' states (see Beyeler 2002; Marcussen 2001, 8; Strang and Chang 1993). In this sense, peer pressure constitutes an instrument for international organisations' explicit or implicit attempts to stimulate the national bureaucrats' willingness to aspire to what the organisation defines as a 'modern', 'democratic' or 'economically efficient' public sector (March and Olsen 1998). Peer reviews must balance between on the one hand a high level of ownership felt by the reviewed country, which requires that the

country be closely involved in the process, and on the other, diminished credibility resulting from attempts by the reviewed State to unduly influence the final outcome.

Four main functions can be attributed to peer reviews (OECD 2003j, 17-18):

1. *Policy dialogue* to enhance the transmission of information on policies and their implementation.

2. Enhancing *transparency* towards peer countries and the public opinion, by engaging the reviewed country to make available to other countries and to the reviewing organisation information concerning its policies.

3. Contributing to *capacity building*, through a mutual learning process, both concerning the policies themselves and the peer review methodologies, (e.g. benchmarking or the use of quantitative indicators).

4. Constituting a system of 'soft enforcement', enhancing *compliance* of countries with internationally agreed policies, standards, and principles.

4.2 Evaluation approaches and the 'paradigm wars': lessons from evaluation research

In the field of evaluation, dichotomies such as those between positive and normative, objective and subjective, positivist and post-modern or hermeneutic, society and individual, structure and agency, macro and micro, judgement and explanation or understanding, evaluation and performance measurement, have been perhaps even more visible than in other disciplines, because of the place of evaluation at the crossroads between research and practice. While the former has largely abandoned the simplistic positivist model in favour of a constructivist approach of one kind or another, the rational view of policymaking still dominates evaluation practice and forms the basis for expectations about the utilisation of evaluations (e.g. Khakee 2003).

4.2.1 Experimental evaluation

A number of lessons relevant to the present study can be drawn from the history of evaluation research and the different evaluation approaches. The earliest approach, 'experimental evaluation', is based on the 'natural science' approach evoked above in chapter 2 and has close links with 'hard' social sciences, such as economics and statistics, as well as medicine and some branches of psychology and educational sciences. The approach fails to meet our criteria, since it is based on the false assumptions implied in the 'rational paradigm' of policymaking: systematic application of reason; unitary public interest; comprehensive analysis of available means and ends; predictable and controllable environment; role of the planner/evaluator as an independent expert detached from the political process; and a pluralistic society where competing interests all have access to power (Lawrence 2000). The paradigm also assumes that policymakers can clearly formulate policy goals. It fails to consider resource and cognitive limits, overestimates our ability to predict and control the environment, gives insufficient consideration to non-rational elements and various types of knowledge and wisdom, neglects dialogue, lacks attention to equity questions, and fails to take into account the institutional context. Most notably, the failure of the experimental approach to deliver on its promise – to provide reliable information on policies that would be successful in alleviating social problems – resulted in disillusioned conclusions of the type "nothing works" or that "most things have been found sometimes to work" (Pawson and Tilley 1997, 10). This failure has its origins in the inability of, experimental evaluation to open the 'black box' of policy implementation and explain *why* a particular programme does or does not work. Likewise, by its obsession with

random allocation, experimental evaluation, paradoxically, tries to eliminate what is essential to a programme: social conditions favourable to its success. (Pawson and Tilley 1997, 50-52; Kazi and Spurling 2000).

4.2.2 Counter-reaction to positivism: constructivist and pragmatic evaluation

Two evaluation approaches – constructivist and pragmatic evaluation – arose as direct responses to the shortcomings of experimental evaluation and as a part of the 'debate on positivism' in the 1970s, based on the rejection of the basic assumptions about causality, objectivity, and value-neutrality underpinning the experimental approach.

Constructivist evaluation encompasses a wide range of different approaches, many focusing primarily on community and local level participation and 'empowerment', seeking dialogue and dynamic interaction between the observer and the observed subject, and emphasising the importance of the context. It urges evaluators to look beyond mere outputs and focus on the causal mechanisms behind the programme results, the internal dynamics of programmes, and the differing everyday meanings attributed to such programmes by the various stakeholders. (Guba and Lincoln 1989; Schwandt 1997; Pawson and Tilley 1997, 17-18; Kushner 2002a, 2.)[66]

Constructivism has sometimes been criticised for neglecting the crucial importance of asymmetries of power, and naïvely assuming that genuine and open discourse could be carried out without the interference of unequal distribution of power.[67] While such critique may be relevant for many constructivist evaluations, it hardly applies to its empowerment-oriented, structural variants, which take precisely the unequal distribution of power as their primary point of departure.[68] According to Patton (1997b), the defining characteristic of empowerment evaluation is precisely the focus on self-determination, which is at the heart of its explicit social and political change agenda.[69]

---

[66] Constructivism encompasses a wide range of approaches, such as the responsive (Stake 1975; Stecher and Davis 1987, 36-39), illuminative (Parlett and Hamilton 1977), various types of democratic (MacDonald 1976; Floc'hlay and Plottu 1998; House and Howe 1999; Ryan 2004), empowerment (Fetterman 1994; 1996), stakeholder (Weiss 1983), or 4th Generation evaluation (Guba and Lincoln 1989).

[67] As seen above, this criticism has often been levelled against the Habermasian idea of the 'ideal speech situation'.

[68] Fetterman (1996, 5, cited in Patton 1997b, 148) defines empowerment evaluation as "the use of evaluation concepts, techniques and findings to foster improvement and self-determination".

[69] In the field of planning, the 'critical pragmatist' planning mode can be seen as a 'cousin' of empowerment evaluation. Forester (1989, 46) sees the former as a refinement of the traditional advocacy planning, a refinement based on the practical recognition of systematic sources of misinformation. An advocacy planner should point out the nature of the bias underlying information that is presented in plans set forth by the establishment, thereby performing a task similar to the legal technique of cross-examination. (Sager 2001, 212.)

Another criticism attacks the 'situational relativism' inherent in constructivism: since every context is considered unique, conceptualising and eventually measuring contextual similarities and difference becomes impossible (Pawson and Tilley 1997, 23). Moreover, constructivism tends to be suspicious of outcome-based methodologies, thus limiting its usefulness to only certain dimensions of practice (Kazi and Spurling 2000).

The third, and perhaps the most fundamental criticism concerns the assertion that since all beliefs are constructions, it is impossible to get beyond them, to the 'real' world. However, in line with the premises underlying this study, most advocates of constructivism today accept the idea that reality 'out there' does indeed exist, but maintain that all information we can obtain on this reality is unavoidably a social construction. Constructivism has evolved considerably over the past couple of decades, and has abandoned the simplistic extreme relativism it is often accused of (Ahonen 2001, 19). As Dahler-Larsen (2001, 336) argues, constructivist evaluation

"acknowledges institutional arrangements and power differentials, it does not refuse generalisation of evaluation results, and it accepts the concept of causality as such. Constructivists merely insist that programmes, programme theories and the contexts in which they operate, are all social constructions that could have been made otherwise. Moreover, it suggests ways to understand institutions, values, roles, norms, actors, and preferences as constructions so that we can understand why A leads to B in some contexts and not in others."

The merit of the *'pragmatic evaluation'* was that it placed utilisation at the centre of evaluation. The bottom line was that evaluators should stop being naïve about the organisational and political context of policymaking. The view of rational, knowledge-driven policymaking process was abandoned in favour of research seen mostly as ammunition for debate and intra-organisational arguments, rather than as a valid, tested body of propositions based on causal logic (Weiss 1980). Evaluation and its object were seen as 'power games', the evaluator's task therefore being to penetrate the networks of power that enable or obstruct the incremental change processes that evaluation is supposed to enhance. (Pawson and Tilley 1997.) Like constructivism, in its extreme version pragmatic evaluation adopts a relativistic epistemology: it abandons the search for truth in favour of a theory of social acceptability of ideas.

The ethical problem with pragmatic evaluation is the risk of blurring of the distinction between evaluation and politics, since serving the policymaker is taken as the ultimate quality criterion of evaluation, thus tending to privilege political expediency over scientific rigour. We shall come back to the question of evaluator's role in chapter 5. In the field of methodology, Pawson and Tilley (1997, 16) accuse pragmatism of not providing a solid theoretical basis for methodological choices, but leaving all to depend on the 'skill' of the evaluator. Methodological choices would be largely dictated according to the political practices and the expected use of evaluation results. (Pawson and Tilley 1997; Ahonen 2001, 108.)

Finally, in the same way as pragmatic planning, pragmatic evaluation seems to adapt to rather than critically examine the prevailing structures of power, thus fostering conservatism and inertia rather than action. The approach also seems incapable of addressing fundamental long-term choices and inequities – questions crucial from the point of view of sustainable development. (Ahonen 2001, 106; Lawrence 2000.)

4.2.3 Attempts at synthesis: combining it all or theory-building?

The proposed solution to the shortcomings of the previous evaluation approaches – to merge all approaches into a *'pluralistic' evaluation*, seeking simultaneously breadth and depth in evaluation[70] – is in line with the present thinking in many other fields of social science, which seek to overcome dichotomies and simplistic categorisations. It is also in tune with the 'methodological pluralism' underpinning this study. However, the Achilles' heel of pluralistic evaluation lies in its theoretical foundations. To be capable of managing the great quantity of information, the evaluator would need clear guidelines for setting priorities, 'seeing wood from the trees' (Pawson and Tilley 1997, 25, 154-155). A crucial question is whether a single methodological purpose should underpin each evaluation (Kushner 2002b),[71] or whether the tensions created by a confrontation of different methodological approaches and the use of different methods enables one to better take into account the concerns and values of the various stakeholders – an idea defended for instance by Greene (2002).

---

[70] For instance, through the so-called 'mixing methods' approach (e.g. Greene et al. 2001), 'pluralistic' evaluation (Baron and Monnier 2003), or 'multiple constituency evaluation' (Vartiainen 2001). Pluralism therefore refers both to the methods used and the participants involved in evaluation.

Instead of trying to merge all the previous approaches to a pluralistic mixture, a number of scholars (e.g. Chen 1990; Pawson and Tilley 1997; Julnes et al. 1998; Pawson 2002a; 2002b) have proposed the *'realistic'* and *'theory-based' evaluation*, which have recently attracted plenty of attention in the evaluation community. These approaches emphasise the importance of understanding the processes and mechanisms of change within programmes as well as the outcomes desired and achieved (e.g. Chen 1990; Julnes et al. 1998). It builds upon scientific realism or 'critical realism', in an attempt to avoid the problems inherent in the mutually opposing approaches of positivism and relativism, by focusing on the logic of explanation.

Realistic evaluation, as defined by Pawson and Tilley (1997, xii-xiv), rests on three fundamental premises:

- *reality* exists (in contrast with the view held by the strong version of constructivism),
- researcher can and should aspire *objectivity* through the use of scientific methodology, and
- there is *no universal logic of evaluation*, but the evaluator should try "to perfect a particular method of evaluation which will work for a specific class of projects in well-circumscribed circumstances".

In view of the premises of the present study, the first and the third point are more or less obvious and acceptable, whereas the second – objectivity as the use of scientific methodology – is more problematic.[72]

In explaining social phenomena, realistic evaluation refers to the so-called *'context-mechanism-outcome' configurations*. Explaining social regularities (outcome) is seen as the basic task of social inquiry, explanation taking the form of positing some underlying *mechanism*, which generates the regularity and thus consists of propositions about how the interplay between structure and agency has constituted the regularity. Workings of such mechanisms are contingent and conditional, and thus only triggered in particular local, historical or institutional *contexts*. (Pawson and Tilley 1997,

---

[71] According to Kushner (2002b), failing to define a single methodological purpose for an evaluation, the evaluator risks being co-opted by the dominant groups in the evaluation situation, and disguising inequalities of power as simply divergent views on epistemology.

[72] This is so in particular as it tends to perpetuate the futile endeavour of social sciences to emulate the natural sciences in producing cumulative and predictive theory, instead of grounding the analysis in the context and focusing on issues such as values and power. Methodological guidelines should be seen as "cautionary indicators of direction" at the most, rather than as "methodological imperatives" (Flyvbjerg 2001, 129).

71.) In contrast with the experimentalists that adhere to a 'successionist' view causality,[73] realists adopt the notion of *'generative' causation* (Harré 1972, 116-119), arguing that causal powers do not reside in particular objects or individuals but in the social relations and organisational structures. To illustrate the role of the context, Pawson and Tilley (1997, 57-58) take an example from the field of physics, noting that gunpowder does not always ignite in the presence of flame or spark. Hence, not only the chemical composition of the gunpowder, but also the right conditions (dry, adequately compacted gunpowder; sufficient oxygen, sufficient duration of the application of heat, etc.) are what provoke an explosion.

Following the logic of the context-mechanism-outcome configuration, realistic evaluation stresses the careful enunciation of the *programme theory* as a prerequisite of evaluation. A programme theory is a construction of a plausible and sensible model of how a public programme is supposed to function (Dahler-Larsen 2001).[74] Reconstructing the programme theory may help identify unnecessary programme components, locate intermediary changes, raise new questions, contribute to a paradigm shift, and provide clarity and focus for the evaluation (Birckmeyer and Weiss 2000.) Making explicit the assumptions that actors hold concerning the causal mechanisms between the inputs, outputs and outcomes of a policy, reconstructing the intervention theories through a participatory process helps make explicit the differences in the underlying perceptions and enhance understanding among different actors. This may also help clarify policy goals and analyse their level of ambition.

Realists do advocate *methodological pluralism*, but stress that the method has to be chosen carefully following the exact form of hypotheses developed through theory building. Generalisation sought in the realistic approach does not rely on replication in the experimentalist sense, but rather on the accumulation of knowledge through evaluations, achieved by developing and improving theory, not

---

[73] The basic idea of this view is that causation cannot be observed directly, but it has to be inferred from the repeated succession of one event by another. Systematic exclusion of every conceivable rival causal agent from the experiment leaves one with a single, certain, causal link. Various types of monitoring and statistical control are needed in order to ensure that the observed changes in the experimental group are not caused by other factors than the 'treatment' being studied. (see e.g. Harré 1972, 116-119.)

[74] Some authors (e.g. Mickwitz 2002; 2003; Vedung 1997, 301; Roman and Vedung 2000) prefer the term 'intervention theory', which they consider a more general term including theories of programmes, policies and policy instruments. Yet others talk about 'programme logic' (Scriven 1998). Vedung (1997, 301) defines intervention theory as "all empirical and normative suppositions that public interventions rest upon". Scriven (1998), in his criticism against the term 'theory-driven evaluation' and the conception of theory therein, notes that Chen's (1990) conception of theory is not consistent with the conventional way of defining a theory, which usually only explains how an entity produces is outputs. Instead of a theory, Chen should talk about 'programme logic'.

by carrying out a series of studies with reliable, replicable, and universal applicability (Pawson and Tilley 1997, 86, 116).[75]

However, opening the 'black box' of policy mechanisms may not be a universal requirement or a quality guarantee in evaluation. As Scriven (1998, 59) has noted, sometimes theory or the resources and expertise needed to conduct a theory-driven evaluation may not be available, and that one does not necessarily need a theory to establish a set of principles that govern the relations between programme components and explain the results of a programme by appeal to these principles. Explaining and identifying the underlying mechanisms is indeed an important, yet only one among many types of knowledge an evaluation is expected to provide (others including programme monitoring and inquiry about values). Julnes et al. (1998) have therefore argued that penetrating ever deeper into explaining the stratified reality of social phenomena may be justified in basic research, but in evaluation, utilisation should be the practical determinant of how far to go in explanation. In the same vein, Ahonen (2001, 108-109) argues that realistic evaluation is not a universally applicable solution, but each evaluation approach has its merits and its own range of application, which depends on the evaluation context.[76]

A key issue in the debates between realists and constructivists is to what extent the 'theory-based' evaluation approaches actually rely on the currently dominant logic of evaluation, linked to the modernist paradigm of reason, underpinned by technical reasoning or 'instrumental rationality', and concerned with the most effective means to given ends (Sanderson 2002). Critics accuse realistic evaluation of reductionism; of neglecting the mediating and contingent factors in explanations and the practical character of knowledge, grounded in everyday construction of meaning; failure to appreciate the contested character of scientific knowledge; and of the tendency to overlook the importance of philosophical assumptions behind the chosen theoretical approach (Schwandt 2003; Shaw and Crompton 2003). This perspective sees theory-based approaches mainly as a source of respectable scientific rationale to evaluation, helping to legitimise the position and status of an evaluation community, which has become increasingly disillusioned with its role as a 'servant' of

---

[75] "Generalization is not a matter of understanding the typicality of a program in terms of its routine conduct. Rather the process of generalization is essentially one of abstraction. We move from one case to another, not because they are descriptively similar, but because we have ideas that can encompass them both." (Pawson and Tilley 1997, 119.)

[76] Indeed, Ahonen (2001, 109) considers that the usefulness of realistic evaluation is linked to the particular political context of the UK, with its two-party system producing a situation of hegemony by one party. The interest in realistic evaluation has also been seen as a parallel to the rise of the 'Performance Management movement' and the recent calls, notably in the UK, for the so-called 'evidence-based policy', (e.g. Martin and Sanderson 1999; Sanderson 2000; Gray 2001; Davies and Nutley 2001; Pawson 2002a; 2002b). This approach argues that the role of evaluation is to promote

government. Theory-based approach would be inadequate as a unifying basis for evaluation, since it is conditioned by dominant structures of power and interest in governance processes and reinforces the conception of evaluation as a form of technical rationality, instead of seeing evaluation as a form of 'practical reason'. (Sanderson 2002; Schwandt 2003.)

4.2.4 Institutionalist evaluation?

In the light of these experiences from evaluation research, the challenge for an institutionalist evaluation is to construct a framework for a complexity-embracing approach, which would not, however, completely deny the existence of rational elements. Following Sanderson's (2000; 2002) ideas, one can argue that institutionalist theories could help build bridges between the illusory certainty involved in the modernist-rationalist order underpinning much of today's 'evidence-based' evaluation, and the danger of a pessimistic nihilism of some postmodernist ideas. An institutionalist approach endorsing complexity can be one way of overcoming the dualisms by emphasising the role of an intermediary level that regulates perception, interpretation and action at the individual level.

The experience from evaluation research lends support to the principle of methodological pluralism, since most scholars today agree that there is not, and probably can never be, one single 'correct' approach to evaluation. Public policies should continue to be evaluated from various disciplinary, theoretical, and philosophical perspectives. However, the question of limits to 'eclectism' is ever present: how far can one go in combining, or using in parallel, mutually contradictory approaches? This study adheres to the view that within a single evaluation or research, "the philosophical commitments, theoretical premises and the methods for analysis must form a consistent whole, so that its findings can be consistently asserted, both theoretically and empirically" (Leskinen 1994, 10). Methodological pluralism should therefore not be interpreted as a call for mixing, within an individual evaluation, methods whose philosophical and theoretical starting points conflict with each other.

Another lesson from evaluation research and practice underlines the importance of the context in choosing an appropriate evaluation approach. Each of the evaluation approaches presented above has its own qualities and disadvantages, and none of them can be declared completely inappropriate,

---

mutual understanding about what constitutes evidence, and in what context, for addressing different types of policy/practice questions (Davies and Nutley 2001, 2).

but there is no single, universally best approach either. The scientific rigour of the experimentalist approach is valuable especially in fields such as medical research; the constructivist evaluation is particularly suitable to situations involving widely varying values and worldviews among participants; pragmatic evaluation is at its best in highly politicised environments where practical utility of evaluation and sensitivity to power relationships are more important than explanation and scientific rigour. Pluralistic evaluation, in turn, tries to combine the best of all evaluation approaches, when no single dominant purpose for evaluation can be established. Finally, the theory-based approaches have brought attention to the need to ground evaluation in a consistent theoretical basis, which would facilitate learning through the refinement of theories.

A key criterion for an institutionalist evaluation is the extent to which it is capable of enhancing deliberative democracy. At the most fundamental level, deliberative democracy requires that an evaluation bring forward and facilitate communication between multiple realities, perspectives and worldviews, instead of putting forward a single judgement. The judgements should always be 'tested' in the light of alternative starting points, values, and worldviews. As we argued in the previous chapter, in the context of sustainable development, this would imply a capacity to question fundamental tenets such as the faith in progress, modernity, economic growth and technological development as sources of sustainability.

Causal theories and 'intervention theories' would best be used as tools for enhancing reasoned debate and deliberation among the parties involved, rather than serving as 'objective' descriptions of reality (Toulemonde 1997). These theories, or simply conceptions concerning causal relationships between different elements of policymaking process, would constitute parts of a coevolutionary framework of explanation, and a basis for the accumulation of knowledge across different evaluations. Such an approach underlies the attempts to combine 'scientific' modelling with broad participation of stakeholders through deliberative evaluation processes.

A major challenge in approaches attempting to foster reflexivity and deliberation stems from the tensions between, on the one hand, the usefulness of evaluation to the client – in the spirit of pragmatic evaluation – and, on the other, the critical reflexivity called for my constructivist and theory-driven evaluation approaches. The evaluator needs to have a clear understanding of the expectations of the main clients of evaluation – who may not always be interested in hearing different truths, but may have a number of political and strategic interests along with the cognitive ones. At the same time, being aware of the prevailing power relations is a key requirement enabling

the 'empowerment' of the weaker and poorly organised actors. The evaluator should be sensitive to the expectations of the main clients and be sufficiently diplomatic to engender the necessary mutual trust, yet the approach should not be too cautious and diplomatic if the evaluation were to retain its potential to provide an opportunity for critical self-assessment and reappraisal of the fundamental starting points behind policies.

Finally, taking into account the perspectives of a broad range of stakeholders with their different values, interests, worldviews, and disciplinary backgrounds necessarily involves an approach perceiving conflict as an essential element of any evaluation situation. Conflict should therefore not only be seen as unavoidable 'bad', but also a potentially powerful positive force that has the capacity of enhancing learning and deliberation by bringing to the table crucial issues of discord. Rather than seeking for consensus around the smallest common denominator – which indeed may be the only feasible solution in some situations – evaluation should seek to create forums for 'reasoned disagreement' and facilitate debate and understanding between sometimes radically conflicting perspectives.

Some of the means for managing complexity, through methodological pluralism, and the recognition of multiple rationalities in the context of plurality of values were already mentioned above as key elements of the institutionalist approach. The following characteristics can be added as crucial for an institutionalist evaluation:

1. *Overcoming dichotomies*. In the spirit of methodological pluralism, evaluation should combine different approaches, and seek syntheses over dichotomies such as substantive vs. procedural, positive vs. normative, yet ensuring sufficient consistency between the philosophical, theoretical and methodological choices within an individual evaluation.

2. *Causality*. Construction, examination, and refinement of causal models – programme or intervention theories – is a useful tool facilitating deliberation among participants in an evaluation situation.

3. *Context-dependence*. Recognising that each of the evaluation approaches has its strengths in specific situations, the evaluation approach should be chosen in function of the prevailing institutional context.

4. ***Power and conflict***. Awareness of the prevailing power relations is a precondition for combining usefulness to the main client groups and the facilitating the entry of the weaker groups and poorly represented arguments to deliberation. Rather than seeking consensus and agreement at any price, evaluation should first of all aim to increase the mutual understanding of the varying worldviews involved.

On the basis of the philosophical and theoretical principles underpinning sustainable development laid out in the previous chapter, and the insights from evaluation research presented in this chapter, we can now summarise the criteria for an 'institutionalist' evaluation of environmental policy for sustainable development. The criteria are classified under four headings: the planetary limits (the substantive side of sustainable development), sustainable development as a process (the procedural element), measurement of sustainable development, and deliberative democracy. This framework will subsequently be applied in chapter 8 to analyse the evaluation approach of the OECD Environmental Performance Reviews.

Box 4.1 Criteria for an institutionalist evaluation of environmental policy for sustainable development.

Planetary 'limits'

- Recognise and seek to define, in each particular evaluation context, the limits posed by the ecological and biophysical life-support systems to the maximum acceptable scale and quality of human activity, taking into account phenomena such as irreversibility, uncertainty, thresholds, and synergistic effects.

Sustainable development as a process

- See sustainable development not only as an end-state to be achieved, but essentially as a *coevolutionary process* of *institutional transformation*, with an emphasis on institutions, history, and phenomena such as cumulative causation, path dependence and lock-in. Pay attention to both formal and informal institutions and their multiple functions, as well as to the pervasiveness of power and conflict.

- Facilitate the understanding of the *links* – synergies and trade-offs – *between the different dimensions* of sustainable development, and at different scales, as well as help distinguish essential from the non-essential; take into account the qualitative differences between the dimensions of sustainable development, therefore conceptualising *the 'social'* through values, meaning, identity, and culture, rather than merely through distributional equity. Apply *causal explanation* as a tool for facilitating deliberation, rather than assuming it as an accurate description of reality.

- Recognise *complexity*, *multiple rationalities* and explanatory frameworks in the spirit of *methodological pluralism*, instead of relying exclusively on a single explanatory and methodological framework based on, for example, efficiency or maximisation of individual utility. Ground policy recommendations in a careful examination of the context, instead of universally valid 'axioms'.

- *Enhance critical reflexivity* by not taking premises such as progress, modernity, economic growth and technological development for given, as phenomena inherently desirable for sustainable development, but stimulate debate on such fundamental assumptions.

Measurement of sustainable development

- Ground the measurement of environmental performance and sustainable development in the idea of *incommensurability of values*, instead of seeking a single measure of performance or well-being.

- Pay equal attention to *qualitative and quantitative* elements.

- Seek appropriate level and degree of *aggregation*.

- Adopt a *broad conception of internalisation*, referring to the entire political process, instead of advocating internalisation simply within a framework of Pareto efficiency.

Deliberative democracy

- The evaluation process is *inclusive*, *deliberative*, and *dialogical*.

- The process and the recommendations are designed so as to maximise their potential to *reduce inequalities in the distribution of power*.

- The process and the recommendations *empower civil society* and foster a bottom-up approach to policymaking.

- The process strengthens the *'mediating' type of civil society*.

- The process and the reports help *overcome the dichotomies* such as positive-normative, objective-subjective, quantitative-qualitative, etc.

4.3 The 'ultimate' test: evaluation's influence

Given the practice-oriented character of evaluation, it is not surprising that the issues of utilisation and usefulness have come to dominate much of the debates on evaluation. They likewise further illustrate the topics briefly touched upon in the previous chapters, namely the uneasy relations between modernity, rationality and evaluation, notably the tensions between evaluation as policy advice on the one hand, and as a type of social critique, on the other. This section draws upon theories and experiences concerning evaluation use and evaluation influence, so as to develop a typology for analysing evaluation influence.

4.3.1 Types of evaluation use

The assumption of *direct use* of evaluation underpinned the traditional experimental evaluation accompanied by the technocratic-rationalistic policymaking model, and the idea of the "experimenting society"[77], committed to piecemeal social engineering on the basis of cold, rational calculations (see e.g. Vedung 1991; Pawson and Tilley 1997, 6).[78] By finding out the results of government interventions, evaluations would (1) help governments decide whether or not continue with particular policies; (2) expand and institutionalise successful programs and policies and cut back unsuccessful ones; and (3) find out which programmes or policies to modify and in which manner. The research findings having been published, rational policy-makers – possessing clearly defined preferences and aiming at clearly established policy goals – would adjust policies according to the most recent available scientific information. The primary clients of evaluations would be policymakers and the upper echelons of state bureaucracy. Evaluations were mostly 'summative', providing information about the impacts and results of programmes. (Rajavaara 1999, 43-44.)

---

[77] The vision of policy implementation typical of that time's 'social engineering' spirit is apparent in the opening passage of 'Reforms as experiments' (Campbell 1969, cited in Pawson and Tilley 1997, 6): "The United States and other modern nations should be ready for an experimental approach to social reform, an approach in which we try out new programs designed to cure specific social problems, in which we learn whether or not these programs are effective, and in which we retain, imitate, modify or discard them on the basis of their apparent effectiveness on the multiple imperfect criteria available."

[78] These ideas were inspired by Karl Popper's (1945) 'open society'. From the perspective of planning theory, it is interesting to note that Popper's scientific approach, involving participative, democratic, organic, interacting systems was a direct critique against the ideology of totalitarian and utopian planning, which implied an idea of predictability of future. Therefore, Popper's ideas can be seen as a reaction against the rational planning theory. Instead of aiming at certain goals, the society should rather concentrate on avoiding mistakes and correcting problems. "Locate a problem, try a solution, and check whether the solution has worked" became thus the slogan of piecemeal social engineering. (Taylor Fitz-Gibbon 2002, 141-142.) This view is implicit also in the OECD's (1999, 13) definition of the principal aims of evaluation as "improving decision-making, resource allocation and accountability".

This notion of direct, instrumental evaluation use soon proved overly optimistic: more often than not, policymakers seemed to ignore the results of evaluation and carry on with policies that they had chosen for other reasons. Scholars in many other fields of social science, such as sociology, political science, psychology, anthropology and even economics felt similar disappointment concerning the use of research results in policymaking. This resulted in disillusionment with government and accusations against policy-makers' irrationality, stupidity, and selfishness. (Leviton and Hughes 1981; Albæk 1995, 10-12; Lampinen 1992, 30-37; Muller and Surel 1998, 110; Pollit 1998; Rajavaara 1999, 43-44; Weiss 1999, 470; Hanberger 2001a, 58.)[79]

More recently, studies on research use have provided a more varied picture of the use of social science results, direct use no longer being considered as the only option. Evaluations do influence policies in more *indirect* ways, as one among many sources of information that decision-makers use in going about their daily activities. (Weiss 1999, 470-471.) This indirect use can be either *conceptual* or *political*. In the former case, evaluations and research provide general background information, leading to what Weiss (1980; 1987; 1998; 1999) has called *'enlightenment'*, "the percolation of new information, ideas and perspectives into the arenas in which decisions are made" (Weiss 1999, 471). Research affects decision-makers' problem definitions, and provides new perspectives on and insights into the problem area, instead of providing information for a single moment of decision, or a single decision-maker that, in fact, often simply does not exist.[80] 'Enlightenment' as an objective for evaluation and as a type of evaluation use has been forcefully put forward by the various constructivist evaluation approaches, in which the evaluator is supposed to function as a 'facilitator' in a negotiation process, bringing forward the various points of view involved and promoting consensus. Evaluation becomes a learning process, and the knowledge created through the process becomes utilised already while the evaluation is being carried out. Evaluation results are never unambiguous, but tend to be interpreted differently by different stakeholders (van der Meer 1999).

As opposed to conceptual use, *political use*, in turn, encompasses first of all *legitimisation*, which has frequently been seen in a negative light, evaluations serving as rationalistic rituals aimed at justifying decisions that have already been taken or policies that are already in place (Vedung 1991;

---

[79] 'Insufficient use' in policy-making has been identified as a major shortcoming of environmental impact assessments, as well. The EIA has often been neglected in the decision-making process, and there is often not evidence that the quality of the EIA procedure would have influenced decision-making. (e.g. Sager 2001, 215-216.)

Lampinen 1992, 30-37; Weiss 1999, 477). However, without legitimisation meaningful policy-making would in fact be impossible, because in a democracy long-term success of any policy depends on its legitimacy in the eyes of the citizenry (e.g. Feinstein 2002, 434; Valovirta 2002). Legitimising use also includes cases where evaluations are used to criticise positions of other actors. Empowerment evaluation likewise strives at legitimative use, since evaluation is aimed at supporting the interests of a specific group. A second type of political use is *tactical*, an evaluation being commissioned, for instance, in order to postpone decision-making by referring to an ongoing study. (Vedung 2001, 141.) Finally, political utilisation may be *'symbolic'*, research being used symbolically, to convey an image or a message (Weiss 1999, 477). Evaluation in such a case constitutes a 'façade' intended to give the impression of a rational organisation, which is ready to listen, is prepared for change, has a serious and competent management, and always takes rational decisions on the basis of collected data (Pollit 1998; Weiss 1999, 472-473; Vedung 2001, 141). Also such political battles would, however, eventually lead to 'enlightenment', through a complex process in which policymakers gradually become sensitised to the perspectives of social science (Weiss 1980).

4.3.2 Purposes and uses of evaluation

The different types of evaluation use are intimately linked with the overall purpose that an evaluation is set out to promote. These purposes can be classified into four major categories (e.g. Albæk 1995, 12; Vedung 1991, 74-81; Patton 1997a, 67-68; Weiss 1998; Baron 1999):

1. *Control or accountability* evaluations intend to provide an overall judgement on whether or not a policy has achieved its objectives. Such information may help politicians control the administration, high-level civil servants verify how well their subordinates perform their duties, citizens hold the government accountable for their deeds, or the persons directly affected by a programme or a policy check whether policies actually deliver what they were supposed to, and whether they respect the principles of equity, social justice, etc.

2. Evaluations for *development or improvement* emphasise learning in order to develop and improve the evaluated policy or programme, by seeking explanations for policy outcomes, instead of merely pointing out the results or outcomes. These evaluations are not primarily

---

[80] Public policy is "almost never a single, discrete, unitary phenomenon" (Greenberg et al. 1977, 1533). Instead of decisions being made, one could talk about 'knowledge creep' and 'decision accretion' (Weiss 1980). See also Owens and Rayner (1999).

targeted at the 'commissioning authority', but rather at the personnel responsible for implementation.

3. *Enlightenment or creation of knowledge*. These evaluations seek to improve the understanding of policy-makers, programme or policy personnel or citizens concerning the policy processes and factors affecting the course of events. Evaluations may contribute to the knowledge base of an academic discipline, but can also provide basic information for concrete decision-making, often in areas not directly related to that of the evaluated policy itself. 'Enlightenment' is seldom seen as the primary purpose of an evaluation, and is unlikely to ever suffice alone to justify that an evaluation be conducted (Albæk (1995, 17).

4. *'Potemkin village'*. This 'illegitimate' purpose of evaluation denotes the attempt by policy-makers, government officials, civil servants, or programme managers to give the appearance that everything goes well, to win time, to create an illusion or rational policymaking, or simply to hide the negative aspects and failures of a policy.

In practice, a single evaluation seldom focuses exclusively on any of the above-mentioned goals, but tries to accommodate and combine at least the three first-mentioned 'legitimate' purposes. Thus, the real challenge becomes that of reconciling the contradictory demands posed by these purposes, focusing on the most important ones and minimising the shortcomings associated with each one.

Recently, there has been a renewed interest in developing alternative frameworks of evaluation use in policy-making, focusing on 'process use' (e.g. Patton 1997a; 1998; Forss et al. 2002), argumentation (Valovirta (2002), dialogue (Abma et al. 2001; Schwandt 2001a; 2001b), or learning (Baron 1999; Siebenhühner 2001; Torres et al. 2001; special issue of 'Evaluation', 2000, Vol. 6, no. 3) as key elements and mechanisms of evaluation use. 'Realistic' evaluation, in turn, has stressed the use of evaluation findings as the accumulation of knowledge from different evaluations, consisting of generalisation through the formulation and refinement of 'middle-range theories' that would help understand which particular contextual conditions and mechanisms have brought about the observed outcomes. Utilisation would be seen as a mutual learning process between the evaluator and the different stakeholders. (Pawson and Tilley 1997.)

To sum up, evaluations are only seldom used directly, according to the rational policy making model, but indirect, more subtle ways of utilisation are much more common. At its best, indirect use of evaluation serves to 'enlighten' the various parties involved and thus constitute one of the sources of information used by actors in their everyday decisions. Often, evaluation use is

motivated by political factors and the various participants can consider use as more or less legitimate.

4.2.3 Use or influence of evaluation?

In this study, 'evaluation influence' through argumentative processes is adopted as the organising concept for understanding and explaining the impacts of evaluation on public policymaking. 'Influence' instead of 'use' has been advocated by authors from the 'realist' school as a concept better capable of capturing the various mechanisms that lead to different end-states of 'use'.

While the increased attention to 'use' or 'utilisation' – for instance, in the so-called 'utilisation-focused evaluation' (Patton 1997a) – has fulfilled a useful function (e.g. participation of stakeholders has become commonplace as it has been expected to increase evaluation use), a number of problems are associated with the concept of 'use' (Henry and Mark 2003; Morabito 2002). Particularly problematic is the parallel use of normative and descriptive meanings attributed to the term.[81] The normative focus on use tends to divert attention from general comparative analyses to individual evaluations. It also creates a need to demonstrate that at least some use occurred, thus limiting attention to only one or a few form(s) of influence at a time, even though evidence concerning evaluation influence might require casting a broader net. Often, unintended impacts go unnoticed as attention is geared towards intended use only.

A further problem is what Henry (2000, 8-10) calls the 'paradox of persuasion': focus on utilisation may jeopardise the credibility and legitimacy of the evaluation itself, if the evaluation is planned, designed, and executed to facilitate persuasion as the particular form of use. The evaluator is compelled to begin with the objective of producing persuasive evidence, and the endeavour becomes 'pseudo-evaluation', or political advocacy, evaluators becoming advocates of a particular position, instead of attempting to raise the status of information in public deliberations. As was seen above, this is one of the main points of criticism raised by Pawson and Tilley (1997) against the pragmatic evaluation approach. Henry and Mark (2003) therefore call for more descriptive rather than normative studies on the influence and consequences of evaluation, to capture the different

---

[81] Henry and Mark (2003) identify four meanings in which the term evaluation use appear in the literature: (1) the descriptive sense, attempting to describe how evaluations are used; (2) the normative view, implying the idea that evaluations *should* be used; (3) use as a specific approach to evaluation (Patton's utilisation-focused evaluation), which can be considered as a particular variant of the normative use of the term; and (4) use as an organisational construct in the theory of evaluation.

ways in which evaluations exert their influence, and the conditions in which such an influence could take place.

4.3.4 Evaluation influence and learning

Collective learning is increasingly being considered as the main type of evaluation use, and the ultimate goal of evaluations (see e.g. Baron 1999, 62-65). Learning is likewise at the core of institutional and evolutionary economics (e.g. Hahtola 1990; Hodgson 2002).[82]

Van der Knaap (1995) identifies three perspectives from which learning from evaluation can be examined. *Cybernetic control* represents direct use, evaluation functioning as a dispenser of reliable feedback information that policy-making institutions use in order to correct error and improve their policy theories. *Cognitive development*, in turn, is close to the enlightenment use, evaluations providing stimuli to the learning process conceived of as a cycle, in which the learning subject moves through different stages. Evaluations draw attention to contradictions between the reality of the situation and the individual's perception of it. *Social learning* is based upon communication, in which people exchange knowledge and beliefs through employing shared linguistic notions. In this context, evaluations provide strong and empirically founded arguments in the ongoing political discourse. This type of learning may sometimes come close to what was above defined as legitimisation. (Balthasar and Rieder 2000.) The key element distinguishing social learning from the two previous ones, however, is that it draws attention to the institutional conditions within which knowledge is being produced and used. Successful social learning requires, first, a considerable degree of 'structural openness', i.e. the possibility for the participants to negotiate "problem perceptions, interests, uncertainties and alternative solutions" (Glasbergen 1997, 189), and, second,

---

[82] Different definitions of learning include that of Marcussen (2001, 6), who emphasises behavioural change in defining learning as "the process that takes place when individuals or social groups permanently change behaviour as a result of an experience". Siebenhühner (2001, 3) defines learning as "a process of long-lasting change in the behaviour or the general ability to behave in a certain way that is founded on changes in knowledge". Knowledge is in this definition understood in a broad sense incorporating cognitive, normative and affective elements (ibid.). Finally, Baron (1999, 231) gives the most elaborate definition, focusing on the capacity to change: "the learning processes of an individual or collective entity are characterised by a complexification of its mental models resulting from the acquisition, creation or internal transfer of knowledge and are translated into an increasing capacity of the entity to respond to perturbations in its environment" ("Les processus d'apprentissage d'une entité (individuelle ou collective) se caractérisent par une complexification de ses modèles mentaux issue d'une acquisition, d'une création ou d'un transfert interne de connaissances et se concrétisent par un accroissement de ses capacités de réponse aux perturbations de son environnement.")

commitment on the part of the participants to solving common problems instead of driving narrow self-interests. (Valve 2003, 18.)[83]

While collective learning is clearly more interesting from the point of view of this study, impacts at the individual level are a necessary, albeit not alone sufficient condition for learning at the collective level (Baron 1999, 211-216; Siebenhühner 2001). Little research has been done on the interaction between individual and collective learning, and opinions diverge about the exact mechanisms of transfer of learning from individual to collective level, but the transfer of knowledge between the members of the organisation is recognised as a key element. This is one of the main arguments for involving various stakeholders in evaluation, since participation enhances chances of interaction and communication between individuals. (Baron 1999, 215, 222-223.)

4.3.5 Dialogue, argumentation and 'process use' as pathways of evaluation influence

The typologies presented above are helpful in categorising different types of learning effects, but do not tell much about the mechanisms and pathways through which learning takes place. *Dialogue* and *argumentation* are among such key mechanisms. Dialogue has been identified as a major vector of especially collective learning, and as a key element of deliberative democracy.[84] Interaction between individuals, i.e. communication and dialogue, is a crucial element of organisational learning. Dialogue can be seen either as an instrument towards an end or as an event fostering understanding and respect across difference (Abma et al. 2001, 166; Schwandt 2001a, 268-269).[85] From the former, *instrumental* perspective, dialogue is a matter of how to proceed rationally in the exchange of messages for the purposes of obtaining information, bringing some state of affairs to pass, reaching agreement, settling a dispute, persuading the other, etc. The latter approach sees dialogue in *substantive* terms, relating to ethical or moral aspects, to identity, to normative ideas of being a 'good human being', to the very idea of our 'understanding' one another, etc. (Abma 1998; Abma et al. 2001, 166; Schwandt 2001a, 268-269.) Indeed, for Schwandt (1997), dialogue is the essence of evaluation. The purpose of evaluation largely determines which one of the two is more

---

[83] Other disciplines have developed similar typologies. In organisation research, the classification of Argyris and Schön (1996), consisting of single-loop learning, double-loop learning and deutero-learning is very close to that of van der Knaap. Glasbergen (1997) makes essentially the same distinction in referring to *technical*, *conceptual* and *social* learning.

[84] See e.g. the special number of the journal 'Evaluation', devoted to dialogue: year 2001, Vol 7, no. 2.

[85] A similar distinction is made by Abma (1998, 436-437), who writes about dialogue dominated by argumentative rationality on the one hand, and dialogue as an open conversation on the other.

important: the instrumental aspect dominates in an accountability perspective, whereas a learning-oriented evaluation gives more emphasis to the substantive type of dialogue.

Our third key concept, *argumentation*, links with the ideas of learning through dialogue. Valovirta (2002) has focused on 'social learning' in stressing the role of argumentation as a fundamental, yet often neglected element in the utilisation of evaluation (see also Feinstein 2002, 434).[86] He considers that "perhaps the most important value of evaluation is that it incites argumentation and can direct it towards a more reasoned debate" (Valovirta 2002, 77). He therefore lays emphasis on the instrumental role of argumentative dialogue. Nevertheless, he agrees with Schwandt (2001a) in rejecting the idea of evaluative information as indisputable knowledge, seeing it instead as a collection of arguments, which can be debated, accepted and disputed.

Learning through dialogue – both in its instrumental and substantive forms, through argumentation and more 'understanding' type of conversation – is an essential part of what has become known as the *'process use'* of evaluations (see e.g. Patton 1997a; 1998; Forss et al. 2002; Morabito 2002; Valovirta 2002).[87] Process use is distinguished from what is conventionally understood as utilisation in that it is stimulated, not by the findings of an evaluation, but by interaction between individuals participating in the process of evaluation. While a common sentiment in the field is that instead of being used, many evaluations simply end up gathering dust on the shelves, the matter seems different if one looks at the ways in which the evaluation process itself may serve policymaking. Evaluations rarely provide completely new insights – often the stakeholders and participants were aware of most of the impacts of policies or programmes already before the evaluation, and evaluations more often confirm existing beliefs than give rise to new ones – but participants nevertheless frequently say "it was useful to go through the process". (Baron 1999, 322; Uusikylä and Virtanen 2000, 55; Forss et al. 2002, 32; Valovirta 2002, 73.)

---

[86] Feinstein (2002, 434) brings attention to the fact that the use of evaluation for persuasion – through an argumentative dialogue – has only recently been acknowledged in evaluation literature, but that there is a school of thought (and practice) in economics that highlights the key role of persuasion (e.g. McCloskey 1994).
[87] For Patton (1997a, 90), process use is a phenomenon that "is indicated by individual changes in thinking and behaviour, and program or organizational changes in procedures and culture, that occur among those involved in evaluation as a result of the learning that occurs during the evaluation process."

4.3.6 Typology for analysing evaluation influence

By combining the framework of Valovirta (2002) on argumentation, that of Henry and Mark (2003) on 'evaluation influence', and Forss et al. (2002) on 'process use' we can now present the typology applied in the present study for assessing evaluation influence.

The influence of evaluation is studied at three levels. First, insights from the evaluation findings or the evaluation process are interpreted at the level of the individual. Such interpretation can result in *confirmation of existing beliefs*, *rejection* of evaluation findings, or *new, transformed beliefs and capacities* (Valovirta 2002). These new beliefs may involve *attitude change*, *increased or decreased salience, elaboration*, or *priming*.

Attitude change is what most evaluators would consider as evaluation use, in the sense of 'enlightenment', but Henry and Mark (2003) see attitude change as a more specific and narrower concept than 'enlightenment'. *Salience* refers to the importance an actor attaches to an issue. *Elaboration* refers to the degree that an individual actually thinks about an issue, in contrast to 'peripheral' or 'heuristic' processing in which an individual relies on other cues, such as who is supporting an issue, to form his/her opinion about an issue (ibid.). The evaluation process prompts actors to elaborate on an issue, thus making them more aware of the goals and objectives of their own work. *Priming* describes the phenomenon by which a concept temporarily becomes cognitively accessible. Priming can affect the very *way* one thinks about an issue. In general, priming a concept makes it and related concerns more influential in the judgement process. (Krosnick and Kinder 1989; Krosnick and Brannon 1993.) The evaluation process can also lead to individual level impacts such as *skill acquisition*, i.e. a process whereby people affected by evaluation acquire new abilities such as evaluation methodologies or cooperation skills. Finally, once the evaluation findings or experiences from participation in evaluation have gone through an interpretation processes, they may also prompt the individual to change his/her individual *behaviour*.

The second level of influence – the process whereby the interpretations enter into a collective process of deliberation, negotiation, argumentation and dialogue – is of key importance for the present study. The attention is focused on the means applied by the various actors (governmental bodies, industry, NGOs, academics, media, etc.) to use evaluations in order to *persuade* others to

adopt a certain line of action, *legitimise* their own actions, *criticise* those of others or the evaluation findings, or *defend* themselves against criticism based on the evaluation findings (Valovirta 2002).[88]

Finally, these argumentative processes may result in effects at the collective level, on the evaluated policies or policymaking more generally. These effects may concern the *decisions and actions* – the level that has usually attracted most attention in studies on evaluation use. These are typically what Henry and Mark (2003) call *policy change*.[89] Evaluation findings or recommendations can therefore be (1) adopted as such; (2) adopted after discussion in a modified form; (3) rejected after a discussion; or (4) rejected without discussion (Valovirta 2002). The latter represents a case in which an actor has the power to impose its own views on other actors.

Often, however, evaluations fail to produce significant policy changes, but may have a significant role in *creating new shared understandings* through the interactive process of argumentation and dialogue, called by Henry and Mark (2003) *'policy-oriented learning'*.[90] Forss et al. (2002) have identified policy-oriented learning as one of the types of process use, stemming from the process whereby differences of values, priorities and objectives held by stakeholders are made explicit, and the arbitration of interests rendered easier. This category also includes modification of social norms ('socialisation').[91] In general terms, creation of new, shared understandings can be seen to consist of three types of alternative outcomes: production of a synthesis, a winner, or a deadlock. Synthesis emerges as a compromise result of an argumentation process, in which each party modifies its initial standpoint. The imbalance of power in argumentative settings implies the potential of producing not only a synthesis, but also a winner, who has the 'last word' (Billig 1987, 106 cited in Valovirta 2002, 67). When views diverge too widely, and no actor has sufficient power to impose her view upon others, the argumentative process results in a deadlock. Even in such a case, just as when a recommendation is rejected, one cannot say that the evaluation would have had no impact,

---

[88] However, evaluations can also be used as background material, without an intention to persuade, legitimise, criticise, or defend. Likewise, government information officers or mass media may report evaluation findings in an innocently 'objective' spirit, with the sole objective to inform the public about the findings of an evaluation.

[89] *Policy change* refers here to major transformations or minor, more incremental modifications of policies (Sabatier & Jenkins-Smith 1993) within the environment in which the evaluation occurred. A relevant point is that the type of evaluation findings seems to affect policy change: negative findings, or findings that show a growing problem, may be more influential for policy change than those that show things are going well (Mintrom 1997).

[90] *Policy-oriented learning* refers to "relatively enduring alterations of thought or behavioural intentions that result from experience and/or new information and that are concerned with the attainment or revision of policy objectives" (Sabatier and Jenkins-Smith 1999, 123). A single evaluation may be unlikely to change pre-existing beliefs because "actors tend to accept information confirming existing beliefs and to screen out dissonant information" (ibid., 145). Therefore, policy-oriented learning is usually a cumulative effect of several evaluations and other sources of influence.

since the argumentative process inevitably leads different parties to improve information on each others' reasoning and sharpen their own arguments. This, in turn, may have further tangible impacts on policymaking in other areas, under other circumstances, etc.[92]

The third category of effects consists of *increased or decreased legitimacy* of past and future actions, actors or organisations. Legitimacy can be internal or external: an example of 'process use' enhancing *internal legitimacy* is that of 'boosting moral': participation in the evaluation process may improve the motivation of the actors by strengthening their commitment to the project, and motivation to pursue their activities, even in cases where evaluation has brought forward shortcomings, mistakes and abuses (Forss et al. 2002).[93] *Strengthening the project* may result from attitude change, changes in perceived salience and elaboration at the individual level. Forss et al. (2002) have noted that the mere fact that a programme or a policy is being evaluated, stimulates reflection on its ultimate goals, aims and means of implementation, especially since evaluation often also brings the managers in contact with intended beneficiaries of the project. Interaction with evaluators, in turn, provides politicians and civil servants with an opportunity to discuss with someone who is interested, who thinks sharply about the issues, and who may be ready to discuss difficult problems of resource allocation, strategic choice and responses. This is worthwhile in its own right, and can lead to changes quite apart from the topics that arise out of the evaluation's terms of reference. (See also Foulkes 1998, 359; Weiss 1998, 25.) *External legitimacy* refers to the legitimacy of a policy, an actor or an organisation in the eyes of outsiders: a ministry or a government policy may gain more legitimacy through positive evaluation findings, and lose it in the opposite case. An increased legitimacy of a policy may stimulate *imitation* by other social actors, or inversely, decreased legitimacy of a policy as a result of negative evaluation findings may incite the evaluated actor to imitate others.

An evaluation process may also help create *professional networks*, or modify the *political agenda*.[94] Impacts on agenda-setting work notably through priming, elaboration and argumentation throughout the evaluation process.

---

[91] *Social norms* are shared standards about what is an appropriate behaviour in a specific context and at a particular moment in time. Social norms can affect behaviour by becoming internalised by the individual, or by affecting his/her conceptions about social expectations towards his/her behaviour. (Henry and Mark 2003.)

[92] While the creation of new, shared understandings is similar to socialisation, the two concepts are not identical. New understandings can be shared, yet without necessarily being internalised as norms, as implied in socialisation.

[93] At the individual level, 'boosting moral' can be classified under the heading 'attitude change' (Henry and Mark 2003).

[94] *Agenda setting* refers to an issue moving onto the media agenda, the public agenda, or the agenda of a government institution (Zhu, 1992).

The different ways of evaluation influence are summarised in table 4.1.

Table 4.1. Summary of the potential types and pathways of influence from an evaluation.

Processes at the individual level

BELIEFS **CAPACITIES**

Confirmed Transformed Rejected **Skills**

Attitude change
Salience
Elaboration
Priming

Evaluation methods
Cooperation

Processes at interpersonal level: dialogue, deliberation and argumentation

PERSUASION LEGITIMISATION CRITICISM DEFENCE
(justification)

Impacts at the collective level

DECISIONS AND ACTIONS **NEW SHARED UNDERSTANDINGS** (~ Policy-oriented learning) **LEGITIMACY** Increased or decreased **AGENDA-SETTING** **NETWORKING**

4.3.7 Influence of evaluation on democracy

While evaluators often wish to see themselves as advocates of a democratic society, they frequently disagree amongst each other on the kind of democracy they would like to promote (e.g. Stake 2004). The matter becomes even more complicated in view of the uncontrollability and unpredictability of many of the impacts of evaluation – including those on democracy. When an external evaluation helps to legitimise a policy or a program, it inevitably strengthens one or

another notion of democracy, and the processes through which an evaluation affects social life have profound impacts on democracy (Hanberger 2001b, 212).

4.3.7.1 Types of democracy and civil society

Hanberger's (2001b) distinction between three types of democracy and four conceptions of civil society will serve as the basis for the analysis in this study. Democracy can be classified into three types: elitist/representative, participatory, and deliberative democracy. In the *elitist/representative* model, citizens can control their government by choosing among competing elites, but are not given a direct role in the policy process. Indeed, an exclusively elitist concept of democracy perceives all participation beyond voting in elections harmful to democracy, as it is seen to promote individual and group interest to the detriment of public interest. The *participatory* notion of democracy assumes that citizen involvement is the most important quality of a democracy. Individuals are encouraged to participate in the policy planning process, but no qualitative requirements are linked to participation. The *deliberative* type of democracy, described briefly above in chapter 2, goes one step further in asserting that the idea of democracy can only be realised through discussions among free and equal citizens. Democracy is not an aggregation of opinions of the majority, but is founded on a common commitment to a mode of reasoning on matters of public policy. Deliberative democracy is usually also associated with some kind of decentralised society and participation by ordinary citizens.

A *libertarian civil society* is composed of two opposing 'cells' – the private sector and the state – confronting each other in a zero-sum game. Individuals are presumed to act as 'economic animals' maximising their own interest. A policy or a programme strengthens a libertarian view of civil society if it promotes competition among firms and the market forces, facilitates public and individual choices, advocates 'liberty', 'self-reliance', and a society in which 'economic man' and 'Gesellschaft-relations'[95] are viewed as ideals. A *communitarian civil society* is fostered, when collaboration within the community/group increases, and when the community becomes empowered. Values underpinning this notion are self-determination, identity/meaning, and

---

[95] The German sociologist Ferdinand Tönnies (1855-1936) made a distinction between two basic types of social groups. He argued that there are two basic forms of human will: the essential will, which is the underlying, organic, or instinctive driving force; and arbitrary will, which is deliberative, purposive, and future (goal) oriented. Tönnies called groups that form around essential will, in which membership is self-fulfilling, *Gemeinschaft* (often translated as community). Groups in which membership was sustained by some instrumental goal or definite end he termed *Gesellschaft* (often translated as society). Gemeinschaft was exemplified by the family or neighborhood; Gesellschaft, by the city or the state. (Truzzi 1971.)

'Gemeinschaft'. The *mediating* model of civil society is defined as "civic communities that qualify as membership communities but are sufficiently open and egalitarian to permit civic participation on a voluntary basis". A mediating civil society needs a vital democratic sphere partly outside and partly inside the formal democratic institutions. Self-governance, civic virtue, and 'collaborative society' are among its underlying values. A mediating civil society is supported if policies benefit target groups and citizens, that is, those whom the policy (or programme) was intended to serve, and if they lead to closer engagement and greater responsibility of ordinary citizens and voluntary organisations. The key unifying feature in the variety of *feminist* theories is the notion that the private sphere is part of the civil sphere, and no clear distinction is made between the two. Strengthening a feminist civil society would require that all types of households benefit and social or gender inequalities be reduced. (Hanberger 2001b.)[96]

4.3.7.2 Ways of 'taking evaluation to the people'

Following his distinction of democracy into three models, Hanberger (2001b) describes three approaches to strengthening democracy through evaluation, or "taking evaluation to the people". The *positivist or technocratic approach,* close to experimental evaluation, sees evaluation essentially as scientific work that aims to keep facts and values apart and seek the truth. Such expert-oriented approach would promote a libertarian civil society and elitist/representative democracy. Evaluators are expected to provide decision-makers with reliable, value-free 'facts', and assess the extent to which goals have been achieved. The evaluator serves politicians and bureaucrats directly and the people indirectly. However, while the technocratic approach strengthens elitist democracy, it may also reinforce the common (mis)perception of science and evaluation as value-free activities.

The *advocacy approach* to evaluation is close to the pragmatic evaluation approach, but many constructivist evaluations fall within this category as well. Such an approach would promote participatory democracy and a libertarian, communitarian, or feminist notion of civil society, depending on who is being empowered. The evaluator can be an advocate serving the client's needs, or a facilitator, educating and assisting the client to become a skilful self-evaluator. The

---

[96] Hanberger's definition of feminist civil society can be criticised for being somewhat limited, representing the idea of feminist scholarship as a viewpoint 'of women'. Nelson (2003, 56) argues that feminism should instead be seen as a "viewpoint that includes previously denigrated dimensions of experience, such as change, emotions, and the sense of moral and aesthetic value, in the reality to be studied, and even more, recognises their importance in the reality to be created."

problem is that instead of serving the broader multicultural society, this approach might promote a fragmented society, giving precedence to the interests of a deprived group or community. Empowerment processes may also expose the evaluator to ethical dilemmas, if the client acts in ways that conflict with the evaluator's democratic values.

While evaluators in the two above-described approaches do not explicitly take account of different arguments unless asked by their clients to do so, the *mediating approach* implies that some kind of practical deliberation between stakeholders needs to take place. This approach promotes a mediating civil society and deliberative democracy (a feminist civil society may also be compatible with this approach). The evaluator would act as a counsellor and/or mediator inquiring, learning, and mediating together with various stakeholders, trying to describe the current situation in fair ways, taking into account critical arguments in the face of difference and conflict, and finding practical solutions to collective problems.

Improving the technocratic approach to evaluation – one possible way of taking evaluation to the people – must be linked to and be part of the greater project of restoring legitimacy for elite democracy and enhancing a libertarian civil society. The advocacy and mediating approaches to evaluation could be perceived as threats to the technocratic approach as well as to elite democracy, yet in fact, they do not challenge democracy as regime, but rather illuminate the need to reform elite democracy. As was seen above, all three models of democracy endorse representative democracy, but differ regarding citizens' role and involvement in public affairs.

Hanberger (2001b) considers the mediating approach as the most promising way of 'taking evaluation to the people', as it could help (de)legitimise a policy or a programme by means of openness, discourse, and critical assessment. By contrast, the technocratic approach supports existing structures, and would not solve the legitimacy crisis of authorities, whereas broadly applied, the advocacy approach could lead to a fragmented society and would only partly resolve the legitimacy crisis.

Hanberger's (2001b) typology illustrates the complementary roles that the different ways of 'taking evaluation to the people' can play in enhancing democracy. To enhance deliberative democracy, a political process does not need to involve participation at all levels and throughout the whole process. For instance, poorly organised participation can, in fact, do more harm than exclusion of certain groups from decision-making. In a situation characterised by obvious imbalances of power,

deliberation may serve as an incentive to manipulate, and produce pathologies rather than better decision-making (Stokes 1998; Przeworski 1998). The essential question is to what extent a political process enhances deliberative democracy as a whole – 'empowerment' of the weaker groups being one of the primary ways towards such a goal. Likewise, for example, an accountability evaluation, even if conducted in the spirit of technocratic-positivist approach, can serve an important democratic function in enhancing citizens' possibilities to control the government. To be credible, an accountability evaluation needs to be carried out by external evaluators that are sufficiently independent of powerful interest groups. This may require limiting the participation of groups having significant interests at stake. In order to grasp the whole panoply of impacts of an evaluation on democracy, the whole 'evaluation cycle' should be examined within its political and organisational context, as suggested by Segerholm (2003). Table 4.2 describes the impacts of different types of evaluation – different ways of 'taking evaluation to people' – on democracy and civil society.

Table 4.2. Impacts from different ways of 'taking evaluation to people' on democracy and civil society.

WAY OF 'TAKING EVALUATION TO PEOPLE'	THE TYPE OF DEMOCRACY ENHANCED	THE TYPE OF CIVIL SOCIETY STRENGTHENED
Positivist-technocratic approach	Representative-elitist	Libertarian
Advocacy approach	Participatory	Communitarian
Mediating approach	Deliberative	Mediating Feminist

4.4 Factors affecting evaluation influence

The third question addressed in this chapter is why evaluation influence varies from one situation to another, one country to another, etc. This also entails asking whether it all depends on the country-specific contingent factors, or whether one can identify more general regularities. In the following, literatures on evaluation, political science and international relations shall be drawn upon in order to identify the factors that have the potential to condition the influence of evaluations on policy. The categories of factors retained here are

- the role of an international organisation as evaluator;
- the involved actors with their expectations, interpretative frames and policy networks;
- the general policy context; and
- the evaluation design.

4.4.1 Setting the scene: Role of the evaluator

While the role of the evaluator is a decisive factor in any evaluation, it has particular importance when evaluations are carried out by an international organisation. Three issues are of special relevance. First, the evaluator may be a part of the evaluated entity – an *internal evaluator* – or an independent outside expert, an *external evaluator*. Both have their advantages, internal evaluation having greater likelihood of being salient and legitimate in the eyes of the individuals involved in policy implementation, while external evaluations tend to be more credible and legitimate in the eyes of other stakeholders (e.g. Mabry 2002, 149).[97] Therefore, internal evaluations are usually preferred when the main purpose is policy improvement and development, whereas accountability evaluation requires external evaluation (see e.g. Rajavaara 1999). Since most evaluations aim at both objectives, the issue is complex and striking an appropriate balance between the two functions is not easy.

The second question relates to the perceived *objectivity of the evaluator* and the degree to which she should provide political recommendations and/or openly defend the interests of a particular constituency. According to the positivistic view on scientific objectivity the evaluator should detach himself as much as possible from the evaluated entity in order to ensure the validity of the findings.

The debates between the advocates of a 'truth-seeking' model of evaluation, and those inspired by the constructivist ideas of value-ladenness of all human activity[98] have focused, on the one hand, on the evaluator's role, notably her ethics, credibility and utility (e.g. Scriven 1997; Chelimsky 1998; House 1983; Weiss 1999; Leca 1997; Patton 1997b, 157-159; Barbier 1999; Geva-May and Pal 1999; Mabry 2002, 149-151; Stake 2004) – problems associated with the evaluator taking on the role of a policy advocate. A related strand of discussion has concentrated on the perceived 'misuse' of evaluation as a tool for exercising power over other stakeholders (e.g. Cook et al. 1980; Pollit 1998).

A major dividing line goes between those who claim that a clear distinction should be made between policy analysis and evaluation (Geva-May and Pal 1999) or between 'political assessment' and 'evaluation of policies' (Barbier 1999, 375), and those who argue that drawing such distinction is impossible, given the different ways of evaluators to see and interpret the socially constructed reality (e.g. Stake 2004). The former call for strong professional evaluator ethics and point out that even if the evaluator can never be totally objective, since evaluations enter into the political process, competing with other factors influencing the political game, she can nevertheless aim at supplying an honest report of her premises, her objectives, her research protocol, the techniques used to collect data, her inferences, her generalisation and the way she has drawn the practical conclusions (Leca 1997, 11).

Mabry (2002, 149-150) presents the dilemma in a nutshell by noting that since evaluators are part of the machinery of politics, "explicit advocacy may threaten the credibility of an evaluation, while failure to advocate may threaten its ethics". Failure to advocate would therefore imply sacrificing the social good for philosophical viability (Mabry 1997). There is a continuous need to balance between clients' needs, and truthfulness to the evaluator's own ideals and intellectual honesty – between utility and uprightness in the face of manipulation (Mabry 2002, 151).

The present study takes the view that emphasises the importance of the context, rejects the possibility of drawing a sharp distinction between the positive and normative claims, and argues that the quality of science should be judged not only on the basis of internal criteria relating to the

---

[97] The questions concerning the credibility, legitimacy and salience of evaluations shall be further deal with in section 4.4.2.2.

[98] Michael Scriven (2001) has, in fact, claimed that evaluation and judgement are essential to any scientific endeavour. Without 'intradisciplinary evaluation' one could not distinguish between science and pseudo-science, let alone good science from bad.

appropriate use of a scientific method, but also by external ones, concerning its expected consequences and social relevance. Consequently, in line with the 'affirmative postmodernist' stand of Mabry (2002, 143, 152), it is seen as impossible to make a sharp distinction between 'scientific' evaluation and 'political' assessment; the evaluator as a social scientist should be seen rather as a philosopher (the one who questions) than an expert (the one who concludes). Her task should therefore be to empower the weak, promote divergent interests, articulate multiple perspectives, and direct attention to multiple realities.

4.4.1.1 Role of international organisations as creators of norms and identities

The *role of an international organisation as evaluator* is the third, and the most specific aspect that merits attention in our study. Balancing between policy advocacy and the need to maintain a sufficient degree of integrity and independence in order to safeguard credibility is particularly delicate when evaluator is an international organisation. Usually, evaluation studies are carried out by independent researchers or consultants. In such cases, the evaluator can be considered as an autonomous, monolithic actor that certainly has her own preferences brought along by her educational, professional, and cultural background, but without the burden and the backing of an organisation, with its own institutionalised practices, explicitly stated objectives, rules of conduct, etc. These aspects may have crucial importance when evaluation is done by an international organisation.

Relatively little research has been conducted on the influence of international organisations that have no regulatory power, but must act through what can be described as 'idea-games', that is, bringing about policy reforms by diffusing ideas, shaping a certain 'repertoire' of reform, and providing domestic actors with arguments that allow them to legitimise their actions (Marcussen 2001). Over the past decade or so, in particular the institutionally oriented scholars of international relations have underlined the importance of such channels of influence (e.g. Finnemore and Sikkink 1998; Haas 1992; March and Olsen 1998).

The idea of international organisations as deliberate *creators of international norms and identities* can be useful in helping understand their role as evaluators (March and Olsen 1998, 961). These organisations "define problems, construct conceptions of causal knowledge, and create frames for action that integrate across nation states" (March and Olsen 1998, 963). Such expert organisations are not only decision-making institutions, but also, and above all, institutions for socialising

individuals, creating meaning, and for promoting specific concepts of the nature and role of the state, markets, human rights, and international organisations (ibid., 964).[99] They exert this influence through the mechanisms of *socialisation*, *imitation*, and *coercion* (Marcussen 2001).

Mechanisms of *socialisation*[100] include both positive and negative feedback – 'stick and carrot'.[101] In international politics, socialisation involves diplomatic praise or censure reinforced by material sanctions and incentives. International organisations also stimulate socialisation by pushing targeted actors to adopt new policies and laws, ratify treaties and by monitoring compliance with international standards. Reasons that make socialisation work and states comply with norms relate to their identities as members of an international society. (Finnemore and Sikkink 1998.)

Imitation is a learning mechanism that does not necessarily involve direct contact between social actors. Imitation is triggered by dissatisfaction or the discovery of new possibilities. In the midst of an overflow of information, international organisations could play a role in helping actors distinguish the relevant from the less relevant. (Marcussen 2001, 22.)

An international organisation without regulatory power can hardly exercise direct *coercion* over its member countries through binding sanctions, but is much better placed to indirectly coerce its member countries to comply with certain norms and principles, through regular peer review mechanisms, for instance (Marcussen 2001, 21-22).

These mechanisms of norm creation operate at the three phases of what Finnemore and Sikkink (1998, 895) call the *norm 'life-cycle'*. First of all, international organisations can function as *'norm entrepreneurs'*: along with NGOs, as parts of larger trans-national advocacy networks, or transnational 'epistemic communities', international organisations can advocate the adoption of certain international norms to a large extent filtered through the training and educational background of the officials (see also Haas 1992, 17).

---

[99] For instance, when an international organisation points at differences between 'leaders' and 'laggards' among countries in terms of their willingness and ability to adopt what is defined as a modern, democratic, and economically efficient public sector, they modify the reference groups, aspirations and behaviour of national bureaucrats.

[100] *Socialisation* has been defined by Ikenberry and Kupchan (1990, 289) as "a process of learning in which norms and ideals are transmitted from one party to another"; as "the induction of new members […] into the ways of behaviour that are preferred in a society" by Finnemore and Sikkink (1998, 902); and simply as "internalisation of norms as a result of interaction with other people" by Marcussen (2001, 6).

[101] Finnemore and Sikkink (1998) apply the concepts 'emulation', 'praise', and 'ridicule' to describe the mechanisms of socialisation.

The second stage, *'norm cascade'*, is a result of an active process of international socialisation intended to induce norm breakers to become norm followers. This is where the concept of *peer pressure* is at its most typical, operating through the above-mentioned mechanisms of socialisation, imitation and coercion. As "custodians of the seals of international approval and disapproval", international organisations can play a crucial role in establishing and assuring adherence to international norms, and *legitimising* or *delegitimising* state actions. No country wants to be seen as a 'rogue state' or 'laggard'. A government can also use external evaluations to acquire legitimacy for its policy. The needs of states to demonstrate, by complying with norms, that they have adapted to the social environment – that they belong to a group – tend to promote conformity.[102] Finnemore and Sikkink (1998, 903-904) suggest that one reason for conformity and socialisation is related to *esteem*: state leaders sometimes follow norms because they want others to think well of them, and they want to think well of themselves.[103]

At the highest level of norm lifecycle, the norm becomes *internalised*, i.e. it acquires a taken-for-granted quality and is no longer subject to broad public debate. Sources and mechanisms of internalisation include professions, iterated behaviour and habit, professions carrying different normative biases systematically instilled by their professional training. Confidence-building represents a typical example of the power of iterated behaviour and habit. Once trust becomes habitual, it becomes internalised and hence, in turn, effects change among the participants. Procedural changes that create new political processes can thus lead to gradual and inadvertent normative, ideational, and political convergence. (Finnemore and Sikkink 1998, 904-905.)

Table 4.3. International organisations as creators of norms and identities.

PHASE OF THE NORM LIFECYCLE	MECHANISMS OF NORM CREATION	REASONS FOR OPERATION
Norm entrepreneur	Socialisation	Legitimisation
Norm cascade	Imitation	Conformity
Norm internalisation	Coercion	Esteem

---

[102] Finnemore and Sikkink (1998, 906) evoke the hypothesis that states that are insecure about their international status or reputation, are the most eager to embrace new international norms. For instance, they argue that Japan's greater openness to endorsing international norms about refugees, as opposed to the relative reluctance of Germany stems from Japan's 'insecurity' with regard to its political leadership role.

[103] Social norms are partly sustained by feelings of embarrassment, guilt, anxiety, and shame that a person suffers at the prospect of violating them. Obviously, it is difficult to generalise to the state level findings from research done at the individual level, but at least to some extent even the individual level mechanisms play a role, since norm entrepreneurs frequently target their criticism against state leaders. (Finnemore and Sikkink 1998, 903-904.)

In conclusion, international organisations are purposive actors that have power by virtue of being able to *structure knowledge*. They can classify the world, create categories of actors and action, fix meanings in the social world, as well as articulate and diffuse new norms, principles and actors around the world (Reinalda 2001, 24). Of course, international organisations can also produce pathology (Barnett and Finnemore 1999, 719). For the purposes of the present study, it is essential to ask what is the role of evaluations carried out by an international organisation in the process of norm creation; are evaluations an efficient means for such an organisation to promote its interests and play a role in the 'idea-games'?

4.4.1.2 International organisations in networks of governance: policy networks and epistemic communities

Above, we have deal with international organisations in isolation of the context in which they operate. If such organisations promote the adoption of certain norms by the states, we must nevertheless ask where the norms come from. International organisations are not islands, but are influenced by their operating environment. Like policymaking at the national level, international evaluations take place within the frameworks of governance dominated by what have been called by different authors alternatively as advocacy coalitions, policy networks, interest groups, iron triangles, advocacy coalitions, or issue networks. The denomination used for such groups or networks has little significance for the present study, since the fundamental point is the same: "small networks of policy specialists congregate to discuss specific issues, set agendas, and formulate policy alternatives outside the formal bureaucratic channels, and they also serve as brokers for admitting new ideas into decision-making circles of bureaucrats and elected officials" (Haas 1992, 31).

The theory of epistemic communities is particularly relevant for explaining the processes of environmental policies involving international organisations that have no regulatory power over their member states, but rely on the power of persuasion and knowledge. Haas (1992, 3) defines an epistemic community as "a network of professionals with recognized expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within that domain or issue-area." Such a community does not need to consist only of natural scientists or of professionals applying the same methodology as natural scientists do, but is bound together by "their shared belief or faith in the verity and the applicability of particular forms of knowledge or specific truths"

(ibid.). An epistemic community may consist of professionals from a variety of disciplines and backgrounds, who have

1. a shared set of normative and principled beliefs;
2. shared causal beliefs that serve as the basis for describing the linkages between possible policy actions and desired outcomes;
3. shared notions of validity; and
4. a common policy enterprise (ibid.).

What distinguishes an epistemic community from various other groups is its shared set of causal and principled (analytic and normative) beliefs, a consensual knowledge base, and a common policy enterprise (common interests) (Haas 1992, 18). An epistemic community does not have to be large to have an impact; "what matters is that the members are respected within their own disciplines and have the ability to influence those within their immediate disciplines and extend their direct and indirect influence in an ever-widening pattern" (Adler and Haas 1992, 380).[104]

Epistemic communities function as creators of identities in the same manner – through socialisation – as international organisations. The ideas of a transnational epistemic community may take root in an international organisation or in various state bodies, and be further diffused to other states via the decision-makers who have been influenced by the ideas (Haas 1992, 17).

Epistemic communities play an important role in facilitating the action of international organisations in the processes of norm creation at its different phases as described above.[105] This process of diffusion, selection and persistence of political innovations is described by Adler and Haas (1992) as an evolutionary process, leading to collective learning, or "changes in the epistemological assumptions and interpretations that help frame and structure collective understanding and action". Underlining the political character of learning in this context, they argue that it is "crucial to know who learns what; whose learning gets translated into policy and why; whose learning gets a chance to affect other countries; and how political processes determine whose

---

[104] A hypothesis behind the theory of epistemic communities is that the greater the extent to which these "communities are mobilised and are able to gain influence in their respective nation-states, the greater is the likelihood that these nation-states will in turn exert power on behalf of the values and practices promoted by the epistemic community and will thus help in their international institutionalisation" (Adler and Haas 1992, 371-372).

[105] Epistemic communities can influence 1) *policy innovation*, by framing the range of political controversy surrounding an issue, defining state interests, and setting standards; 2) *policy selection*, the policy-makers seeking advice from

interpretations of reality are more viable in a particular historical context" (ibid., 386). Epistemic communities can therefore be seen as catalysers in the processes of policy development and norm creation, thus making them crucial actors conditioning the influence of policy evaluation.

4.4.1.3 Organisations vs. states: principal-agent dilemma

If international organisations have power through the 'idea games', norm creation, and peer pressure, one must ask whose interests and ideas they actually promote. The first perspective is to look at the relationships between the member governments and the organisation's secretariat, the latter having in principle simply the duty of executing the orders of the member governments.

From the *Public Choice* perspective, Frey (1984) argues that international bureaucracies have greater room for discretionary action than national ones, because controlling them is difficult: their 'output' is undefined and hard to measure, and national governments would run into conflict with other governments if they tried to interfere with the workings of international organisations. In the absence of effective control, none of the layers of the organisation has real incentive to work toward the 'official product'. This leads to inefficiency, profusion of red tape, and international bureaucracies expanding quite independent of the tasks to be performed, because expanding budgets and winning the intra-organisational power battles become dominant objectives. (Frey 1984, 221.)

Verbeek (2001) suggests that the international organisations operate in the context of confrontation between the principal (the nation states), who wants to control the behaviour of the agent (international organisation), which in turn tries to carve out as much leeway as possible to have things the way it wants. A key question is to what degree states succeed in keeping control over the structures they create. Principal-agent analysis suggests that an agent has substantial margin of manoeuvre mainly by virtue of asymmetrical distribution of information in its favour. This may lead to *'shirking'* ('bureaucratic shift'), or *'slippage'*. The former denotes situations where agents pursue interests of their own. 'Slippage', in turn, refers to incentives that may follow from the specific manner in which the principal delegates the tasks to the agent. To curb the potential for agent's freedom, principals develop various oversight procedures or construct situations in which one agent is set against the other. These procedures entail, however, additional costs to the principal

---

epistemic communities in order to get support for their own ideas and build coalitions; and 3) *policy persistence*, by helping institutionalise new ideas which thus gain the status of orthodoxy (Adler and Haas 1992).

and their application therefore depends, among other things, on the cost-benefit ratio. (Verbeek 2001, 7.)

Several factors favour the independence of the agents. In a situation with several principals the agent may exploit differences between them – a possibility that is compounded if decision-making rules (e.g. consensus requirement among principals) allow the agent to take advantage of the situation. While the unanimity rule may also lead to a stalemate unfavourable to the agent, often it perpetuates the status quo, which frequently plays to the advantage of the agent. It can therefore be expected that an international organisation's room for manoeuvre is negatively correlated with principals' *economic power*, and positively correlated with the *number of principals* an agent has to deal with, whereas the *information advantage* of the agent obviously increases its margin of manoeuvre (Verbeek 2001, 7, 9.) The tools available for international organisations include the use of technical knowledge and juridical language in order to define issues and alternative solutions, and the building of coalitions with domestic and trans-national actors. (Reinalda 2001, 10.)

The principal-agent framework displays several advantages: simplicity, capacity to provide clues as to why international organisations seek margin of manoeuvre and as to the sources of influence available to (organisational units within) international organisations and their member states. Nevertheless, the premise that agents seek expansion of their missions or budgets, or both, is too limited, and needs to be complemented by more constructivistic approaches. Principal-agent approaches fail to explain what actually constitutes the interest of an international organisation (Barnett and Finnemore 1999).[106]

4.4.1.4 Inside an international organisation

What constitutes the interest of an international organisation is not as obvious as it might seem at the first sight. While the assumption inherent in Public Choice and 'bureaucratic politics' (e.g. Kozak and Keagle 1988) literature that an agent will maximise its tasks, budget, and personnel may hold true when the survival of the organisation is at stake, in most cases the motives of an international organisation are more complex: there are several players focusing not on a single strategic issue but also on many diverse intraorganisational problems; and players do not act in terms of a consistent set of strategic objectives, but according to various conceptions of national,

---

[106] Moreover, Finnemore (1993, 566) argues that international organisations can, in some cases at least, be seen as principals, rather than agents (see also Reinalda 2001, 9).

organisational, and personal goals. Political leaders are joined by officials who occupy positions on the top of major organisations to form a circle of central players, central in relation to the particular decision or outcome the analyst seeks to explain. (Stern and Verbeek 1998; Verbeek 1998; Reinalda 2001, 18.)[107]

It is here that the theory of epistemic communities becomes relevant again: instead of being 'monoliths' with clear objectives cutting through the organisational hierarchy from the top to the bottom, internalised by each and every individual and working unit, international organisations are often split into more or less rivalling fractions. The lines of fracture often follow the boundaries between different 'epistemic communities', even though the desire to defend and increase budgets, personnel and tasks – as postulated by the bureaucratic politics literature – and the attempts to forge strategic alliances with key member states are significant as well.

4.4.1.5 Summary: international organisation as evaluator

In this study, we are interested in the role of an international organisation as an *external* evaluator. An international organisation faces – like any evaluator – a fundamental choice between two extremes: to advocate a predetermined policy or the interests of a particular group, or to adopt a detached and 'objective' view. This choice is complicated by the need of the organisation to take into account its position in the international governance structures, the internal dynamics of the organisation and its relations with its member states. This study focuses on the influence that an international organisation without direct regulatory power exercises as a creator of norms and identities in 'idea-games', in particular through the processes of socialisation, imitation, and 'soft coercion'. Concretely, international organisations exert their influence as key players in 'epistemic communities'. While the motivations relating to the maximisation of tasks and budgets explain a part of the international organisations' behaviour, epistemic factors such as the educational and professional background of the international bureaucrats are probably equally important in explaining the behaviour in the context of 'idea games'. Finally, international organisations are not monoliths, but the different epistemic communities cut across them.

---

[107] Moreover, not every individual bureaucrat is loyal to his/her organisation, but often experiences *role conflicts* in having to choose between loyalty towards the principal (home country government) and the actor (the international organisation). Unlike the Public Choice theory assumes, socialisation processes and the time needed to consult home government often push individuals to adopt the international organisation's perspective (Finnemore 1993, 566; Reinalda 2001, 19).

4.4.2 Through networks from international to the national level: actors and 'repertoires'

The rational-technocratic view, which often underpins experimental evaluation, tends to attribute the absence of evaluation use to the lack of rationality of the potential users. These should listen to the evaluators' advice that is based on the available scientific evidence, and implement policies supposed to correct the identified problems. If they do not, they are considered "devious, inadequate, ignorant or some combination of all three" (Pollit 1998, 215). Those stressing the political character of evaluation, by contrast, consider that users:

- should not be expected to act rationally, but instead in function of political contingencies,
- need to make decisions with extremely short delays, under constant time-pressure, and
- need to satisfy multiple constituencies such as fellow ministers, party colleagues, key pressure groups, other governments, the media, etc. (ibid., 216-217).

From the perspective of government accountability towards citizens, policies should be evaluated against the assumption that policymakers do in fact act rationally, attempting to reach the official policy goals. However, to understand the 'whys' of policy-making processes, one has to endorse the idea of multiple types of rationalities – the influence of factors such as political expediency, routine, educational background, and ethics, in a context where decision-makers are torn between multiple loyalties. Before looking at the motivations of actors more closely, let us first define which are the main actor groups in a typical evaluation situation.

4.4.2.1 Main categories of actors

An evaluation usually involves a number of different actor categories, some of which have commissioned the evaluation, while others can be characterised as 'intended beneficiaries', or policy target groups. Yet others may be officials in charge of executing the policies, some can be characterised as stakeholders, whereas the bulk of the costs and non-intended effects of the policies may fall upon yet another category of actors. Identifying the different groups or communities involved or having an interest in the issues at hand is one of the essential first steps of any evaluation (e.g. OECD 1999, 6; Montague 2004, 11).

It is useful to look at three spheres of potential user groups. The traditional evaluation models saw the *commissioning authority* as the self-evident and the only user of evaluations, with other

stakeholders serving as sources of information at the most. Pragmatic evaluation brought attention to the second category, the *'intended users'*, which may include a wider range of stakeholders. Finally, the transformation of the policy environment, including the increasing importance of the Internet, has forced evaluators to consider the fate of evaluations in the *public arena*. The slogan 'intended use by the intended user' (Patton 1997a) loses its feasibility, because evaluators have no means of controlling the use of evaluations in the context of multiple media, audiences, and streams and forms of data, anonymous relations between actors, and long chains of utilisation.

Active citizens, community opinion leaders, etc. function not only as vectors of evaluation influence, but can also play an active role in giving a more positive image of evaluated social programmes (Weiss 1998, 28-29). Likewise, NGOs, for example, may serve as a critical voice in relation to the evaluations or the evaluated policies, and may contribute to 'peer pressure'.

The media play a key role as actors in the 'public sphere', in 'agenda setting' and general dissemination of evaluation results (e.g. Weiss 1987; Henry and Mark 2003). Whether they adopt a critical stance towards evaluations or function simply as disseminators largely depends on the general policy context, notably on the reputation and the perceived independence of the evaluator. The greater the perceived independence of the evaluator in relation to public authorities, the less critical the media tend to be towards the evaluations.[108] The media power largely relies on agenda-setting: it has less influence on people's opinions on particular issues than on what issues the public think are most important (e.g. McCombs and Shaw 1972). In raising the salience of certain issues, the media influence the standards by which the public estimates the performance of the government or public officials, through 'priming' (Ivengar et al. 1982; Krosnick and Kinder 1989; Krosnick and Brannon 1993).[109] If the media pay, at any specific moment, particular attention to one environmental problem – climate change, for example – in judging the government's environmental policy, the citizens are more likely to give more weight to the government's record in combating climate change than to issues that receive less media attention.

---

[108] For instance, Gunvaldsen and Karlsen (1999, 462-463) discovered that the media tended to simply reiterate the results and conclusions of the evaluations carried out by the Norwegian Office of the Auditor General as if these were the objective truth, instead of seeking to take a critical distance to the findings. The authors identify two probable reasons for such lack of criticism. Firstly, the media did not perceive the Office of the Auditor General as a part of the operative power apparatus, but rather as an independent agent controlling the exercise of power. Secondly, the journalists did not have access to all the sources of information they would have needed to critically examine the findings of the audits conducted by the Office.

[109] Priming is based on the ideas introduced in chapter 2 concerning rationality and motives behind individual behaviour. When making judgements or decisions, people seldom 'optimise', in the sense of taking into account the entire array of available relevant evidence. Rather, they rely on limited subsets of the available information pool so as to make satisfactory judgements without expending a great deal of effort (Krosnick and Brannon 1993, 963).

4.4.2.2 Expectations of actors: credibility, legitimacy, and salience of evaluations

The influence of evaluations to a great extent depends on the expectations, interests, informational backgrounds and worldviews of the actors involved. Research suggests that international scientific assessments are most influential when they are *credible* as to their scientific methods, *salient* to the potential users, and *legitimate* in the way the assessment is designed (Clark and Dickson 1999; Eckley 2001). *Credibility* of an evaluation can be defined as its scientific and technical believability to a defined user. It can be based either on the process through which the information for the evaluation was created or on the credentials or other characteristics of the evaluator. *Salience* or relevance refers to the "ability of an evaluation to address the particular concerns of a user". *Legitimacy* means the political acceptability or perceived fairness of an evaluation to a user. Not all of these conditions can be simultaneously fully satisfied. Moreover, the importance accorded to each criterion varies according to the situation. For instance, the more controversial the policy issue, the more credibility tends to be appreciated. The criteria of credibility and legitimacy are also influenced by the relationships between the actors. For example, an evaluation is likely to be more influential if brought to the public debate by a group enjoying high credibility among the other actors. (Eckley 2001.)

What different actors consider credible, salient and legitimate depends on their interests and expectations. Barbier (1999, 377) has classified actors' expectations into three categories: *expectations of cognitive nature* include the need for information and knowledge concerning the way a policy is being implemented, its outputs, results, outcomes, impacts, etc. In reality, most actors' expectations are, however, either *strategic* or *political*, the former relating to actors' desires to preserve their respective powers, and the latter to the political activity as commonly understood in the sense of the political actors' legitimisation efforts (see also sections 4.3.5 and 4.3.6 above). Cognitive and strategic expectations come together when a policymaker commissions an evaluation in order to obtain 'counter-information' to balance information fed by other sources, which she judges less trustworthy (e.g. by a rival political party). Different actors are also interested in different criteria. High-level political decision-makers tend to take an interest in economic effectiveness, local level officials seek information on policy outputs, and non-governmental stakeholders see evaluations primarily as an avenue for feedback and involvement (Wimbush and Watson 2000, 304).

Sometimes an evaluation is in reality not directly commissioned by any actor, but is instead being imposed upon actors from outside. Often this emanates from a supervisory body within the same administration, but increasingly also from a supranational organisation, such as the EU, the IMF, the World Bank, or the OECD.[110] Coupled with the multiplication of reporting obligations under the various international conventions, the increase in the frequency of evaluations has engendered 'evaluation fatigue', where any new evaluation is perceived as a burden, a mere obligation that needs to be fulfilled. Such a situation tends to enhance dissimulation rather than transparency, as actors do their best to hide negative aspects in the policies. (e.g. Perrin 1998; 2002b; Baron 1999, 84; Vedung 2001; Taut and Brauns 2003.)

4.4.2.3 Repertoires

A concept that brings together actors' expectations, belief systems (Sabatier 1987), mental models and the more operational codes and practices of organisations is 'repertoire', defined by van der Meer (1999, 390) as "stabilised ways of thinking and acting (on the individual level) or stabilised codes, operations and technology (on other levels)".[111] The concept resembles the term 'institutions' as employed by institutional economists, recognising the importance of the multiple layers of human consciousness in shaping individual preferences and human behaviour. Because of differences in their histories, experiences, and positions in the relations of power, actors have different repertoires, which they use in the process of sense-making and construction of behaviour. The impacts of an evaluation are therefore not primarily determined by the logic of the evaluator, but by the repertoire-based interpretations and (re)actions of the agents involved. Contrary to the assumptions implicit in a lot of the research on evaluation use, evaluation results are far from unequivocal and have widely varying meanings to various actors involved. Moreover, sense-making and organisational patterns interact and mutually influence each other. (van der Meer 1999, 390-392.)

Since actors' repertoires differ, their behaviour is rarely self-evident for others. Actors are constantly faced with ambiguity, which in turn triggers change in their repertoires.[112] Ambiguity stems not only from outside of one's own organisation or group, but also from within, because

---

[110] Of course, evaluations are not officially imposed upon countries, given that for instance carrying out of an OECD review requires an official invitation by the country's government. For most of the civil servants involved, however, these reviews are 'imposed upon' from above by their superiors.

[111] The concept is close to 'frame' (Giddens 1984, 87; see also Valve 1999) or 'ideological orientation' (Söderbaum 2001).

individuals are generally involved in multiple, partly related social contexts. A minister functions in political circles in a repertoire dominated by party ideologies, parliamentary majorities and accountability, but on the other hand, he is also an administrator and therefore focused on economy and effectiveness, or a partner in consultations with societal groups, in which cooperation, harmony and support are central elements of the repertoires. The dominant repertoire in a group, organisation or a network, is never fully shared by its constituent individuals, subgroups, or departments, but different individuals or subgroups are included in different repertoires to a greater or lesser extent.

Changes in the repertoires do not necessarily arise from the evaluation, but from meaningful reactions of third parties, whose repertoires overlap with or are linked to those of the main actor. For example, a governmental unit can consider the results of an evaluation irrelevant (e.g. because they come from 'another world'), yet react to them because of attention given to the evaluation in the press or in Parliament. Evaluation judgements that fit in the repertoire of the unit evaluated will be taken as supportive and reinforce the repertoire, irrespective of how such assessments are used in the arguments of the evaluator. Since evaluated entities host a number of overlapping and competing repertoires, the intensity and direction of the impacts of evaluations depend on the extent to which they link to the repertoires of certain individuals or factions in the organisation. Understanding the repertoires of key actors not only helps the evaluator to produce evaluations that are influential, i.e. that link with the repertoires of one or more actors, but might also help initiate explicit discussion on central underlying assumptions and value orientations. On the other hand, evaluators should be aware of their own repertoires, and examine their own positions from the point of view of other actors. (van der Meer 1999, 390-392, 402-404.)

Evaluation contexts are typically impregnated with different forms of power based on information, prestige and other symbolic attributes. What becomes crucial then is the capability of actors to harness the opportunities to exercise symbolic power – through persuasion, legitimisation, defence, etc. Whose repertoires, belief systems and ideologies are dominant in society determines in large part who exercises cultural power and who is being 'empowered' by an evaluation.

4.4.2.4 Summary on actors in evaluation

Experience from policymaking in general, and evaluation in particular, has shown that the 'users' of evaluations do not adhere to a single conception of rationality, but are guided by a range of

---

[112] Van der Meer (1999, 390-392) calls this process of change learning.

cognitive, strategic and political motivations. The political environment, within which evaluations are carried out, is becoming increasingly complex, making it virtually impossible to control the fate of evaluation findings in public arenas. Crucial for the influence of evaluations are actors' repertoires, i.e. their various expectations, interests, belief systems, and worldviews, combined with the way these have been institutionalised in the codes, operations and technologies of the collective entities within which the individual actors operate. A key requirement for an evaluation to promote learning is that it links with the dominant repertoires and stimulates the development of new links between different repertoires or repertoire elements. Analysing the repertoires of different epistemic communities is necessary for understanding the role of different actors in evaluations carried out by international organisations.

4.4.3 Institutional context

Evaluations do not take place in a political vacuum, but a number of contextual factors influence the way in which the evaluation findings and the experiences from evaluation processes enter the policy system. Thus, a recurrent theme in evaluation literature is the exhortation to evaluators to obtain a better understanding of the overall political context (e.g. Pawson and Tilley 1997; Julnes et al. 1998; Pollit 1998; Barnes et al. 2003). In a coevolutionary framework, taking into account complexity requires that context be recognised as a part of the open system within which the evaluated policies are operating, instead of being external to the system. The context itself is subject to change as a result of actions or activities beyond the scope of the policies, but also in response to the intended or unintended consequences of policy implementation. (Barnes et al. 2003, 269.) *Institutions* such as history, culture, policy style, tradition, and standard operating practices constitute the context, which delimits the range of issues that can be addressed in a particular evaluation. 'Path dependence' created by previous decisions further limits the scope for possible action. In this section, we shall focus on the more 'structural' factors related to the general political environment within which an evaluation is carried out. In producing its effects on policy, the context interacts with the characteristics of actors and participants. (Weiss 1999, 478.)

4.4.3.1 Fundamental long-term framework conditions

Sabatier (1987) has divided the external factors within the broader political system affecting policy change into two categories: on the one hand, the relatively stable parameters that usually only change over a period of several decades, and, on the other, aspects that are susceptible to fluctuate

significantly over the course of a few years and thus serve as major stimuli to policy change. Among the former are:

1. ***The basic attributes of the problem area*** or of the 'good', such as excludability[113], or the problem's susceptibility to quantitative measurement. In evaluation research, this is closely related to evaluation design, and will therefore be further elaborated in the next section.

2. Basic ***distribution of natural resources*** is obviously relevant especially in environmental policies. In Finland, problems related to forestry and freshwater quality are likely to be at the top of the policy agenda. The energy policy has in the U.S. taken a significantly different path than in France, largely because the latter has no domestic oil resources.

3. ***Fundamental cultural values and social structure***. While the terms 'repertoire' and 'belief system' referred to groups of individuals, or organisations, fundamental cultural values and social structure encompasses a broad area of issues such as the 'political culture',[114] and 'policy styles'[115] at the level of the society as a whole. For example, the attitudes towards environmental taxes tend to be a much more favourable in Europe than in the U.S., largely because of the historically shaped social norms. Weiss (1999, 480-482) has argued that factors enhancing evaluation use include a high degree of *functional specialisation* in the political and administrative system;[116] a high proportion of *social scientists in the administrative system*; a thriving *social science community*, especially if committed to empirical work instead of mere theory-building; and a 'climate of rationality'. This last point merits a comment, as Gambetta (1998) has suggested that societies based on 'analytical' assumptions about knowledge constitute a more fertile ground for deliberative processes than those with 'indexical' beliefs, prevalent in the political and intellectual elites of many 'Latin' countries. The former entails the assumption that knowledge is a result of good reasoning, empirical verification and hard work, whereas the indexical conception is based on a holistic view: ignorance in a certain area is taken as a sign of ignorance of the

---

[113] A good is excludable, if a way can be imagined, given the current level of technology, to prevent somebody from consuming it.

[114] Political culture can be defined as the totality of "the attitudes, beliefs, emotions and values of the society that relate to the political system and to political issues" (Ball and Peters 2000, in Sairinen 2000, 47).

[115] Van Waarden (1995, cited in Sairinen 2000, 51) has identified six dimensions defining a policy style: (1) market-oriented vs. state-oriented; (2) active vs. reactive; (3) comprehensive vs. fragmented; (4) adversarial vs. consensual vs. paternalistic; (5) legalistic vs. pragmatic; and (6) formal vs. informal relations between public and private actors.

[116] Yet, in a highly specialised political system, a cross-sectoral evaluation may fall between different sectoral competencies, accountabilities not having been clearly defined for the evaluated policy. Furthermore, research seems to indicate that the longer the policymaker has worked in the field, the less he is interested in keeping up with evaluation information.

whole.[117] This distinction is in the following referred as one between different *discussion cultures*. To these aspects one could add the independence of *the media*, as well as the strength and the status of the NGOs and the *civil society* in general, which tend to promote critical discussion in the society.

4. ***Basic legal structure***. This includes the role of the courts, the degree of transparency of the political and administrative system, the degree of concentration of decision-making powers, etc. Weiss (1999, 480) posits that democracy, a competitive, multi-party political system, and a high degree of decentralisation of policymaking encourage the use of research and evaluation for policy-making, given the greater openness of such systems to evidence and argument. In the context of comparative international environmental policy evaluation, the level of development of policies, the degree of institutionalisation of environmental concerns into administrative structures, and the extent of policy integration in the environmental field are likely to be crucial.

4.4.3.2 More ephemeral, short-term background factors

The political background factors that can vary substantially over the course of a few years or a decade include:

1. Changes in socio-economic conditions and technology (e.g. the oil crises, innovations in new information technology or pollution abatement technology).
2. Changes in systemic governing coalitions. Left and right wing governments tend to have different policy approaches in environmental matters, for instance.
3. Policy decisions and impacts from other subsystems. For example, the increasing harmonisation of European economic and other policies have brought about thorough changes, such as greater reliance on free market principles.

While evaluators can hardly influence either the long-term structural framework conditions or the ephemeral, short-term factors, awareness of these aspects is essential. Moreover, understanding

---

[117] Gambetta (1998) argues that the indexical beliefs are globally more widespread than the analytical ones. People having such beliefs are prone to have strong opinions on everything, right at the outset of a discussion. In the societies with an indexical set of beliefs, dubbed by Gambetta as cultures of 'Claro!', discussion culture is based on a desire to win an argument rather than on being willing to listen and learn from others; questions of pride often take precedence over reason; 'intellectuals' assumed to be knowledgeable in just about all matters of life are highly respected rather than viewed with suspicion; and rhetoric rather than valid arguments tend to dominate, therefore easily leading to bargaining instead of deliberation.

particularly the short-term situational policy context will help the evaluators to exploit the political 'windows of opportunity' that may open new possibilities to enhance evaluation influence. Since the focus of the present study is on international peer reviews, and given the increasing internationalisation of environmental policies, the political framework conditions must be placed in an international context. Trends in environmental policy are largely set by 'forerunner' countries – international organisations and conventions playing a crucial catalysing role.

4.4.4 Evaluation design

Finally, let us look at the obvious, i.e. the impact of evaluation design on influence. Often this is taken as the primary or even the only aspect conditioning evaluation influence. However, the contextual factors, 'repertoires' of actors and power relations are often more significant either directly or through their impact on evaluation design. Aspects concerning evaluation design addressed here include the quality of the report and its recommendations; the character of issues addressed in the evaluation; the scope, criteria and purpose of evaluation; organisation of the evaluation process; and participation of stakeholders.

4.4.4.1 Quality of the report and the recommendations

Although the quality of the report mostly influences the degree of direct evaluation use, it is plausible to assume that even the indirect effects depend on the final report meeting certain basic quality criteria. An evaluation may fail to have an influence on policy-making because of *lacking relevance* and other deficiencies in the evaluation report itself. The evaluation may have overlooked important issues, such as side effects, the results may be internally contradictory or unclear, the methods of data collection may be inappropriate, or the evaluators may have failed to come up with concrete policy recommendations.[118] *The language* of the report must be calibrated according to the potential users. Too technical language using expert 'jargon' compromises legitimacy and salience. (Vedung 1991; van der Meer 1999; Weiss 1999, 479; Eckley 2001.)

---

[118] The inclusion of *recommendations* in the evaluation report is a double-edged sword, however. On the one hand, they increase salience, but on the other, policymakers can perceive them as an illegitimate intrusion to their 'territory' (Eckley 2001). Some constructivist evaluators (e.g. Kushner 2002a) consider that even an evaluation report is unnecessary and can be detrimental to the influence of evaluation. However, these critics seem to overlook the fact that an evaluation report enhances social learning through argumentative processes (Patton 1998, 226-227; Baron 1999; 233-236; Valovirta 2002, 78).

The credibility of the *source of evaluation data* plays a vital role. Data coming from a source internal to the evaluated entity tends to increase legitimacy and salience in the eyes of the evaluated entity, but decrease the credibility in the eyes of other groups (e.g. van der Meer 1999, 387-388).

Finally, the evaluators must have the necessary expertise and competence to carry out a high quality evaluation. Mere expertise usually is not enough, and symbolic issues may be at least as important, depending on the context. The image that the evaluation team gives to the evaluated entity and the individual evaluators' institutional allegiance, political status, etc. may be decisive for the credibility and legitimacy of an evaluation.

4.4.4.2 Character of issues

A reference was made above to the 'basic attributes of the problem area' as a factor affecting evaluation use. This reasoning can be expanded to considerations relating to the degree of disagreement on and the technicality of the issues selected for evaluation, as well as the 'fit' between policies and recommendations of the evaluation. In general, findings that are relatively non-controversial and do not deviate too much from the organisation's dominant 'repertoire' are more likely to be adopted than those that require large-scale overhaul of thinking or organisational practices (e.g. Weiss 1998, 23-24). Findings or recommendations concerning politically very sensitive and controversial issues stand little chance of being taken up by policy actors (e.g. Weiss 1999, 482; Leknes 2001; Beyeler 2002). However, evaluations can help build consensus in controversial cases, if the evaluator is considered as neutral and independent (Schubert et al. 2000). Evaluation can be expected to influence policies much more directly in highly technical and 'professional' issues than in 'political' issues involving a large number of stakeholders with diverging views. Likewise, the greater the extent to which the management of the issue has been institutionalised into the decision-making processes, the higher the prospects for evaluations having an impact. (Leknes 2001.)

It has been suggested that the *'misfit'* between the policy recommendations on the one hand, and the domestic policies and institutions on the other must be sufficient to stimulate action, but not wide enough to give rise to outright rejection (Beyeler 2002; Radaelli 2000). A policy recommendation or an evaluation finding is likely to be influential if it is radical enough to generate debate and

interest, but sufficiently anchored in the prevailing context so as not to be discarded as unrealistic. Such a 'fit' of course also depends on the 'repertoires' of the actors involved.[119]

4.4.4.3 Scope and focus of evaluation

The range of issues covered in an evaluation has an ambiguous role. A narrow focus enhances credibility, because the evaluators are likely to be experts in the relevant policy domain, and have the time to carry out an in-depth analysis using scientifically reliable and tested methods. A narrow focus would also better ensure that sufficient resources are invested in answering the most important questions, which might be forgotten in a broad, comprehensive evaluation (Lang 2001, 12). The drawbacks are reduced salience to decision-makers if evaluations only focus on a narrow range of issues, and the loss of legitimacy, because actors from other fields are unlikely to consider legitimate an evaluation focusing exclusively on one or few disciplinary perspectives. (Eckley 2001.)

Obviously, the *length* of the report must be appropriate in view of the potential use and users. In general, the shorter the report, the more likely it is to be salient and read by decision-makers, while too concise reports, on the other hand, may not be perceived as credible. (Vedung 1991; van der Meer 1999; Eckley 2001.)

4.4.4.4 Criteria for policy success

On a general level, evaluations can be distinguished in function of the degree to which they consider the goals of the policy as the main criterion of success – the *goal-achievement model* – as opposed to relying on evaluative criteria external to the evaluated policy. Early writings on evaluation in fact defined evaluation as an examination of the extent to which programmes achieve their stated goals (Peled and Spiro 1998, 457). Subsequent criticisms against the pure goal-achievement model have included the contention that programme goals must be deliberately vague, so as to accommodate multiple interests (Cronbach et al. 1980); warnings against the 'goal trap' into which evaluators and their clients might fall if pursuing goals which were never actually intended to be implemented; 'tunnel vision' and the resulting neglect of important but non-

---

[119] Beyeler (2002) has found that the OECD Economic Surveys do not shape social welfare policies in the OECD member countries in a uniform manner, on the one hand because the recommendations vary from one country to another, and on the other because the impact of the OECD ideas depends on the fit of national institutions, policies, and other external pressures.

anticipated ancillary benefits and harmful effects of the programme (e.g. Peled and Spiro 1998; Perrin 1998; 2002a; Leeuw 2002), as well as bias in favour of the interests of powerful groups (Scriven 1991). The latter prompted the idea of *goal-free evaluation*, focusing on all intended and non-intended impacts of a programme. In practice, however, goal-free evaluation has hardly been tested empirically (Chen 1990).

To date, most evaluations continue to take programme or policy goals as their main point of departure, but the goal-achievement model is becoming more sophisticated. Hence, the goals are no longer taken as exclusive standards of success, but attention is paid to a host of other criteria as well, including those listed in box 4.2.

Box 4.2. Possible criteria for judging policy success
(adapted from Roman and Vedung 2000, 21-27; Ahonen 2001, 111; Vedung 2001).

- *Relevance* or *problem-solving effectiveness* measures the relation of outcomes to the needs that should be satisfied or problems to be solved.
- *Impact* measures the degree to which the observed results stem from the policy being evaluated.
- *Sustainability or persistence* refers to the extent to which the results are expected to last even beyond the policy timeframe.
- *Flexibility* refers to the capacity of the policy (instrument) to cope with changing conditions.
- *Predictability* denotes the extent to which different actors can foresee the administration, outputs and outcomes of the policy measures.
- *Goal-attainment effectiveness* reflects the traditional goal-achievement approach, measuring the extent to which the (stated or non-stated) goals are attained over time.
- *Efficiency* can be measured in terms of the *cost-benefit* ratio, benefits being measured either in monetary or other terms. Another dimension of efficiency is *cost-effectiveness*, in which the policy objectives are taken as given, and the evaluation concerns the most economic means of achieving the objectives.
- *Addressee effectiveness* measures the changes in the knowledge, attitudes, or behaviour of the groups at which the policy was targeted.
- *Administrative effectiveness* or *process effectiveness* turns attention to the implementation phase, asking whether the administrative system, at its different levels, is optimised to its

task, whether it has adequate resources available, etc.

- *Constitutive effectiveness* resides in the regime's operation as an institutional arena where actors meet and exchange ideas and information, thus enhancing problem solving in other areas as well.

- *Legitimacy* refers to the degree to which the policy is acceptable to the different stakeholders.

- *Transparency* requires that the outputs, outcomes and the implementation processes are observable to outsiders.

- *Equity* relates to the equitable distribution of benefits, negative impacts, and costs of a programme or a policy, as well as to the equality of opportunities to participate in and influence the policy processes.

Yet other proposed criteria include effects on *empowerment*, *subsidiarity*, *partnership*, *coherence* with other policies, *congruence* with general policy principles, and '*good governance'* (Ahonen 2001, 113).

Perhaps the main lesson to be drawn from the long list of potential evaluation criteria is the need to look beyond the traditional goal-achievement model, and adjust the evaluation criteria according to the context. There is no universally accepted set of criteria, but a large number of partially overlapping and competing elements of policy success. For instance, Roman and Vedung (2000) advocate the use of the term effectiveness as the overarching concept for measuring policy performance, covering all kinds of impacts. However, it is hardly feasible to translate issues such as legitimacy, transparency and equity into the terms of efficiency. Reducing all criteria to effectiveness tends to overestimate instrumental rationality and the function of evaluation as straightforward measurement of effectiveness. Likewise, labelling certain policy effects as 'soft', as done for instance by the OECD (1997d), gives the impression of a hierarchy of effects between the 'hard', concrete, 'real' impacts, and 'soft', secondary effects. However, the 'soft' impacts – defined by the OECD (1997d, 94; Jones 2004, 5) mainly as changes in attitudes and awareness – may often be difficult to measure, but also the most important or even the only impacts generated by a policy.

Even though the pure goal-achievement model has been criticised, focusing on goals has another virtue, namely the identification of program or policy goals. This can be seen as an evaluation end in and by itself, not only as a starting point for the measurement of effectiveness, since it can

enhance mutual, informed dialogue and deliberation among stakeholders on the objectives to be pursued. Various participants may not only perceive differently the causal relations between policy means and ends, i.e. have different intervention theories, but they may also pursue divergent goals. Furthermore, declared goals may differ from those actually held by an organisation, goals may be instruments utilised in power struggles, and they may constantly evolve and change. (Peled and Spiro 1998.)

4.4.4.5 Purpose of evaluation: the juxtaposition between learning and accountability

The rising popularity of performance measurement has increased the tension between, on the one hand, the call for greater accountability through evaluations and, on the other, the purposes of 'enlightenment', development, and learning. The main risks involved in increased use of performance measurement can be summarised as follows (Perrin 1998; 2002b; Blalock 1999, 118-119; Davies 1999, 154-155; Greene 1999; Lang 2001, 19-21; Feller 2002, 439-447; Leeuw 2002; Taylor Fitz-Gibbon 2002, 140):

1. *Diversion of funds.* Money that could be used for programme improvement, goes to various types of reporting and monitoring activities.

2. *Goal displacement.* Since performance measurement focuses on measuring quantifiable results, it may divert attention from the actual outcomes to outputs and results. Managers may, for instance, focus on 'easy' cases in order to ensure that quantitative indicators show good results. They may also set less ambitious goals, merely to ensure compliance. The chosen indicators may therefore not measure what they were intended to measure.[120] Finally, misleading aggregate indicators often disguise differences between subgroups.

3. Performance indicators are often technically complex and therefore *not well understood by the potential users*. Causality between activities and outcomes is often taken for granted by the users of performance information, even when such causality has not been demonstrated. As a result, programme personnel may be held accountable for outcomes, which they cannot influence.

4. *Poor relevance to policy-making.* The level of ambition of policy objectives is not easy to take into account, and unintended consequences as well as the continuous changes of objectives are often ignored. Since performance measurement disregards the 'why' and

---

[120] Often indicators are, indeed, selected on the basis of the availability of selected forms of data rather than on the relevance of the data for policies (e.g. Lavoux 2003).

'how' questions, it is often of scarce utility for decision-making and resource allocation. Moreover, tracking past performance and monitoring dimensions of current activities has limited value for prospective decisions.

5. *Misplaced emphasis on control*. The accountability-based performance systems rely on an outdated top-down command-and-control approach to management, which has been proven ineffective, and engenders organisational paralysis, 'tunnel vision' and 'sub-optimisation'. The emphasis on control tends to discourage responsibility, engender resistance and risk aversion instead of fostering innovation, creativity and motivating people to do their best. Paradoxically, people then pay less attention to outcomes, as they seek to justify their work instead of improving it. To meet the performance objectives, entities may be tempted to dissimulate and distort data, or even lie and cheat. If evaluations are felt as nothing more than an obligation coming from above, their expected benefits are likely to remain largely untapped.

6. Performance measures are an *authoritative*, closed way of representing the quality of programmes and therefore *inadequate for advancing the ideals of deliberative democracy*. Programme quality is too complex to be reduced to end-points, because it is irreducibly pluralistic, and is not available as an objective, permanent measure of a fixed reality. The use of administrative rhetoric associated with performance management in an area with a non-administrative tradition (e.g. science and technology policy), may collide with the institutional, functional, and policy/political constraints specific to this area.

7. Performance measurement tends to *legitimise and reinforce the prevailing power structures*. For instance, the use of citation measures as an indicator in science serves to reinforce the existing perceptions of a hierarchy among scientific journals.

On the other hand, and partly in response to criticisms, a number of conditions have been identified for successful performance management systems, including (Davies 1999; Greene 1999, 162-163; Winston 1999; OECD 2003j, 19-20):

1. Top leadership support, including sufficient personnel and monetary resources for the programme.
2. Personal involvement of senior line managers.
3. Participation of the relevant stakeholders, including the parliamentarians when nation-wide policies are being evaluated.
4. Consensus among all participants on the performance measures to be used.

5. Clarity about the process of establishment and intended use of performance information (e.g. to improve programmes and policies or to make budget decisions).

6. Mutual trust, confidence and credibility in the process by those involved. These are preconditions for, among other things, sufficient disclosure of data, information and documentation by the evaluated entities.

The key issues concerning performance measurement can to a large extent be summarised into two points. Firstly, policymakers face a wide array of *information needs*, and performance measurement obviously cannot answer all of them. It mainly serves the purpose of accountability, whereas one of the main – if not *the* main – purposes of evaluation is improvement and development of the evaluated policies. (Feller 2002, 449.)

Secondly, the usefulness of performance measurement systems crucially depends on the *context* within which they are applied. Some institutional conditions seem to lead to more perverse effects than others (Leeuw 2002, 10). Feller (2002, 438) has argued that performance measurement is most appropriate in organisations whose outputs and outcomes are not easily observable. By contrast, science agencies and universities, for example, typically have multiple goals, loosely specified production processes, and probabilistic, long-gestating and loosely coupled linkages between outputs and outcomes, and therefore performance measurement systems are likely to benefit them much less or even be harmful. Because of the complexity, uncertainty and time lags inherent in problems related to the environment and sustainable development, systems and organisations dealing with these issues also often fall into the latter categories. The case is even clearer when evaluation focuses on a broad range of policies, rather than on a single organisation.

The juxtaposition between evaluation and performance measurement can be seen as a conflict between the developmental and the accountability purposes of evaluation, but at a deeper, more philosophical level, it reflects the 'paradigm wars' between the positivistic and hermeneutical or post-modern ideas. Some writers see the purposes of accountability and development/learning as mutually exclusive (e.g. Greene 1999), while others emphasise the considerable learning potential stemming from performance measurement and auditing (e.g. Bukkems and de Groot 2002). Learning and accountability can be seen as complementary as well, a view inherent in Schwandt's (2001b, 235) remarks stating that the element of judgement is actually present also in dialogical evaluation practice. From this perspective, evaluation dialogue can be seen as conversation, exchange of arguments, in which parties involved engage in constructive self-criticism – judgement

of one's own circumstances, pre-understandings and practices – that hence promotes reflexivity. Therefore, to the extent that it promotes informed dialogue accountability evaluation can be considered as an essential element of learning. (e.g. Foulkes 1998; Perrin 1998, 376; Blalock 1999; Davies 1999; Biott and Cook 2000; Wimbush and Watson 2000; Martin and Sanderson 1999; Feinstein 2002; Feller 2002; Owen 2002.)

The measures suggested to achieve complementarity between evaluation and performance management and to minimise the potential negative consequences of the latter have included expanding the notion of performance from programme performance to organisational performance by reinterpreting accountability as the commitment by the organisation to continuous evaluation and dialogue; involving stakeholders in developing, reviewing, and revising performance measurement systems; carrying out periodic comprehensive evaluations to check the accuracy of information gathered through performance management systems; better understanding the context in order to maximise use, but minimise the potential for misuse; testing, reviewing, revising and updating indicators frequently; using multiple, both quantitative and qualitative indicators; and considering alternative approaches such as expert panels. (Perrin 1998, 376; Blalock 1999; Davies 1999; Feller 2002, 450; Owen 2002; Montague and Allderings 2004.)

Peer review as a 'soft enforcement' system attempts to combine the accountability and developmental (learning) purposes of evaluations, thus avoiding some of the negative consequences of strict and binding performance measurement systems (e.g. Perrin 1998, 1999; 2002a; 2002b; Blalock 1999; Davies 1999; Feller 2002). While such soft enforcement mechanisms are not sufficient in all situations, and more coercive measures are sometimes needed, they are nevertheless practically the only tools of influence available to several intergovernmental organisations, such as the OECD.

4.4.4.6 Organisation of the evaluation process, dissemination of results, and follow-up

The organisation of an evaluation process must fulfil the minimum criteria of legitimacy, equity of treatment, transparency, and technical quality. All participants should understand the purpose of the evaluation and what is expected from them. The timing and frequency of evaluations to make them fit into policy and decision-making cycles is crucial for evaluation influence (OECD 1999, 32).

The findings of an evaluation must be adequately communicated to the potential users or interested parties, in multiple forums, using methods appropriate to the particular target group and type of evaluation. Evaluations must be timely and keep to the deadlines. Too often evaluations are only symbolic, stand-alone tools, because at the time they are carried out, nobody needs them, while evaluation results are not available when policy decisions are on the agenda (Lang 2001, 11). However, while lacking or insufficient dissemination has often been identified as a major weakness (e.g. Uusikylä and Virtanen 2000), dissemination is not a panacea. The policymakers' unwillingness to use evaluation results directly seldom stems from lack of awareness of the available policy options. Moreover, better dissemination is unlikely to fend off whatever resistance there might be against the findings of an evaluation or research – rather, it is likely to strengthen the opposition. (Knott and Wildavsky 1980.)

The impact of evaluation must likewise be monitored. In particular, as a part of an 'accountability structure' (Montague and Allredings 2004), the implementation of eventual policy recommendations must be followed-up. While binding sanctions for failing to implement the policies recommended in an evaluation rarely apply, broad publicity is likely to create pressure to change policies (OECD 1999, 36).

4.4.4.7 Participation of stakeholders

Participation is not only an essential element of deliberative democracy, but also a key factor determining the consequences of an evaluation. The involvement of stakeholders in the evaluation process is generally considered as a necessary, even if not alone a sufficient condition for influential evaluations (see e.g. Patton 1997a; 1998; Baron 1999; Eckley 2001; Weiss 1998; Torres et al. 1997), as well as for sustained policy results and empowerment in development projects in general (e.g. Lyons et al. 2001). However, it is not enough to just advocate the broadest possible citizen participation, but one needs to look at the conditions under which such participation takes place, in particular, the relations of power. Key concerns here relate to the ways of organising participation that would best enhance deliberative democracy; the potential limitations and negative impacts of participation; and the tools for measuring participation's contribution towards better evaluations and discursive democracy.

Arguments for participation have ethical, cognitive and normative dimensions. These include people's right to decide on their own destiny, the improved quality and legitimacy of evaluations,

and the capacity of evaluations to empower actors, encourage the development of democratic values, build capacity for self-evaluations, identify and elaborate values and problems of various actors, develop strategies for resolving them, and identify the impacts on all stakeholder groups. (Gregory 2000; Sanderson 2000; McKie 2003.) More specifically, broad involvement of stakeholders is expected to (Floc'hlay and Plottu 1998; Baron 1999, 86-88; Eckley 2001; Boaz and Hayden 2002; Froger and Oberti 2002):

1. Help take into account the complexity of the issues at stake – uncertainty, risk, plurality of values, interests and knowledge – and thus enhance collective and individual learning, improving the quality and the relevance of evaluation findings.
2. Establish a climate of mutual trust among participants, necessary for reaching consensus. This will improve the legitimacy and acceptability of decisions and evaluation findings, as well as create a sense of 'ownership' of the evaluation results among the stakeholders.
3. Limit the influence of powerful groups with vested interests and bring forward the public interest.
4. Promote the acceptance of the evaluation process by the stakeholders and hence facilitate information gathering.

However, criticism towards the 'participation orthodoxy' (Gregory 2000) has become increasingly frequent. Henry and Mark (2003) go as far as to argue that the call for greater participation is often founded more on the belief rather than empirical evidence of participation's beneficial impact on evaluation use. Other research has, by contrast, provided ample evidence of the capacity of participation to enhance evaluation use (see e.g. Baron 1999; Patton 1997a, 87-113). What seems clear is that participation is not a panacea, but it has its problems and limitations of both practical and more fundamental character. They can be classified under three main headings (Gambetta 1998; Baron 1999, 127-132; Eckley 2001; Hisschemöller et al. 2001, 63-64; O'Connor et al. 2001/2002, 35-36; O'Neill 2001; Parkinson 2001; Sager 2001, 210; Henry and Mark 2003; McKie 2003, 309):

Problems related to scarcity of resources and lack of motivation to participate

- Well-designed participatory processes require considerable amounts of time and resources and may be perceived as an unnecessary burden slowing down the execution of policies.

- Budgets for evaluations are often limited, which necessarily restricts the range of issues that can be handled, and engenders frustration among the weaker groups, whose ideas are often considered as marginal.

- If the weaker groups have inadequate resources, capacities, and knowledge in relation to experts and powerful groups, a false image of participation on equal terms may emerge. The 'code of presentation' or 'rhetorical spaces'[121] largely defined by the funding agencies and evaluators tends to delineate who and which types of ideas are credible, thus limiting the contribution that the ideas, insights and knowledge of participants can bring to evaluation.

- The benefits of participation often appear only in the long term. Therefore, it may be difficult to motivate the groups not directly affected to make the personal 'investment' participation would require. The motivation to participate may not last over long periods, and tiredness and exhaustion may eventually lead to low public activity.

Risk of poorer quality and hence lesser credibility of decisions

- Insufficient competence of the participants may compromise the quality of evaluations.

- While broad participation tends to enhance the legitimacy of the evaluation, it may jeopardize the evaluation's credibility, if the process thus becomes too 'politicised'.

Problems related to the political and administrative setting

- Broad participation may prompt parties to seek consensus around the smallest common denominator, making participatory processes inherently conservative. For the same reason, the most critical perspectives tend to be excluded.

- The political and policymaking environment can restrict participation, rendering it superficial and increasing the risk of manipulation.

- Project managers, as well as commissioning and funding bodies can be reluctant to relinquish control of the information and processes crucial to the collection of data and the conduct of the evaluation.

- Since participation of all affected individuals is seldom feasible, participation needs to build on at least partial representation, with the concomitant difficulties of determining 'who is to represent whom?'

---

[121] McKie (2003, 308) refers to 'rhetorical spaces' as the "tacit social knowledge and territorial imperatives that structure and restrict how knowledge and experiences can be voiced".

Sometimes inappropriately organised participation may not only reduce the instrumental rationality (e.g. through compromising the quality of evaluations), but may result in what Sager (2001, 207-208) has called 'parapraxis', where the limitations of communicative and instrumental rationality mutually reinforce each other, leading to accumulated irrationality in policymaking. For instance, under time constraints, the planning and evaluation processes may be limited to a very general analysis, thus compromising the instrumental rationality. The lack of careful analysis may alienate local people, discourage participation, and thus reduce communicative rationality, generate protests and conflicts, which again further hamper the instrumental rationality of the policy process. (ibid.)

Therefore, it may sometimes be desirable, in the interest of deliberative democracy, to limit participation, in order to safeguard against the use of a seemingly participatory process as a means whereby powerful groups legitimise their own actions (e.g. O'Neill 2001; Baron 1999, 129). This would imply leaving 'model-weak'[122] but well-organized groups out of the evaluation process, letting them to carry out direct protests outside the process.

To minimize the negative aspects and problems associated with participation in evaluation, Baron (1999, 141-145) has proposed to calibrate participation in function of the phase of the process. Broad participation at the *early stages* should reduce the risk that relevant aspects are overlooked or forgotten. *During the evaluation* process, participation should be limited to a small number of highly motivated and sufficiently skilled participants so as to ensure a high quality of the evaluation. At the stage of *collecting information* participation should be relatively deep, i.e. interaction should be intensive, but not extended to a very large number of participants. *After the evaluation* report has been released, the broadest possible range of actors should be engaged in public debate and search for adequate solutions to problems identified in the evaluation. Even when participation in the early stages of the process has been modest, presentation of the evaluation results through simulations or question-answer sessions with evaluators may enhance the salience of the evaluations (Eckley 2001).

Eckley (2001) has argued that even symbolic participation – formal opportunity to be present – may be enough to ensure sufficient *legitimacy* of the evaluation, whereas genuine participation is needed for enhanced *salience*. Moreover, reduced *credibility* of the evaluation process, as a result of broad

---

[122] The term 'model' should be broadly interpreted to encompass both highly analytic and elaborated and more rudimentary and simple mental pictures of a part of reality (Sager 2001, 210).

participation of stakeholders sometimes lacking requisite skills and knowledge may be the price to be paid for enhanced legitimacy, when issues are conflict-laden and highly contested. Therefore, it is necessary to examine the policy context in order to strike an appropriate balance between credibility, salience and legitimacy.

4.4.5 Summary: factors potentially affecting evaluation influence

This section has listed the factors that have the potential to shape the influence of an evaluation carried out by an international organisation. The factors were classified into four broad categories: the role of an international organisation as evaluator; actors with their expectations, interests, and 'repertoires'; the political and cultural context; and the design of evaluations. Finally, the problems and possibilities that a greater involvement of different parties in the evaluation process would present were evoked. This presentation does not constitute a theory of how the various factors interact to produce their impacts, but the primary purpose has been to draw up a 'checklist' of factors that merit attention when examining concrete evaluation situations, and to identify some of the relations between the factors. This list will serve as the framework of analysis in chapter 7, and should facilitate the construction of more specific hypotheses concerning factors that affect evaluation influence and their interaction.

Figure 4.1 presents schematically the factors affecting evaluation influence. Actors have each their own repertoires, they act as members of epistemic or other policy communities, in which the evaluator as an international organisation plays a key role. These repertoires, in interaction with the context and evaluation design, determine the extent to which actors perceive the evaluations credible, salient and legitimate. Credibility, salience, and legitimacy are the criteria through which actors look at evaluations and assess their potential value, while repertoires provide the framework of interpretation through which actors look at evaluation. Finally, the positions of actors in policy networks, be they epistemic communities or otherwise, influence their repertoires and repertoire-based interpretations of evaluation situations.

Figure 4.1. Factors affecting evaluation influence.
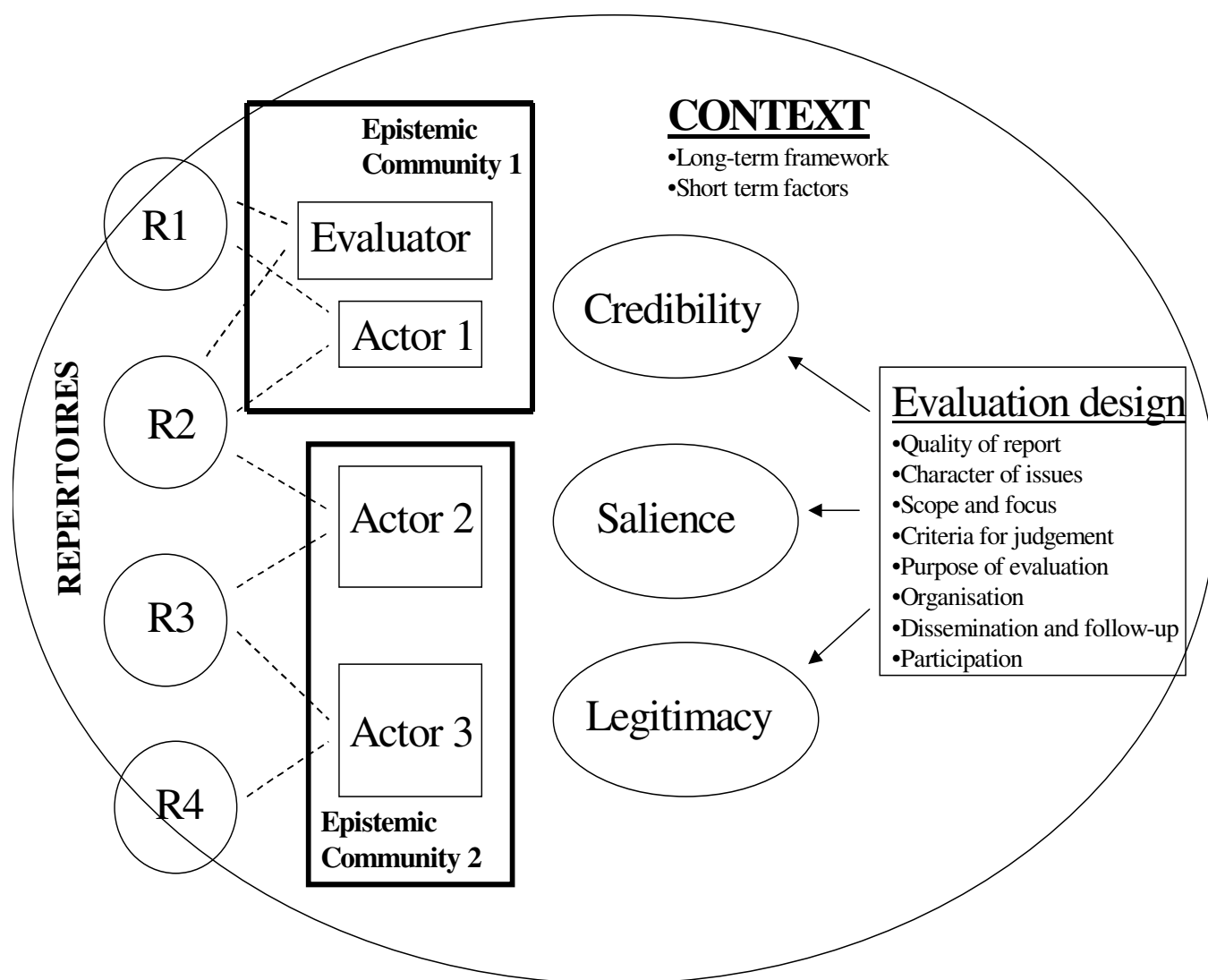
This chapter completed the theoretical framework that shall be applied for analysing the influence of the OECD Environmental Performance Reviews in the second part of the study**.**

PART II

INFLUENCE OF THE OECD ENVIRONMENTAL PERFORMANCE REVIEWS ON SUSTAINABLE DEVELOPMENT

This second part of the study applies the concepts developed in the first part in order to analyse the OECD Environmental Performance Review (EPR) programme. The case study sets out to establish 1) which have been the impacts from the EPRs; 2) which factors have affected the degree of influence the EPRs have had; and 3) to what extent the EPRs' evaluation approach corresponds to the ideals of an 'institutionalist' evaluation of sustainable development, as defined in chapters 2, 3 and 4 of the present study.

Chapter 5 presents the OECD as an international organisation, its work on the environment and sustainable development, and the case study – the Environmental Performance Reviews. Applying the analytical framework and the typology developed in section 4.3, chapter 6 examines the types of influence from the EPRs, whereas the factors that condition the influence of the reviews are analysed in chapter 7, drawing upon the typologies developed in section 4.4. Chapter 8 looks at the EPR 'message' and the EPR process from the perspective of the principles of 'institutionalist' evaluation of sustainable development, listed in the end of section 4.2.

Chapter 5

Description of the Case Study: OECD Environmental Performance Reviews

This chapter will first describe the OECD's role as an international organisation – its mandate, working methods and different roles. It then goes on to give an overview of the OECD sustainable development work and of the role of the OECD Environment Directorate followed by a presentation of the objectives, structure, and history of the OECD Environmental Performance Review programme. Finally, in the last section, the EPRs are compared with the oldest and most prestigious OECD peer reviews, the Economic Surveys.

5.1 OECD as an international organisation

The OECD can exert neither regulatory nor financial pressure to influence the behaviour of actors in its member countries. It is therefore a typical organisation in the sense described in section 4.4.1, exercising its power through 'norm creation', 'idea games', moral persuasion and peer pressure. The most important characteristic of the OECD is that no decisions are made – the OECD committees are not mainly engaged in preparing decisions for the OECD Council (Marcussen 2001, 1). Julin (2003) has described the OECD as simultaneously fulfilling the roles of a think-tank, a databank, a problem-solver, a pathfinder, a policy adviser, a forum and a meeting place, and sometimes a deal broker. Gass (2003) defines analysis, impartiality, professionalism, and innovation as the OECD's key characteristics, and sees the organisation as a "policy laboratory".

5.1.1 Mandate, working method, and reform: the search for a lost identity

The OECD's mandate consists of three objectives: to promote policies designed to achieve the highest sustainable economic growth; to contribute to sound economic expansion in member as well as non-member countries; and to contribute to the expansion of world trade.[123] There is a virtual consensus among the member countries about the continuing relevance of these objectives as the foundation for OECD work. (e.g. OECD 2003h). However, over the past decade, changes in the international policy arena have lead to a situation in which the organisation seems to have somewhat lost its original identity, and its legitimacy is increasingly being questioned, both within and beyond the member governments. While the workload and the demands placed upon the OECD are increasing – six new countries have joined the organisation since 1994,[124] and ever new areas are being proposed and added to the OECD list of activities – the available resources have been dwindling.

The OECD can be seen as a relatively 'lean' organisation compared with some other international organisations, such as the UN, the World Bank or the IMF. The permanent professional staff working in the OECD secretariat is only 4.6 per cent larger today than it was in the late 1970s, and almost 9 per cent less than in 1995 (OECD 2003a). This is so despite the near doubling of the GDP volume of its member countries since the late 1970s. Today, OECD members have about 60 % of world GDP and account for three quarters of world trade (Julin 2003). The lack of resources can be seen partly to reflect the declining importance that the member countries seem to attach to OECD work in relation to that carried out by other organisations and bodies. Among the reasons for the OECD needing to 'fight harder for less' and look for its lost identity are the following (Julin 2003; OECD 2003a, 20; OECD 2003h):

1. Along with **globalisation**, the number of important players in the global policy arena has increased considerably, which has accentuated the competition in the arena of international policy analysis and advice. The urgent short term issues tend to crowd out the longer term concerns in which the OECD has its comparative advantage.

---

[123] The aims of the OECD as set forth in its founding Convention are: (a) to achieve the highest sustainable economic growth and employment and a rising standard of living in member countries, while maintaining financial stability, and thus to contribute to the development of the world economy; (b) to contribute to sound economic expansion in member as well as non-member countries in the process of economic development; and (c) to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations. (http://www.oecd.org/document/7/0,2340,en_2649_201185_1915847_1_1_1_1,00.html)

2. The ***changing role of the state*** has lead to the expansion of the role of non-governmental actors, such as business and civil society.

3. The ***end of the cold war*** and the intensification of the European integration have moved the geopolitical priorities away from the OECD.

4. The extensive ***pressure on governments to cut spending*** also affects the OECD, which has made the OECD staff feel that governments and the public at large are not listening to them.

These problems have been partly reflected in the discussions and outright disputes over the financing of the organisation's work. The main disagreements have concerned the sharing of burden in financing the organisation's activities and the division between the budgetary and voluntary contributions. In slightly simplified terms, the OECD budget consists of two parts: 1) the general budget, financed by member countries' contributions that are proportional to their share of the sum total of member countries' GDP, and 2) voluntary contributions that can be added to any budget line. The latter are not bound by the consensual budget decisions adopted by the organisation's highest decision-making body, the Council.[125] At present, voluntary contributions cover about 25 per cent of OECD's financing needs, and the share is growing rapidly, while the general budget is being cut or kept stable in nominal terms. In practice, voluntary contributions are a way for the wealthiest countries to steer the OECD work according to their interests, by avoiding the consensual decision-making within the Council, thus reducing transparency and compromising the possibilities of long-term planning of activities. A zero real growth objective has been established for the OECD general budget in the upcoming years (Ulkoasiainministeriö 2003c).

The OECD has, in fact, evolved and undergone changes in the 1990s. New members have joined; cooperation with non-member countries has grown significantly; communication has been improved; new, more flexible working methods have been introduced; and dialogue with civil society has been strengthened through greater involvement of NGOs. However, these changes have not been sufficient to erase doubts concerning the organisation's legitimacy and usefulness. The following list contains some of the main criticisms towards the OECD:

---

[124] These new members were Mexico (1994), the Czech Republic (1995), Hungary (1996), Poland (1996), Korea (1996), and the Slovak Republic (2000).
[125] Consisting of the permanent representatives of member countries to the OECD and meeting once a year at the ministerial level, countries usually being represented by their ministers of finance and/or foreign affairs.

1. It lacks global legitimacy, being perceived as a 'club of rich countries', too inward-looking an organisation, despite the attempt to increase stakeholder consultations.[126] Therefore, its utility to non-OECD countries or to non-governmental actors has been judged poor. According to this argument, the organisation has practically fulfilled its main mission – promoting the market economy – but has failed to contribute to global development, as the gap between the rich and poor countries has kept growing, apart from a few exceptions.[127]

2. There is duplication of work among the different global players, and many of the OECD's tasks and functions could be better handled by other international organisations.

3. The OECD has been incapable of responding quickly to emerging, new policy issues, because of its inflexible organisational structure. Because practically all decisions must be taken by consensus among the member countries, discontinuing the activity of a committee is virtually impossible, since there are always groups opposed to such changes.

4. The OECD has low visibility, because of its overly prudent communication approach, reflecting the need to ensure consensus among its member countries. For the same reason, the organisation tends to avoid "politically potentially explosive products" that would for instance compare countries' performance in a particular sector. (OECD 2003h, 20).[128]

There is a broad consensus concerning the *functional core competences of the OECD*: economic analysis of policies; production of comprehensive, internationally comparable statistics; multidisciplinary expertise; policy surveillance and peer review; benchmarking and best practices; standard setting and development of 'soft law'; and the provision of a forum for discussions in a non-negotiating context. Matters become more complicated when trying to define what *activity areas* constitute the core. Trade, development, taxation and issues related to corporate practices are among the most often mentioned subject areas, while views among member countries are quite divided on the importance of environmental work, notably work that is not directly linked to economic aspects or possibly linked to sustainable development. In the area of sustainable development, 'decoupling' of environmental pressures from economic growth is the theme that enjoys the broadest support (OECD 2003b, 35). The work on environmental peer reviews, indicators and outlooks has over the years received stable support (ibid.). The United States tends to

---

[126] This was, for instance, the unanimous judgement of the stakeholders in the meeting preceding the OECD Environment Ministerial meeting in April 2004.

[127] Often, the OECD has been accused – or praised – for promoting neoliberal, monetarist policies in all areas of its activity, from agriculture to technology or fiscal policies (L'OCDE est un…2001).

[128] Exceptions to this rule exist, of course, the OECD Programme for International Student Assessment (PISA) being among the clearest example of studies that can have a decisive impact on policies, by virtue of the public interest and

see much of OECD's environmental work as unnecessary, especially when it comes to the environmental aspects of trade, agriculture, biodiversity, and greater harmonisation of national environmental policies across countries. It argues that instead of producing normative guidelines, the OECD should focus its work in the environmental field on the provision of reliable background information. (e.g. Ulkoasiainministeriö 2003a.) In general terms, the new American unilateralism has greatly affected OECD work, not least in the areas of the environment and sustainable development.

One characteristic that distinguishes the OECD from most other intergovernmental organisations and which is both its strength and its weakness is that its constituencies in the capitals are spread throughout government administrations (Julin 2003). It is a weakness in so far as the absence of a clear 'lead' ministry looking after the interests of the organisation in national governments creates problems of co-ordination. However, the positive side of the arrangement is its potential to increase inter-ministerial co-operation and foster environmental policy integration, for example. At present, the rather rigid and compartmentalised organisational structures of the OECD – reflecting the structures of its member country administrations – compromise its ability to carry out truly cross-cutting analysis (Julin 2003).

To tackle the problems it is faced with, the OECD launched a *reform process*, with the intention to better define its core activities, reorganise its committee structure, and thus enhance its impact in the member country capitals (e.g. OECD 2003h).[129] The current reform proposals within the OECD have two main objectives. The first one is to enhance the visibility and the influence of the organisation's work in its Member countries and beyond, whereas the second relates to managing and preparing for the likely enlargement of the organisation to new member countries.[130] The OECD sees its comparative advantage in its capacity to provide analytical tools necessary for the management of globalisation processes.[131] This would imply a more horizontal approach in areas such as sustainable development and enhancing the coherence of policies across countries.

---

political pressure they generate. PISA is a three-yearly survey of the knowledge and skills of 15-year olds in the principal industrialised countries aimed at producing valid comparisons across countries and cultures.

[129] The final communiqué from the OECD Council Ministerial meeting in May 2001 (OECD 2001g) mentions co-operation and dialogue with countries beyond the organisation itself as the first item, thus signalling the desire of the OECD to enhance its external legitimacy.

[130] A recent report on the OECD's future directions mentions 45 as the medium-term ceiling for the number of member countries (OECD 2003h, 6).

[131] Globalisation has been mentioned as a promising, but so far inadequately exploited opportunity for the OECD to higher its profile (OECD 2003h). In an interview to the French newspaper 'Libération', the Secretary General of the OECD, Donald Johnston, sees the organisation's main task the promotion and "humanisation" of globalisation, by seeking to strike a balance between the objectives of economic growth and social cohesion.

Cooperation with the big emerging players such as China, Russia, India, Brazil, Indonesia, and South Africa would need to be strengthened, with a view to preparing them for possible membership. Other potential new member countries are Chile, Argentina, Singapore, and the new EU countries not yet members of the OECD. It has been pointed out that the enlargement should not lead to excessive 'Europeanisation' of the OECD, and that the above mentioned 'big six' should be among the new members (Ulkoasiainministeriö 2003b).

In order to remedy the legitimacy problem, the OECD created a new Directorate for Public Affairs and Communications in 1997 and employed soon after the failed MAI[132] negotiations in 1998 a director responsible for public relations and communication, with a view to raising the visibility of the organisation's work and increasing the interaction with civil society (OECD 2004b, 30). Despite recent initiatives both by the OECD and by the civil society (see e.g. OECD 2001i; http://www.oecd.org/department/0,2688,en_2649_34495_1_1_1_1_1,00.html), the NGO community still remains less integrated in the OECD structure than the business and trade union circles, which have their own permanent consultative organs in the OECD: the Business and Industry Advisory Committee (BIAC) and the Trade Union Advisory Committee (TUAC). Establishing such a consultative body for the NGOs has been discussed, but the member countries have remained divided on the issue, the United States being open to the idea, Japan and France adopting a more reserved attitude (Les institutions internationales…2001). Neither are all NGOs wholeheartedly behind such an idea, given the financial burden participation would impose on them, and the potential prejudice to their credibility, if the NGOs' independence in relation to the OECD is called into question. For the time being, the NGOs are invited to participate, on an *ad hoc* basis, in many of the OECD's meetings as unofficial observers. They are, however, often in a disadvantaged position in relation to the national delegates, notably with respect to the access to background documents and meetings in which real policy discussions are conducted.[133]

The reform work concluded – unsurprisingly – that many of the charges levelled against the OECD were unfounded. For instance, the OECD benefits not only its member countries, but a wide range of non-members as well, since virtually all the information comes into the public domain, and is available to those who seek it. Furthermore, the importance of the OECD co-operation with developing countries and the rapid growth of the 'outreach' activities testify to the organisation's

---

[132] The Multilateral Agreement on Investment.
[133] Personal communication, 14 September 2003, with the then Finnish delegate to many of the OECD's bodies dealing with the environment, Heikki Sisula, and the NGO representative Hanna Matinpuro from the Finnish Association for Nature Conservation.

openness towards non-member countries. (Gass 2003, 30.) The gradual move from consensus to qualified majority voting is expected to enhance the efficiency and capacity of the organisation to respond to new challenges. Competition with other global players does indeed exist, but the OECD claims to possess a combination of qualities unparalleled by that of any other organisation, by virtue of:

- its intergovernmental character;
- broad range of its Member countries (as compared with, for example, the EU or the G7);
- more focused capacity for policy analysis among a like-minded group (by contrast with the UN agencies);
- non-negotiating context in which most OECD bodies meet;
- unique cross-national data sets and deep analytical capacity of many OECD divisions; and
- highly developed capacity to carry out benchmarking and peer reviews (OECD 2003a, 9).

Hence, the OECD's fundamental purpose and design – a policy R&D organisation with (non-elected) officials as the primary audiences, augmented periodically by the direct engagement of ministers – remains valid, and the organisation continues to balance a creative tension between orientation towards action on the one hand, and reflection on the other (ibid., 21). However, no matter what the OECD officials or other 'insiders' might say, the fact remains that the organisation has an image problem, which affects its capacity to make an impact. To put into perspective the discussion on the OECD's role, and its means of influencing policies through 'norm creation', let us now look at how Marcussen (2001) has analysed the OECD's different roles in international policymaking.

5.1.2 OECD's traditional roles: ideational artist and ideational arbitrator

The fundamental OECD 'mission' can be seen to consist of two roles, both essentially based on consultation. Firstly, as an *'ideational artist'* the organisation is an enormous 'think tank' that formulates, tests and diffuses new policy ideas. Since the workforce of the OECD consists of well-paid experts independent of their home countries, and given its financial and political independence, the OECD would be able to distance itself from national controversies and possess considerable room for manoeuvre when inventing new ideas in old policy areas. In the environmental policies,

the OECD's most well-known contribution was the invention and 'marketing' of concept of polluter pays principle. (Marcussen 2001.)[134]

Secondly, the OECD plays the role of an *'ideational arbitrator'*, allowing national civil servants to meet each other in supportive surroundings, thereby helping them acquire skills and competence, and sometimes even develop their personality and feelings of belonging, through **learning processes** such as socialisation, imitation and coercion. The hundreds of OECD meetings held in its headquarters in Paris, which attract each year some 40 000 civil servants, vary in their character from tough negotiations, in which the delegates' room for manoeuvre is constrained by strict mandates from their capitals, to meetings that allow delegates practically complete freedom to exchange with their peers from other countries. In the latter kind of meetings, delegates often do not even totally grasp their own role, the problem at hand, or possible solutions to the problems treated, yet through deliberation, they slowly learn more about these matters. Most of the meetings fall between these two extremes. On some occasions, an issue becomes politicised, thus introducing an element of bargaining and negotiation to the meeting. On other occasions, the sessions become increasingly routinised, social norms of appropriateness regulating delegates' behaviour. (ibid.)

5.1.3 From idealism to realism: OECD as an ideational agent and agency

According to Marcussen (2001), these two core functions, which fully correspond to the functions of the OECD envisioned by its first secretary-general, Thorkil Kristensen, have been complemented by two other, less idealistic roles, namely those of an 'ideational agent' and an 'ideational agency'. As an **ideational agent**, the OECD picks up ideas that are prevalent among the most prosperous members states and then links these ideas to department-specific areas of activity and transfers the final product to other, less centrally placed member states. In essence, the OECD becomes an agent promoting new ideas developed by those member countries that have the ability and willingness to do so.

As an ideational agent the OECD takes up an idea once it has been developed by a particularly influential source, replicates it through its own official discourse, and helps to diffuse the idea

---

[134] Marcussen (2001, 3-6) uses the example of the shifts in the OECD's economic policy doctrine as an example to show that the organisation indeed plays such a role of ideational artist. After the years of neo-Keynesian policy doctrine, in late 1970s and early 1980s the OECD developed what became the organisation's new economic policy 'dogma' with focus on low inflation, budgetary discipline, a medium-term perspective, and structural adaptation. This

amongst its member countries. Once the idea reaches a 'tipping point' of sufficient support among the member countries, no member country wishing to remain legitimate within the international economic elite society can afford to ignore it. Subsequently, the various OECD directorates exploit the reigning idea as it fits their immediate purposes, thus effectively protecting themselves from a closure that might have threatened had the department failed to find new arguments for survival in a changing context.[135]

Finally, as an ***ideational agency***, the OECD constantly surfs around among the national political debates in order to make a good 'business deal'. When it discovers that a set of ideas has gained ground among the member states, it picks up these ideas, operationalises them so that they end up taking the form of causal ideas that can be resold in the member states. In other words, the OECD acts as a grocery or an agency that provides its clients the kind of products that happen to be in general demand. It is continuously scanning the entire OECD territory in order to better grasp the political atmosphere in the respective countries, with the view to diffusing ideas that easily fit into the local contexts.[136] This last role, in which the OECD agrees to diffuse any idea for which there is general demand in the member countries, seems to be in a direct conflict with the ideals of Thorkil Kristensen. As opposed to the perception of an organisation sufficiently independent to produce ideas that have not yet been thought of in the member countries, the OECD would thus risk becoming just another consultant among many.[137]

The categorisation of the four roles of the OECD brings us back to the discussion on the role of international organisations as 'norm creators' throughout the different phases of the 'norm life-cycle' that were discussed in the previous chapter. While the concepts of norm creator and norm life-cycle focused on the international organisation as a relatively independent actor, the four roles evoked above draw attention to the sources of norms and the influence of the general policy

---

new policy line was adopted despite the fact that an alternative would have been available, notably developed in the mid-1970s within the European Commission.

[135] As an example, Marcussen (2001, 10-15) describes the way in which the so-called structural determinist globalisation discourse, developed initially in the U.S. during the early days of the Clinton administration, was effectively diffused through the OECD. Even though such a discourse by no means enjoys a consensus among globalisation scholars, a consensus was discursively constructed – with major assistance by the OECD – among national political elites regardless of their place on the political spectrum.

[136] Marcussen (2001) takes the concept of ethics in public administration as a prime example of the OECD Public Management Committee acting as an ideational agency. Instead of creating a unique and inflexible concept of ethical public administration, PUMA developed a conceptual apparatus – a framework of analysis – to which nobody could effectively be opposed, whether advocating a value system based on economic effectiveness or on political principles of regularity, simplicity, openness, and equal treatment.

[137] UNESCO successfully redefining its role in science politics can be evoked as another example of an international organisation operating as an ideational agency (Finnemore 1993).

environment on the functioning of an international organisation. We can look at the four roles from the perspective of the independence and autonomy of the OECD, illustrated through figure 5.1. As an ideational artist, the OECD is the source of new ideas, as ideational arbitrator it continues to play an important role as innovator and transmitter of policy ideas, whereas ideational agent and agency are rather in receiving role in the processes of norm creation.

Figure 5.1. The four roles of the OECD.
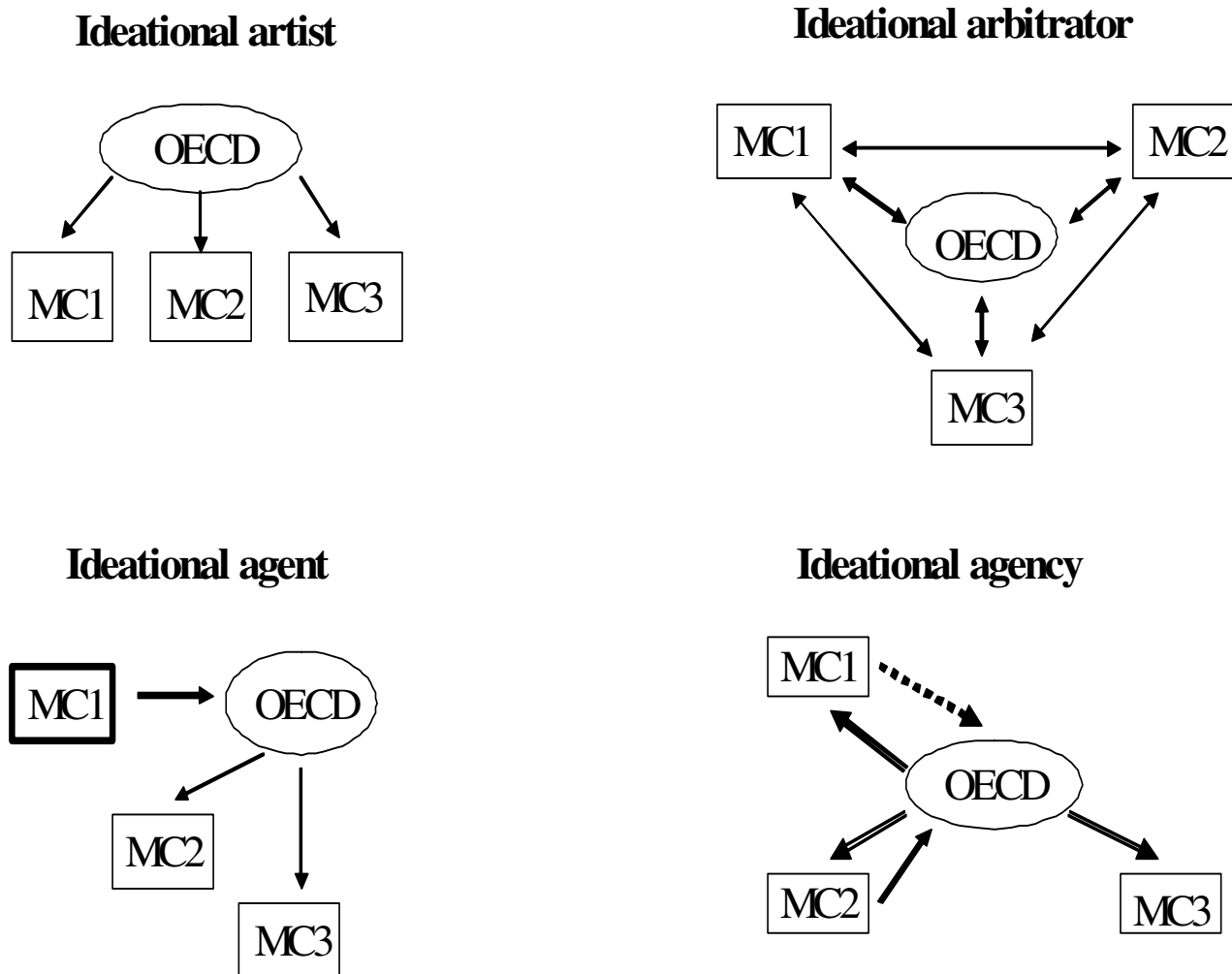
A key concept in Marcussen's typology is ***credibility***, which he considers as an indispensable resource for the OECD. As an ideational artist, the OECD gains credibility among its member countries by helping them learn in situations in which they have gone through a painful process of unlearning. The credibility of an ideational arbitrator rests on the effectiveness with which it is

capable of communicating with national civil servants, drawing these into a learning process. As an ideational agent, the OECD gains credibility by playing a game defined by the countries that carry the main burden for financing the organisation. Finally, the credibility of an ideational agency stems from the constant operationalisation of ideas for which there is already a demand among the member states and the assistance given to countries in the concrete implementation of these ideas.

The independence of the OECD therefore declines from the fully independent 'ideational artist' towards the 'ideational agency' – a role where the OECD is completely at the mercy of its member countries. The mechanisms of 'norm creation' – socialisation, imitation, and coercion – as well as the reasons why such mechanisms are operational (legitimacy, conformity, esteem) remain the same, but their locus tends to move forward along the norm life-cycle. Ideational artist ideally operates as a norm entrepreneur, whereas the ideational agent and agency no longer have the capacity to innovate, but take up innovations developed elsewhere and rather contribute to 'norm cascade' and 'norm internalisation'. In these two latter roles, the OECD has clearly entered the 'power game', in which its independence as a source of legitimacy and credibility in promoting norm adoption may be called into question.

The OECD work in the area of sustainable development, the theme that will be addressed in the following section, provides an example of the organisation's quest for credibility in a changing world, and helps situate the Environmental Performance Reviews in a broader context.

5.2 OECD sustainable development work

As early as 1997, a High-Level Advisory Group convened by the Secretary-General of the OECD concluded its work by an overall recommendation that the OECD, "given its unique ability through systematic analysis and peer review to develop a shared framework of strategic policy […] should become the key intergovernmental organisation providing the industrialised nations with the analytical and comparative framework for policy necessary for their economies to make the transition to sustainable development." Furthermore, it called upon the organisation to "re-interpret the 1961 OECD Convention – which calls on the organisation to pursue policies that promote sustainable economic growth and employment – in the light of the 21st Century conditions and challenges so that sustainable economic growth takes on a new meaning". (OECD 1997c). Following these recommendations, the 1998 Ministerial Council "agreed to interpret the term

sustainable as including social and environmental, as well as economic considerations" (OECD 2004a). In the following year, the first cycle of horizontal work on sustainable development was initiated, a project that culminated three years later with a declaration from the OECD Ministerial Council and the publication of two reports on the OECD approach to sustainable development (OECD 2001e; 2001f). The Council decided to extend the project by three years (for 2001-2004) with a mandate to:

- develop agreed indicators and incorporate these in OECD economic, social and environmental peer review processes;
- identify how obstacles to policy reforms can be overcome;
- analyse the social aspects of sustainable development; and
- provide guidance on policy coherence and integration (Lorentsen 2004; OECD 2003d, 13; OECD 2004a, 8).

The work was overseen by an Ad Hoc Group on Sustainable Development, composed of delegates from the member countries. It reported back to the OECD Ministerial Council meeting in 13-14 May 2004 and recommended further work on

- obstacles to reducing environmentally harmful subsidies;
- obstacles to further use of economic instruments;
- sustainable resource use (material flow accounting, decoupling and resource productivity); and
- emerging issues (e.g. to feed into the work of UNCSD) (OECD 2004a, 57).

Later in 2004, the Ad Hoc Group was transformed to an Ad Hoc Experts Meeting on Sustainable Development, with a more limited scope of action, the precise mandate and budget still being uncertain in early 2005. This change significantly reduced the weight of sustainable development work within the OECD.

In addition, a Ministerial-level Round Table on Sustainable Development, an independent body hosted by the OECD, composed of representatives from different stakeholder groups, was established in 1998 to enhance ministerial interest in sustainable development, attract ministers from member and non-member countries to meetings in Paris, and provide intellectual stimulus to

the work of the Secretariat (OECD 2004a, 8). In practice, the work of the Round Table has been poorly integrated with that of the horizontal Ad Hoc Group, and its influence on the OECD work has been modest (Suomen pysyvä edustusto 2003).

The sustainable development work in the OECD has faced a number of difficulties, related to the difficulties of agreeing on a common definition of sustainable development, notably on the meaning of the social dimension of sustainability (Caccia 2003); governments' tendency to prioritise activities carried out by a single OECD Committee at the cost of 'whole-of-government' themes; the lacking authority of the Ad Hoc Group, dominated by delegates from the politically 'lightweight' environment ministries; the scarcity and instability in the allocation of resources to the work; and the lack of a clear 'lead directorate' in the OECD (OECD 2003h; 2004a). Together with the reform leading to the creation of an Ad Hoc Experts Meeting, a decision was made to nominate a sustainable development coordinator in the Environment Directorate for the upcoming two-year work period, a decision, which can be seen to reflect the perception that sustainable development, in the end, is an environmental, rather than a cross-cutting issue.

5.3 Role of the OECD Environment Directorate

The mission of the *OECD Environment Directorate*, founded in 1970, is to "assist countries to improve their environmental policies by providing policy advice based on sound analysis, cross-country experience, and reliable data, and by fostering accountability through peer pressure". The staff of about 100 members operates on a budget of some 12 million euros per year, half of which comes through voluntary contributions. The main areas of work of the Environment Directorate consist of (Lorentsen 2004):

- Environmental Performance Reviews (EPR), indicators and outlooks;
- Natural resources management and climate change
- Environment, health and safety; and social linkages; and
- Decoupling environmental pressures from economic growth (Lorentsen 2004).

The OECD Council adopted in 2001 the OECD Environmental Strategy for the First Decade of the 21st Century, which provides the foundation for the Environment work programme. The strategy

can be seen as the Environment Directorate's contribution to the OECD's search for its lost identity. It consists of five key objectives:

1. Maintaining the integrity of ecosystems through the efficient management of natural resources.
2. Decoupling environmental pressures from economic growth.
3. Improving information for decision making: measuring progress through indicators.
4. The social and environmental interface: enhancing the quality of life.
5. Global environmental interdependence: improving governance and co-operation.

The work of the Environment Directorate is overseen by the Environment Policy Committee (EPOC), the Chemicals Committee, and their subsidiary bodies, such as the Working Party on Environmental Performance (WPEP), responsible for the Environmental Performance Reviews. The Environmental Performance and Information Division (EPI) is in charge of running the review programme.

5.4 Environmental Performance Review Programme

The OECD has carried out systematic reviews of its member countries' environmental policies since 1992. During the first cycle of the reviews (1992-2000), the principal aim of the EPR programme was "to help Member countries improve their individual and collective performances in environmental management" (OECD 1997a, 5). The reform of the programme for the second cycle of reviews introduced the notion of sustainable development as the ultimate objective of the reviews.[138] The specific goals of the programme remained largely unchanged, apart from the strengthened emphasis on the social aspects and links with all sectors of society. The specific goals for the second cycle therefore read as follows:

1) helping individual governments assess progress by establishing baseline conditions, trends, policy commitments, institutional arrangements and mechanisms for evaluating policy results in their social and economic contexts;

---

[138] The new definition of the principal aim of the EPRs states: "[t]he principal objective of the review programme is to help Member countries improve their individual and collective performances in environmental management with the goal of achieving sustainable development" (OECD 1998a, 6).

2) promoting environmental improvements and a continuous policy dialogue among Member countries, through a peer review process and by the transfer of information on policies, approaches, and experiences of reviewed countries;

3) stimulating greater accountability from Member country governments towards representatives of all sectors of society both at national and international levels. (OECD 1998a, 6.)

Together with other country reviews of the OECD,[139] the EPRs constitute an element in the organisation's efforts to analyse sustainable development in its member countries. As a result of the changing policy context during the 1990s, and in line with the OECD's three-year effort to place sustainable development at the centre of its work, the 'second cycle' of reviews that started in 2000 was to address all three dimensions of sustainable development – economic, environmental and social. (e.g. OECD 2001b; 2001c.) In practice, the most visible innovation was the introduction of the environmental-social interface as a special chapter in the reviews.

5.4.1 Structure of the programme

The structure of the EPRs and the 'review logic' have been consolidated during the 'second cycle'. They can be summarised in a '4 x 3 structure' as follows (OECD 2000):

1. Objectives. The basic criteria for evaluation consists of 1) the objectives the government has established and the international commitments it has adopted; 2) commonly agreed OECD policy principles; and 3) the cost-effectiveness of policies. These objectives can be divided into (i) general *aims*, (ii) more precise *goals*, and (iii) quantitative *targets*.

2. Performance. When evaluating performance, three aspects are taken into consideration: (i) the *intentions* of the government to implement policies, draw up programmes, etc. (ii) the *actions* the government has undertaken, and (iii) the *results* from policy measures.

3. Framework. The framework of analysing results follows the OECD environmental indicator model (e.g. OECD 2003e), in which attention is given to (i) the *pressures* on the environment from

---

[139] Such as Economic Surveys, territorial development reviews, development co-operation reviews, employment reviews, education reviews, and energy policy reviews, the latter being carried out by the International Energy Agency (IEA).

human activities, (ii) the *state* of the environment, and (iii) government's *responses* aimed at remedying problems.

4. Outputs. Each chapter in the review report consists of three types of information: (i) *description* of the state of the environment, the context, and the actions taken in order to tackle the problems, (ii) *evaluation* of performance in reaching the objectives, and (iii) *prescriptions* given to the country in the form of policy recommendations.

5.4.2 Review process

At the beginning of each Environmental Performance Review the head of the OECD Environmental Performance and Information Division visits the country in order to agree on the detailed scope of the review and explain the steps of the review process to the persons responsible for organising the review process in the country. This is followed by the preparation of background material by the reviewed country, according to the guidelines and more specific questions sent by the OECD secretariat. In particular, the country is asked to provide answers to a questionnaire, prepared by the EPI, the answers then being designed to provide the review team with basic information prior to the 'review mission'.[140] The majority of the background material comes from the country's environmental authorities, usually responsible for the coordination of the review, but includes also information from other relevant ministries and agencies, and often material provided by environmental NGOs.

After having studied the background material, the review team – made up of 5-8 OECD secretariat members and consultants, and 3-4 experts from other OECD countries – travels to the country for a one-week 'review mission'.[141] The team meets relevant authorities in thematic meetings on water and air management, nature conservation, integration of environmental issues into other policy sectors, environmental-social interface, international environmental cooperation, and a special issue chosen together with the country's government (environmental concerns in policy areas such as energy, chemicals, transport, health, agriculture, or tourism). The team also meets with NGOs, business representatives, and environmental researchers, usually in 2-3 hour sessions in which authorities are not present.

---

[140] Until recently (autumn 2004), this background information was provided in the form of a 'Country Memorandum', prepared by the government according to a list of 'Main Themes for Discussion', drawn up separately for each country, according to a standard format by the EPI.

[141] For large countries, the mission may be longer.

After the mission, the team drafts a preliminary review report, which is sent to the reviewed country and to the delegates of the OECD Working Party for Environmental Performance (WPEP) about a month before the final peer review meeting in Paris. The country sends its comments on the report to the OECD secretariat prior to the meeting, in which the WPEP group adopts by consensus the conclusions and recommendations of the review.[142] The changes to the text, proposed by the reviewed country, considered controversial by the secretariat, are highlighted in the draft document and discussed among the group. Only a few changes are usually proposed and accepted, and most of the time in the meeting is devoted to more general dialogue between the WPEP members and the reviewed country delegation.

About three months after the peer review meeting, the report is released in the reviewed country's capital, usually in a press conference and/or a seminar, with presentations and comments given by a high-level OECD official, the Minister of the Environment (sometimes other ministers attend as well), and representatives from others sectors of government, industry, NGOs, and researchers. The reviewed country itself decides how to organise the release event, disseminate the report, and follow-up the implementation of the recommendations. However, the OECD urges countries to draw up a report on the implementation of the recommendations 2-3 years after the report has been published.

5.4.3 Structure of the review report

The EPR reports consist of three parts. The first part, entitled 'environmental management', deals with the 'traditional' environmental issues, namely water and air management, and nature conservation (including protection of biodiversity). Many of the first cycle reviews also contained a chapter on waste management, which has however been dropped in the second cycle. The second part of the report deals with the integration of environmental concerns into sectoral policies, containing a chapter on environment-economy interaction, another one on environmental-social interface, and a third one on the integration of environmental concerns in a specific sector. The third part of the report is devoted to international environmental cooperation. Each chapter contains a

---

[142] In principle, the 29 countries not being reviewed can adopt a formulation or a policy recommendation even without the consent of the reviewed country in the peer review meeting, but in practice this seldom happens, because the objective is to give recommendations that are the most useful possible for the reviewed country.

description of the policy objectives, the state of the environment and the general problem situation, evaluation of performance, and a set of usually 5-8 policy recommendations.

5.4.4 Second cycle of reviews and its objectives

The second cycle of reviews was officially launched in November 2000, with the peer review meetings of Germany and Iceland,[143] and a special session recalling the main lessons learned from the first cycle. The launching of the second cycle had been preceded by a long process of preparation, including a number of WPEP meetings and seminars aimed at generating and gathering ideas from member countries, and preparing the new cycle. A survey was carried out in 1997 to learn from the experiences from the first cycle and to gather ideas for the second cycle among a large number of persons from all member countries having been involved in the EPRs (OECD 1997b).

The preparatory work resulted in basic guidelines for the second cycle adopted by the WPEP and EPOC (OECD Environmental Policy Committee) in 1999. These guidelines included the following salient points (OECD 1998a; 2003i, 3):

- to build on the recommendations to individual countries from the first cycle of reviews;
- to strengthen the performance orientation (including international commitments);
- to better cover sustainable development issues, including social/environmental issues);
- to be selective of sectors and issues with the highest relevance and be more country-tailored;
- to further develop and use indicators;
- to increase contributions by reviewed countries (including from stakeholders);
- to investigate ways to shorten the cycle;
- to strengthen co-operation with other OECD reviews, the UN-ECE and others;
- to stimulate greater accountability from Member countries; and
- to increase the influence of the reviews and derived products.

Later, it was decided that the EPRs would also be one of the main mechanisms for monitoring member countries' implementation of the 2001 OECD Environment Strategy for the First Decade of the 21st Century adopted by the OECD environment ministers in their meeting in May 2001 (OECD 2001j, 3). In practice, the follow-up of the Strategy has not decisively changed the review

framework. The recommendations of the WPEP for the second cycle have been implemented only partially and changes have been gradual.

As the half-way of the second cycle approached, the WPEP produced a mid-term review of the second cycle (OECD 2003i), noting in particular that the objectives established for the second cycle were very ambitious in view of the resources available at the secretariat. Especially the goal of shortening the review cycle seemed elusive, given that the programme already depends on significant extra-budgetary contributions from the member countries. The recommendations in the mid-term review therefore mainly stressed the need to improve the quality, focus and usefulness of the EPRs:

- concentrate more on *policy analysis* by focusing on fewer, *selected issues* for each country, with fewer, sharper and more country-specific recommendations;
- prepare a focused *questionnaire* to be answered by the reviewed country prior to the review mission so as to better target the factual and background material gathered for the mission;
- *shorten the report* to maximum length of 150 pages, structured around some five chapters, reducing the descriptive information and the number of issues covered in the report, while in turn increasing the use of new, more standardised indicators;
- *strengthen co-operation with other OECD review processes*,[144] transfer the OECD experience to other parts of the world through co-operation on reviews of selected non-Member countries;
- aim at livelier peer-to-peer *dialogue* and focus on just a few salient issues *in the peer review meetings* with less presentation of prepared material and less discussion on precise drafting questions; and
- strengthen the *involvement of stakeholders*, including different branches of government, EPOC (e.g. in the discussion of the review reports), countries own self-reporting (e.g. follow-up reports), and aim at *possibly shortening the review cycle*.

These recommendations represent a compromise between the often divergent views among the OECD countries. For instance, some countries advocate a focusing in the reviews on very few themes, whereas others underline the value of the broad overview that the reviews provide at

---

[143] The same countries were reviewed first in the beginning of the first cycle of reviews.

[144] This could achieved, for instance, through the participation in the EPRs by colleagues from other parts of the OECD, and possible follow-up of EPR recommendations in other OECD reviews.

present. The compromise solution was to reduce only gradually the number of themes covered, so as to retain sufficient continuity and comparability. In general, the lack of resources has been pinpointed as the main obstacle to changes, but it has nevertheless been argued that gradual changes have, indeed, taken place already. For instance, the number of themes on the agenda of the peer review meeting has been slightly reduced and some chapters in the review have been dropped (e.g. waste management chapter). Air management was not handled as a separate chapter in the Swedish review, and in the review of the Czech Republic, to be released in the autumn 2005, the 'traditional environmental themes' (water, air, nature) were for the first time merged into one single chapter.

The recommendations in the mid-term review by no means introduced revolutionary ideas compared to those of the paper preparing for the second cycle; they just gave the ideas a more concrete form. The objective of shortening the review cycle was put forward less forcefully. The only more or less new ideas were those of shortening the report by being more selective in the choice of themes to be covered, and the objective of making the peer review meeting more dialogical. The use of indicators and the desire to carry out more comparative analysis continued to be highlighted. The underlying belief is, of course, that these changes would increase the influence of the reviews.

5.4.5 Performance measurement or evaluation?

As the reference to performance in the name of the reviews indicates, the EPRs' main objective is to examine goal-achievement, instead of trying to understand the reasons for policy outcomes. Institutional issues are given a great deal of importance as well, but usually there is no systematic attempt to link the policy results to the underlying policy processes. Rather, the recommendations concerning institutional arrangements rely on agreed OECD principles of 'good governance' (e.g. the application of the pollution pays and user pays principles; policy integration; access to environmental information by citizens). Given the available resources[145] and the broad focus of the EPRs, it would indeed be difficult if not impossible to study the cause-effect relationships in detail. Consequently, the reviews can be classified as 'black-box' evaluations (see section 4.1.2 above), which nevertheless try to establish whether the elements of 'good governance' are in place and whether the policies are being adequately and effectively implemented and enforced.

---

[145] At present, a staff of four permanent administrator-level experts at the OECD secretariat carries out the reviews with the help of temporary consultants and country experts. The regular OECD budget covers three-thirds of the financing needed for carrying out the four annual country reviews, the rest coming from voluntary contributions.

5.4.6 Criteria of evaluation: goal-achievement, cost-effectiveness and policy integration

In principle, the OECD reviews represent a typical case of *goal-achievement evaluation* – the performance of a country is evaluated in relation to the degree to which it has achieved its own policy aims, objectives and targets. This is only a part of the story, however, a large part of the recommendations being based on other considerations than those relating merely to goal-achievement. Looking at the list of possible criteria for evaluating policy success (section 4.4.4.4), the other main criteria (beyond goal-attainment effectiveness) applied in the EPRs are cost-effectiveness, administrative effectiveness, transparency, and equity.

Cost-effectiveness is an explicit criterion in the EPRs, yet it is not applied directly, in the sense that the review team would itself calculate cost-effectiveness of policies. Instead, the team relies, on the one hand, on the commonly agreed upon OECD policy orientations, notably underlining the cost-effectiveness of economic instruments in relation to other policy instruments, and on the other hand, on information from the national authorities and experts as to the existence or not of studies on the cost-effectiveness of environmental policies.[146]

Administrative effectiveness is a major part of the evaluation, since a large share of the EPR recommendations concern policy implementation, urging the reviewed country to better implement and enforce its policies, increase resources to neglected policy areas, improve environmental policy integration in different sectors, etc. Each review contains a separate section or, in recent reviews, a whole chapter, on environmental policy implementation. Environmental policy integration, in turn, is the key theme of the second part of the reviews, called 'sustainable development'.[147]

---

[146] Two examples serve to illustrate the EPR approach to cost-effectiveness. The review of Sweden notes that the economic benefits of policies aimed at protecting particularly sensitive groups, such as children, from adverse health effects of exposure to chemicals "are not systematically calculated" (OECD 2004c, 150-151), and therefore recommends the country to "step up measures to meet environmental and public health objectives, with appropriate attention to cost-effectiveness and precautionary principle" (ibid., 25). As for Spain, the EPR notes that the recent National Hydrological Plan requires "rigorous and transparent cost-benefit analyses for all its water development infrastructure projects", as was recommended in the first EPR of the country in 1997, but wonders to what extent the main purpose of such studies has been simply to justify projects rather than help decision-making (OECD 2004d, 73). The review then recommends Spain to "improve recognition and understanding of the relationships between water and economic variables with: i) better data on expenditures, prices and financing…" (ibid., 21) and more generally, to "increase the use of economic analysis to improve the efficiency of environmental policies…" (ibid., 17).

[147] In the first cycle, the corresponding section was called 'integration of policies'. The contents of this part have remained largely unchanged since the first cycle – the main focus is still on environmental policy integration. The main innovation has been the creation of a new chapter on environmental-social interface.

Transparency lies at the heart of the 'EPR logic' in so far as the purpose of peer reviews is to generate 'peer pressure' through greater transparency of policies and governments towards their citizens. The reviews also evaluate the degree of transparency in the reviewed country, by looking at the availability of good quality information on environmental policies, and by examining public participation, access to justice, and access to environmental information by citizens, based on an OECD Council Recommendation.[148]

Equity has increasingly been addressed in the second cycle of reviews in the chapter on environmental-social interface, which examines, among other things, the distribution of harmful environmental impacts across the population and the incidence of costs of environmental policies between different population groups. Furthermore, the reviews have looked at issues such as the distributional impacts of green tax reforms, impacts of full-cost pricing of water, and the consequences of the privatisation of water management.

Comparison with other countries is a fundamental, yet controversial element of the OECD peer reviews. In the EPRs, comparisons that appear in the conclusions and recommendations are usually made with reference to the OECD average, or to the average of European OECD countries, whereas the main report uses indicators to compare the reviewed country's performance with that of 5-6 other, similar OECD countries. Care is taken to ensure that indicators are not used in abstraction from the country-specific context.

5.5 Evaluation of sustainable development and the OECD Economic Surveys

Before moving on to a more detailed analysis of the influence of the Environmental Performance Reviews, let us briefly look at another OECD peer review having the ambition to integrate environmental and sustainability concerns into sectoral policies, namely the OECD Economic Surveys. Thirteen Surveys carried out in 1999-2001 included a chapter called 'environmentally sustainable growth', whereas a full cycle of Surveys in 2001-2004 addressed sustainable development in a specific section of the review. Given that the Economic Surveys are probably most well-known and politically the most influential of the OECD reviews, at least in theory, they

---

[148] Recommendation of the Council on Environmental Information (3 April 1998 - C(98)67/Final). A frequent recommendation in the reviews is to urge countries to ratify and implement the UN Economic Commission for Europe (UNECE) Århus Convention on Access to Information, Public Participation in Decision-Making and Access to Justice in Environmental Matters, and to increase public participation, broaden communication between authorities and civil

can be expected to have greater potential to promote integration of environmental concerns in economic policies than the much less prestigious and less frequent EPRs.[149]

The OECD Economic Surveys are carried out by the Economic and Development Review Committee (EDRC) with participation of the reviewed country's permanent delegates to the OECD, sometimes assisted by civil servants from the capitals. The committee designates for each review two of its members to lead the discussion in the peer review meeting. The draft survey is prepared by the secretariat, the 'country desk' of the reviewed country at the secretariat carrying out the bulk of the work. The review process includes a high-level mission and one or two expert level missions to the reviewed country, during which the review team of 4-5 officials from the secretariat meets authorities in the reviewed country, to discuss on the basis of a list of questions sent to the government prior to the mission. The draft report prepared by the secretariat is dealt with in the peer review meeting, and further amended by the secretariat in consultation with the reviewed country government on the basis of the discussions during the meeting. Amended comments are then adopted by the EDRC in its subsequent meeting by consensus and published under its responsibility. The reviews were previously carried out with 12 to 18 months interval, but with the recent change of focus from macroeconomic to structural policies, the frequency was reduced to between 18 and 24 months.[150]

The Economic Survey process therefore significantly differs from that of the EPRs, the main differences including the following:

- The review cycle is considerably shorter (18-24 months as opposed to 7-9 years for the EPRs);
- The review team is smaller and does not include country experts;
- The OECD 'country desks' who follow the reviewed country's economic policies on a daily basis carry the main responsibility for drafting the review;

---

society, encourage citizen initiatives, and improve the systems of environmental information and indicators (e.g. OECD 2004c, 23; 2004d, 26).

[149] In line with the mandate given to the OECD by the 2001 Council ministerial meeting to develop indicators of sustainable development, the Economics Department was charged with the task of developing a framework of analysis of sustainable development, to be tested in one cycle of Economic Surveys between 2001 and 2004.

[150] According to a civil servant from Finland's Ministry of Finance, this shift of emphasis was largely a result of pressure from the Anglo-Saxon member countries – the United States, the United Kingdom, Canada, Australia, and New Zealand, which have been in the forefront in opening the public sector for competition (OECD ärsyttää… 1998).

- The secretariat has an important role in the peer review meeting, whereas in the EPRs, the secretariat intervenes only to clarify factual issues;

- The discussion during the peer review meeting is less pre-structured and involves more free-flowing dialogue than in the EPRs;

- The members of the EDRC usually come from their countries' permanent delegations to the OECD, not from the capitals, as the majority of the WPEP delegates;[151]

- The reviewed country officials have a greater influence on the amendments to be made to the draft survey. The assessment and recommendations of the Survey are discussed in the meeting between the reviewed country and the secretariat on the day following the peer review meeting, whereas the corresponding meeting in the EPRs only concerns factual changes to be made to the main report, the conclusions and recommendations having already been adopted by the WPEP in the peer review meeting.

The theoretical framework developed by the Economics Department under the guidance of the EDRC for analysing sustainable development in the Economic Surveys consisted of a set of seven sustainable development themes:

- greenhouse gas emissions
- air pollution
- natural resource management
- water pollution
- waste management
- sustainable retirement incomes
- improving living standards in developing countries

Each Economic Survey then included a short (usually about six pages) sustainable development section covering three of the seven issues, chosen in function of the conditions of the country being reviewed. For each of the seven themes, a range of indicators has been used to measure performance. Since no consensus was reached in the Ad Hoc Group on Sustainable Development on a list of indicators to be used in all reviews, the Group agreed to proceed on the basis of the list, leaving open, however, the choice of specific indicators in each review. The choice of the indicators to the 'long list' was driven by practical considerations, focusing on issues that 1) have a strong degree of permanence and inter-generational concerns, and 2) are analytically tractable (OECD 2002b, 5). In particular, only a limited set of issues allowing meaningful comparison across countries could be addressed in the reviews over the 2001-2004 period. Finally, all indicators focus

on performance, thus excluding indicators of 'policy settings', on the grounds that "different policies will be appropriate under different national conditions" (OECD 2004a, 16). The seven issues therefore do not even intend to be an exhaustive list, but rather represent themes that are common to the policy agendas of all member countries, and that could be addressed in the peer reviews. (ibid., 11.)

Obviously, the list of seven sustainable development issues seems rather arbitrary, consisting of mainly traditional environmental themes – 'end-of-pipe' policies – and two social issues, one of which is acutely relevant for the developed countries, the other concerning the developing countries. The OECD Environment Directorate provided highly critical comments to the draft versions of the framework, but these did not lead to significant changes in the approach.[152] The delegates in the WPEP group have likewise expressed their dissatisfaction with the framework, and the genuine willingness of the Economics Department to carry on with sustainable development issues has been questioned. Indeed, the OECD Economics Department made it clear enough that it would continue analysing sustainable development – which it considered mostly of marginal relevance to the major structural and macroeconomic policies – only as long as it would receive funding specifically allocated for this task. Therefore it was not a surprise that the EDRC decided to discontinue the 'sustainable development experiment' in the Economic Surveys after 2004, once the specific funding for this purpose ceased to come in.[153] Environmental issues – notably those having significant economic implications such as climate change – are likely to figure among themes treated in future Economic Surveys as one of the potential 'structural issues'. Instead of having a in the reviews a specific sustainable development section, the EDRC has decided that each survey would deal with at least one environmental theme.

After this description of the context within which the EPRs are carried out in the OECD, the next chapter moves on to the analysis of the review programme, asking which types of impact the EPRs generate in the reviewed countries.

---

[151] For example, in the Netherlands, a civil servant from the country's permanent OECD delegation participates in two-thirds of the reviews, a delegate from The Hague taking care of every third meeting.

[152] The approach has generated criticism also from beyond the OECD 'insiders' (e.g. Caccia 2003).

[153] This decision also quickly dissipated the often-expressed worry that evaluation of sustainable development by the OECD would become merely a secondary issue as a part of the Economic Surveys and that the Environmental Performance Review programme would be simply terminated.

Chapter 6

Do the OECD Environmental Performance reviews have influence?

The Environmental Performance Reviews aim to influence OECD countries' policies by judging progress so as to promote dialogue and accountability. In this chapter we shall apply the framework developed in section 4.3 to identify the main types of influence that the EPRs engender on policies and actors in the reviewed countries. The chapter describes individual level effects, the processes of deliberation, dialogue and argumentation, and the consequences on policies. Both the effects stemming from the review process and from the review report are considered.[154]

The notes and the general experience from the WPEP meetings since May 1996 and the rather extensive survey of the opinions concerning the first cycle of reviews (OECD 1997b) provided the basis for the identification of the key issues concerning the review influence. This background material was analysed in the light of the different typologies of evaluation influence identified in the literature, in order to prepare more specific questions for the interviews and to elaborate the typology of review influence presented in section 4.3. The interviews were at this stage analysed primarily from the 'fact perspective', i.e. as descriptions of the reality that constituted the main subject of research interest. They were used, on the one hand, to verify the interpretations based on the background material, and on the other, to provide further insights into the types of review influence. The analysis of the interviews entailed, first, classifying the responses into the categories of review influence identified in the literature, and, at the second stage, going through the

---

[154] Direct citations from the memoranda written after the interviews are used to illustrate and substantiate arguments put forward in the text.

interviews with an eye on any impact types that had not been recognised in the literature. Combining insights from the background material and the interviews allowed me to get an idea of the relative importance of different types of influence. However, the main intention was to obtain the broadest possible range of ideas on the possible types of influence – both real and potential – instead of attempting a quantitative analyse of the respective importance of the different types of influence.

6.1 Impacts at the individual level

The effects of the EPRs on individual actors are a necessary, yet not alone sufficient condition for broader, collective level influence. This study primarily focused on collective type of impacts, but some effects at the individual level were identified as well. First, in the reviewed countries, the persons organising the review learn a lot throughout the process.[155] Others who participated in the meetings during the mission, usually did not consider having significantly learned from the process, and some had even forgot having participated in the review a few years earlier. The reports, on the other hand, seldom provide new findings or ideas to the environmental authorities, largely because they are essentially based on information provided by the environmental administration, and the issues in the report have already been addressed in the discussions during the review mission. Therefore, the reviews seldom generate ***change in attitudes***. By contrast, for actors in the non-environmental sectors, and for foreigners interested in environmental policy in the reviewed country, the reports are more likely to provide genuinely new information.[156] By presenting a good overview of a country's environmental policies, the review offers useful background material for those wishing to quickly obtain a general picture of the country's environmental policies, or to learn about more specific instruments and approaches used by other countries. The review may confirm existing beliefs or lead to rejection of the findings. Nevertheless, outright rejection is not very common, because the recommendations are designed to be as useful as possible to the reviewed country's environmental authorities, tend to be based on generally accepted OECD policy principles, and are adopted in the OECD WPEP group by consensus. The reviews do influence ***salience***, by elevating the status of environmental issues, especially among actors in the non-

---

[155] *"Personally, I have of course learned a whole lot during the process."* (a person responsible for organising the review of his country)

[156] *"For us, the review has not only the value of providing us arguments and helping us by giving more weight for our own arguments, but also that of helping us learn a lot on the environmental policy of our country, and on the ways to improve it. For the persons working on specific themes, the OECD review probably does not reveal much novelties, but for us, who have to work across a broad range of policies and thematic areas, the report contains plenty of useful information."* (civil servants from a non-environmental sector)

environmental sectors, but also in the environmental field in so far as environmental authorities attach more importance to issues raised by the OECD.[157]

The review process also 'forces' civil servants, even those from non-environmental sectors, to *elaborate* on environmental issues, and form opinions on issues at stake.[158] However, this influence is reduced in so far as the participants in the review process are mainly environmental experts in their respective organisations and therefore already knowledgeable on matters handled in the review.

Two groups of individuals are the most likely to benefit from the review process. Firstly, together with his/her closest collaborators, the coordinator of the process in the reviewed country – often the WPEP delegate – learns a great deal about the country's environmental policies, as well as about the OECD and the 'EPR logic' and peer review method, and gains organisation, cooperation, and negotiation skills. The extent of such learning depends on the individual's personal background, and the country's policy tradition – in particular, whether such cooperative processes are a novelty or a part of ordinary day-to-day policymaking. Of course, the WPEP delegate is usually rather familiar with the process from having participated in the WPEP meetings, but the more intensive involvement as an organiser of a review or as a national expert in a review team helps understand the review logic – with its opportunities and limitations – on a more concrete level.[159] Secondly, the *OECD review team* members learn about the substance of environmental policies in the OECD countries, and, naturally, about the organisation of reviews. The OECD secretariat can therefore be considered as the 'institutional memory' of the review programme.

---

[157] For example in the Netherlands, the authorities noted that the review helped revive some policy proposals that had already previously been rejected. *"Per km-tax had, about a year prior to the review, been rejected by Parliament, but now the issue is surfacing again. It is difficult to say whether the OECD recommendation to this effect may have any impact. The main impacts are probably that we need to do something about increasing congestion on the roads, and in the current situation there is no money for greater transport investments (road construction to alleviate congestion), so the tax has again become an attractive option. Moreover, the present political setting seems more favourable to the adoption of a tax."* (a Dutch civil servant)

[158] *"For us, as high-level civil servants, the review has the important function of allowing us (in fact obliging us!) to stop, carry out a general evaluation, think what we are actually doing, etc. That's something we would never have the time to do, if the review did not exist."* (a Portuguese civil servant)

[159] Such learning may concern the OECD role in general, helping understand what the organisation can actually do, which are its strengths compared with other evaluators, and what are the key objectives and characteristics of peer review. On a more concrete level, participating in the process makes the national delegates aware of the resource limitations that dictate what is realistically achievable in the EPRs. The officials responsible for the review in the OECD Environment Directorate of course seek to ensure that the national delegates have such general understanding.

6.2 Collective processes of deliberation, dialogue, and argumentation

While the reviews may sometimes engender significant individual learning, the main promise of the EPRs is their potential to promote collective learning and accountability through deliberation, dialogue and argumentation. The review process and the debate following the release of the review report are two ways of enhancing such phenomena.

6.2.1 Review process

Three stages in the review process have the capacity to generate impacts: the preparation of the reviewed country for the OECD team's 'mission', the discussions among participants during the mission, and the debate in the peer review meeting. This process in its various phases involves acts of legitimisation, persuasion, critique, and defence on the part of the various participants.

The ***processes of preparation in the reviewed country*** for the mission of the review team as well as for the peer review meeting were perceived by many as important sources of review influence, since they provide occasions for mutual learning among the various participants in the reviewed country.[160] Drafting the responses to the questionnaire received from the OECD secretariat, as well as the process of 'briefing' civil servants for the meetings with the OECD mobilise a large number of individuals from different sectors of administration to agree on common positions on the issues to be addressed in the review. A number of participants (at least in Japan, Mexico, the Netherlands, and Portugal) thought the first review of the country had been more important than the second in enhancing cooperation and mutual understanding.[161] The Mexican Ministry of the Environment

---

[160] *"In the course of the preparations for the review, we made a significant effort to come into an agreement on what kind of an image we wanted to transmit from our country to the OECD."* (a civil servant organising the review in his country)

"The question of economic instruments was a case in point: we (in the Ministry of the Environment) consider the lack of economic instruments as a major weakness of our environmental policies, whereas the finance ministry is reluctant to introduce them. At a certain point during the process I believed we would fail in finding an agreement with them [the Ministry of Finance] on a common position, but in the end, we succeeded." (a civil servant organising the review in his country)

[161] *"The process of preparing for the first cycle review was a substantial effort, involving civil servants across the government. We asked experts from different agencies to draft sections for the country memorandum, and we sometimes had to send the draft back for revision twice, if we were not satisfied with the proposed text. In difficult cases we even had meetings with the experts in order to explain to them the real character of the EPR and the expectations of the OECD towards us. Before the review mission, we again had meetings with civil servants in order to 'brief' them on the character of the review, on the possible questions and explain what was expected of them. This process required of course a lot of effort and resources, but it was also a powerful tool of 'mainstreaming' environmental policies. By contrast, the second cycle review was, in my view, organised in too centralised a manner, essentially by a small team of a handful of civil servants, headed by the vice-minister of planning."* (a WPEP delegate)

used the country's first EPR process (1997-1998) as a tool for strengthening the involvement of the NGOs in environmental policy, which was one of the priorities of the Minister at the time. However, the NGOs were somewhat disappointed at not having a greater say in the process, partly because they did not fully understand the review was not carried out by the government, but by the OECD.

Similar argumentative processes could, in principle, take place during the ***review team's mission*** to the country, but the set-up of the mission does not always encourage genuine dialogue, partly because of time limitations.[162] The mission has above all the ambition of ensuring the review team can fill the information gaps and answer the open questions that remain after it has studied the background material. The main preoccupation of the reviewed country government, in turn, is to ensure it gets its message across, and, in several cases (especially for the sector ministries), to defend itself in advance in the face of potential criticism.[163] The review team does its best to find points of criticism in the country's policies, but often this is difficult, as demonstrated by the numerous comments evoking the ease at which the team members could be mislead, so as to give an 'embellished' picture of the country's policies.[164] Since various authorities are present in the meetings at the same time, the 'social control' among the national civil servants sometimes makes it more difficult for the review team to get behind the official positions.[165]

In Portugal, the timing of the reviews – both in the first and the second cycle – right after the country's EU presidency facilitated the process, since there was already a team ready to take up the challenge. While this probably facilitated the preparatory work, it may not have been the most effective way of maximising the review influence across the administration and especially among actors previously not well integrated in international cooperation.

[162] *"The meetings during the mission were hardly useful in other ways than perhaps in providing another possibility for stakeholders (e.g. researchers) to build networks with other actors in the field."* (a Dutch civil servant)

"Five days was too short a period for dialogue between the national delegates and the OECD team, and there were too few national experts in the review team. Yet, the field of environmental policies is much broader than that of [our sector]. The review did not provide us sufficient time to make sure the review team had the correct facts available." (a civil servant from a non-environmental sector)

[163] *"[The reviews] …also represent the danger that pressure groups (associations, local politicians, etc.) seize the opportunity to use the reviews in order to criticise us and in this way render our work more difficult."* (a representative of a sector ministry)

"For instance, during the OECD EPR mission, Greenpeace representatives showed an indicator study that incorrectly showed [our country's] environmental performance was among the poorest in the OECD. It took me a long time to explain to the team what was the reality in the matter." (a civil servant organising his country's review)

[164] *"A problem in the reviews is that national civil servants conceal issues, especially in countries such as Portugal and Italy […] It is easy for the civil servants to mislead the national experts in the review team – the experts from the OECD secretariat are much more capable of avoiding being mislead, since they have more time for preparing themselves for the review."* (a WPEP delegate with experience from reviews as a national expert)

"The meetings with the OECD team were easier than I had imagined. I had expected more detailed questions and thought the group would have been better informed about our policies. For instance, they did not ask us about the disparity in investments between road and rail transport, even though an answer to this question would have easily revealed the lack of integration between sectors. I did not conceal facts, but I gave a greater emphasis to aspects that were positive for us, and that I was most familiar with." (a civil servant from a non-environmental sector)

[165] However, the secretariat encourages the team members to organise during the mission also informal meetings with civil servants and experts they have identified during the official meetings as key informants.

On the other hand, many interviewees considered the review team's visit as a valuable opportunity to reflect upon the country's policies in a new light, from an 'outsider' perspective.[166] The increased visibility of environmental issues among stakeholders, notably the non-environmental sectors of administration, was frequently mentioned as the main pathway of influence.[167] An important part of such enhanced visibility consists of meetings with high-level officials and ministers. The political impact of this 'ceremonial' part of the reviews is probably more important in the new OECD countries than in the old ones.

The ***peer review meeting*** is, in principle, a key event for dialogue and argumentation, providing the WPEP delegates with an opportunity to pose critical questions concerning the reviewed country's policies, for the reviewed country to explain and defend its policies (or admit the shortcomings, which is frequently the case[168]), in a non-adversarial atmosphere laying ground for mutual learning from experience and best practices. The discussion is structured so as to elicit critical comments and questions from the other countries' delegates, and defence on the part of the reviewed country. At the end of the discussion on a cluster of issues, the group agrees upon the wordings in the relevant part of the conclusions and recommendations of the review.[169] However, while there is a general consensus in the WPEP group on the need to further reduce the number of items discussed in order to reach a more genuine dialogue (OECD 2003i), the opinions on the usefulness of the event vary from strong scepticism by those who perceived the meeting as a long and useless 'theatre play' with

---

[166] *"The question-response method was particularly useful precisely because it pushed us towards a more profound reflection. For our civil servants it was highly useful to be forced to call into question their own ways of thinking and working. For instance, the questions of the Canadian expert clearly showed that he had experience of problems almost identical to ours, but which have been solved in a very different manner in his country. This way, by providing a new perspective to our own work, he helped us better understand the issues at stake."* (a French civil servant)

[167] *"We made a big effort to use the first EPR **process** as an instrument of attracting more attention from other ministries and stakeholders towards environmental issues, and using the prestige of an 'official OECD review' in order to raise the status of environmental issues. We thus made a '110 % effort' to obtain a maximum impact from the visits during the review mission. Indeed, the results were very positive, and environmental issues received more attention from other ministries."* (a person who coordinated his country's first EPR)

"The review process also provided us with an 'excuse' to reach the 'difficult' actors, such as the national […] agency... The name of the OECD was essential in enabling us to reach these powerful actors, who most likely would have been much more reluctant to move had it been just the ministry itself asking them to do a favour and provide information." (a civil servant from the environmental administration)

[168] A case in point was the peer review meeting of the UK, in which the country's then environment minister, Michael Meacher, openly accepted the criticism expressed by the other countries' WPEP delegates, without trying to conceal facts that might have shown the country's environmental performance in a less favourable light. The same kind of attitude towards the review findings seems to prevail in the country as the follow-up report on the implementation of the recommendations is being prepared (personal communication, UK WPEP delegate, 18 May 2005).

[169] After having received the comments by the reviewed country to the draft document of conclusions and recommendations, the secretariat drafts a new version of the document, which displays in bold letters those of the changes proposed by the reviewed country that the secretariat considers potentially acceptable, further dividing them

prewritten dialogues, to highly satisfied views emphasising the positive and open atmosphere promoting mutual dialogue and understanding. The debate is usually very cordial, and the reviewed country delegation seldom has difficulties to fend off the criticism. Moreover, the debate does not seem to reach very far beyond the peer review meeting to the actual policy discussions in the country.[170]

Some Dutch interviewees considered the meeting more or less futile, since the overall setting (including time limitations) did not allow real dialogue. Others noted that the peer meeting during the first review was indeed highly appreciated and several ministries sent their representatives to the country's delegation. Yet other participants considered even the second cycle meeting as useful and interesting, as it offered a possibility to discuss the apparent retreat of the Netherlands from its long-standing position as a 'forerunner' in environmental policies.

The ministers of the environment, who have led their country's delegation in peer review meetings, have often perceived the event as a positive occasion to get a broad overview of the country's policies in a relaxed and non-confrontational atmosphere.[171]

One may of course ask how big a role the peer review meeting actually plays for the review process as a whole in promoting accountability and policy learning. It probably has mostly an indirect impact as a 'PR event' for the OECD and the EPRs. The meetings should provide learning opportunities both to the WPEP delegates and the members of the reviewed country delegation, in order to maintain the delegates' interest in the review, to ensure they pass on a positive picture of the EPRs to their superiors, and to motivate the reviewed country delegation to carry on with advocating the review at home, implementing the recommendations, and following up their implementation.

---

into relatively uncontroversial ones and those that require more substantial discussion by the group. Such major issues are usually very few; in the U.S. review there was only one such proposal.

[170] The interviews showed that even civil servants working in the same organisation with a person having participated in the peer meeting were not always aware of the discussions conducted during the meeting.

[171] While the positive comments by the French, Swedish, and the UK environment ministers directly after the review meetings cannot be taken at face value, since they can be partly attributed to sheer politeness, the French and Swedish civil servants confirmed this favourable impression in the interviews and 'discussions in the corridors'. Moreover, these ministers seemingly enjoyed the atmosphere of the peer review meeting, and appreciated the possibility of presenting to the OECD delegates a broad picture of the country's environmental policy, and obtaining such an overview themselves.

Box 6.1. Dilemmas of the peer review meeting.

The peer review meeting: diplomatic conversation or critical examination?

The general atmosphere and character of the peer review meeting reflects the dilemmas and divergences of views that the EPR programme is constantly faced with. The character of these meetings is between hard negotiations on specific wordings, and general 'brainstorming'. On the one hand, the interviews and the discussions in the WPEP group have shown that many of the participants appreciate the possibility provided by the peer review meetings for a frank dialogue on 'hot' and controversial issues, without the need to enter into confrontation and a 'negotiation mode'. This perspective was clearly present in the peer review meeting of the United States, in May 2005. Before the event, both the OECD secretariat and the U.S. delegation were highly concerned that the meeting might take a confrontational character, for instance if one of the early questions by the WPEP delegates would have been strongly critical towards the American policies. On issues that are politically sensitive in the reviewed country, or high on the agenda of international politics such as climate change, discussion on the wordings of the 'conclusions and recommendations' document may momentarily approach a negotiation setting, and create tensions. In the U.S. review this did not happen, and on the contrary, the meeting was conducted in an atmosphere even more consensual than on average.

On the other hand, many participants considered that the lack of confrontation and genuine dialogue reduce the relevance of the meetings. In principle, all delegates are free to express their views more or less freely, but in reality, the meeting follows a very carefully defined structure and schedule, which hardly leaves room for spontaneity, improvisation and innovation.[172] Even the questions by country delegates are partly pre-programmed: to make sure that at least some questions are addressed to the reviewed country on each issue on the agenda,[173] the secretariat often contacts national delegates in advance and asks them to pose a question on a certain agenda item. Likewise, a number of WPEP delegates have, at least privately, expressed their frustration about having the feeling of neither learning through dialogue, nor being able to influence the outcome of the meeting. The participation of ministers in the peer review meetings is a double-edged sword: on the one

---

[172] For example, a Dutch civil servant argued *"[t]he peer review meeting was rather useless: no real dialogue, no follow-up questions, no possibility to dig deeper into the most relevant questions; no opportunity for learning from good experiences; and too many items on the agenda to permit real debate."*

[173] The so-called Main Items for Discussion document, which contains a total of about ten broad themes, on each of the three parts of the review – environmental management, sustainable development, and international cooperation. In practice, the agenda hardly excludes any issues treated in the report, and the delegates are entitled to ask questions on practically any theme they wish. The document therefore serves mainly as a general guide for the discussions.

hand, their presence gives more political weight and visibility to the EPRs, and strengthens the commitment of the administration to the process. Ministerial participation may also enable a more frank debate, because in spite of being ultimately tied by the government's policy line, the minister can express himself more freely than a civil servant on issues of political importance. On the other hand, it is easier for the WPEP delegates to argue against high-level civil servants than against ministers.[174] Having the minister heading the delegation may also transform the meeting to a 'one-man-show', in which the rest of the country's delegation is frustrated by not having the possibility to participate in the discussion.

More active participation of the secretariat – as in the peer review meetings of the Economic Surveys – could improve the quality of debates. The discussions in the economic peer review meetings flow more freely, some participants often being considerably more active than others. Of course, imitating this process would have its downsides. New members of the WPEP might find it more difficult to become integrated, some members of the reviewed country delegation might feel frustrated if no questions were asked on the area of their competence. Above all, making the meetings more dialogical would require a lot from the WPEP delegates, given the broad scope of environmental policies covered in the reviews.[175] Focusing the meetings more clearly on a few specific topics would allow countries to send to the meetings experts on the themes on the agenda.

A closely related problem is that of 'unanswered questions', i.e. situations where the reviewed country delegates are allowed to ignore some questions, or give answers that hardly have anything to do with the question. The chairman has a crucial role in making sure that delegates can in such situations pose follow-up questions. In practice, the tightness of the agenda limits the time available for such dialogue. Moreover, the problem is also on the other side of the table: the questions must be sharp and relevant if good answers are to be expected.

---

[174] A Portuguese interviewee argued that the chapter on nature conservation in the EPR report gave a too positive picture of the country's performance, precisely because the WPEP delegates did not dare press harder the Portuguese minister during the meeting.

[175] In the Economic Surveys the range of issues covered is more limited, and above all, the delegates can rely on a common body of economic theory, which provides a framework for the analysis and for the discussions. Moreover, the reviews are conducted at a short interval, which means that the delegates meet each other in peer review meetings at a much higher frequency than their colleagues in the WPEP.

6.2.2 Processes following the publication of the review report

Dialogue, *deliberation* and *argumentation* generated in the reviewed country constitute the main pathways through which the EPRs exert their influence once the report has been published. The ordinary way of dissemination in the reviewed country is to release the review report in a press conference, usually followed by or combined with a public seminar, where various stakeholders are invited to comment on the review. With the exception of Japan, the Minister or Secretary of State for the Environment has hosted the release event in all countries reviewed during the second cycle.[176] As a rule, the release events and press conferences have resulted in a number of articles in written press, and often some TV and radio interviews. Other means of dissemination have included articles written in professional journals and the environment ministry's information newsletter, debates in Parliament on the basis of the report, and communications to Parliament on the actions taken by the government to implement the recommendations. (OECD 2005a, 19-20). Compared with the OECD Economic Surveys, the EPRs usually generate rather modest media attention, and the journalists seldom go beyond repeating the main points of the OECD press release, towards more analytic and critical journalism.[177] The results of the review are most often presented following the logic that governs the EPR reports themselves: the country has made progress in certain areas, but, like in all OECD countries, a lot remains to be done. In the Netherlands, many articles reported as the main finding that the country no longer is a leader in environmental policies in international comparison, whereas in Portugal the reports were more positive, pointing out that the country had 'passed the exam' of the OECD. In Italy, in turn, many of the numerous articles[178] were more critical, pointing out especially the problems of urban air pollution. Many informants for the present study noted that the public debate following the event soon died out, once new policy issues took over the media agenda.

Most frequently the environmental authorities exploit the OECD prestige by referring to the review findings and recommendations so as to **persuade** other sectors to adopt certain policies. The environmental authorities in the first place, and other government sectors to a lesser extent are the most frequent users. Typically, the ministers of the environment use the recommendations in a

---

[176] According to the Japanese interviewees, press conferences are not a part of the political tradition in the country. In Chile, the report was presented to the President, and in Mexico, as many as nine ministers participated in the release event.

[177] This evaluation is based on a review of articles published in Italian, Portuguese and Swedish newspapers right after these countries' release events.

[178] The selection of press articles that the OECD secretariat compiles after the release of each EPR report consisted of 15 articles published in Italian newspapers the day after the release event.

rather selective manner, tactically, drawing attention exclusively to those recommendations that serve their own interests, overlooking the broader picture.[179] The environmental administration, in turn, uses the recommendations in the negotiations with its counterparts in the other ministries, for instance to persuade the Ministry of Finance to adopt specific environmental taxes or broader 'ecological tax reforms'.[180] The experience from parliamentary debates has been mixed. In Sweden, the Parliament environmental committee carried out a lively debate, whereas in Hungary, despite the attempts by some parliamentarians to use the report to criticise the government, the debate was lame. In France, the environment ministry expects to use the report to 'educate' the parliamentarians.[181]

An equally common practice is to refer to the reports in order to explicitly *legitimise* environmental actors or policies.[182] In Portugal, the Minister of the Environment clearly used the report to point out that even the OECD approves of the country's waste management policies, and recognises the success of the government's nature conservation efforts. In the Netherlands, this kind of use was less evident. In a sense, the review was used to legitimise the country's environmental policies and authorities, since according to the Dutch environmental authorities most of the recommendations concerned measures that the government was already implementing or was about to implement in the near future.[183] However, one Dutch interviewee contested this view, arguing it was "a very Dutch way of avoiding further action". Be that as it may, the review did not generate considerable public debate in the country.

The recommendations were much more seldom used to *criticise* the activities of any group of actors. In Norway and Mexico, the new government used the review to criticise the old one or to

---

[179] The Portuguese interviewees argued that three issues clearly stood out as the ones that the Minister of the Environment paid the greatest attention to: the country's waste incineration policy, institutional arrangements for water management, and nature conservation policy. In France and in Germany, the environment ministers wanted to publish the review recommendations and conclusions early on, before the whole report was published, so as to maximise the support the review could provide for the introduction of a 'green tax reform'.

[180] *"I remember only informal use of the EPRs, e.g. in discussions with colleagues of the Ministry of the Environment on pricing mechanisms: '…the OECD supports the idea of a […] tax; why doesn't the Ministry of Transport introduce one?'"* (a civil servant from a non-environmental sector)

[181] *"Parliament is a priority target for the report, because there are still a large amount of parliamentarians that can be considered as 'enemies of the environment'. A message coming from a source as prestigious as the OECD is more likely to pass among these parliamentarians than many other similar reports."* (a civil servant from the French environment ministry)

[182] This type of use was recognised even in the Netherlands, where the review was considered by all interviewees of very little use: *"Our State Secretary used the report in his various speeches during the couple of months after is publication, in order to gain support for his own policies; after that, other issues came up and we were busy with them instead of the ones raised up in the report."* (a Dutch civil servant)

[183] The Secretary for the Environment submitted to Parliament a communication containing this message.

provide a platform for its own policies.[184] Such criticism would be a typical task for the NGOs, but they seem to prefer other sources of support and legitimacy in order to substantiate their arguments.[185] In the Netherlands, the country's own 'Environmental Balance' and 'Nature Balance', produced by the the National Institute for Public Health and the Environment (RIVM), are frequently used by the NGOs, arguably because of their more confrontational approach, although the OECD report may in this context be used to give further strength to the arguments. One Dutch interviewee argued to the opposite: the RIVM reports fail to have influence because they lack the balanced 'carrot-and-stick' approach characteristic of the EPRs.[186] The Portuguese NGOs argued that the EU publications (notably those of the EEA) are more accessible and useful than the OECD reports, whose existence the NGOs are not always even aware of. The Canadian civil servants considered, by contrast, that the NGOs and researchers do in fact use the review to criticise and put pressure on the government. In France, the environment ministry representatives wished to encourage the NGOs to use the reviews in this manner.

The fourth way of using the recommendations – ***defence*** against the criticism voiced in the review – is certainly present, but was seldom evoked in the interviews. The non-environmental sectors are more prone to adopt a defensive attitude than the environmental authorities.[187] In the Netherlands, the most common way of the government to respond to the EPR – to affirm that 'this is what we are doing already' – can be seen as an example of skilful defence against potential criticism. The debate following the release of the report in the reviewed country usually does not involve major confrontation. At least two factors may explain the relative absence of defensive claims in the debates. For the first, the implementation of the recommendations being voluntary, the authorities can focus on those recommendations that the government has already decided to put into effect, or is about to do so. Secondly, the recommendations have a rather lenient tone, they hardly generate

---

[184] Interview with Veronique Deli (see annex I) and personal communication from Gunnar Farestveit, 9 January 2002.

[185] In Portugal, the NGOs complained about the prohibitive price and poor dissemination of the review report. In the Netherlands, it was suggested that the NGOs did not believe the report would help to persuade the government to change its policies anyhow. While no systematic survey was made of the use of the EPRs by the NGOs, the evidence from the interviews with the civil servants suggests that such use is not common. Interestingly, a Portuguese environment ministry official said the industry does not use the country's EPR report, because it considers the report to be too lenient towards the environmental administration, and believes the administration was able to mislead the OECD team.

[186] This divergence of views probably stemmed from the different expectations of these individuals concerning the review influence – an issue we shall address later on in chapter 7.

[187] *"In general, the more economically oriented ministries (Finance, and those having to do with resource use) were more eager to give a favourable picture of [our country's] policies, whereas in the environmental administration, people were rather interested in having the OECD criticise us so that we could use their critique to give credibility to our own arguments."* (a WPEP delegate that organised his country's EPR)

any stronger criticism, and consequently, the authorities are not forced to defend themselves. The EPRs hardly provide potentially 'explosive' material – unlike for instance the OECD-PISA education assessments – that would decisively affect the public debate and the country's environmental policy agenda. Rather, the recommendations are merely used to strengthen arguments in debates that were already being carried out prior to the review and support policies that are still in their infancy.[188] In the peer review meetings, defence by the reviewed country delegation is, in principle, an essential part of the game, but, as noted above, the debate practically never leads to real confrontation, and instead of genuinely defending itself by answering to critical questions posed by the WPEP delegates, the reviewed country delegation can sometimes simply eschew the real questions.

The governments – mainly the environmental authorities – frequently use the report or its recommendations as **background material** in negotiations, or in drafting environmental or sustainable development strategies and programmes.[189] The function of the reviews in such cases is to strengthen and legitimise arguments by 'certifying' them with an OCDE 'quality label'.[190] The reports are likewise used as reference material – by civil servants in their daily work, by researchers, and by foreign experts wishing to obtain concise background information on the country's environmental policies. Moreover, many evoked the process of preparing the review, especially the drafting of the Country Memorandum, as a welcome opportunity, triggered by outside pressure, for the Ministry of the Environment to renew its information material. Often the Country Memorandum is published in the publication series of the ministry or otherwise distributed within and beyond the government. The Economic Survey teams used the EPRs as background material when drafting the sustainable development sections, even though some WPEP delegates

---

"It is true that the sector ministries that were main targets of the OECD criticism adopted a rather defensive attitude towards the review, trying to anticipate likely criticisms so as to attenuate their impact later in the process." (a civil servant from the Ministry of the Environment)

[188] *"The review did not provide novel ideas, but in the informal discussions with the head of the review team, I constantly underlined the importance of certain issues […] I did not directly propose recommendations [that the OECD team could give us] but we simply discussed the importance of certain questions. In this manner, I tried to use the review as an occasion to solidify, and affirm positive […] individual experiments that needed support in order to turn into real policies."* (a civil servant organising his country's EPR)

[189] For instance, the Swedish Government Bill on Environmental Objectives (Regeringens proposition 2004/05: 150), published less than a year after the EPR report, systematically refers to the EPR recommendations to support its proposals. In Japan, the recommendations were used when drafting the government White Papers on the environment and the recycling society. Similar, even though less systematic, use of the report was mentioned, among others, by the Greek, Hungarian, and Slovak WPEP delegates.

[190] *"In the public debate, the 'OECD label' is important. It proves that the arguments come from an independent international organisation. It would be much harder to make pass the same message if it came from the Ministry of the Environment."* (a French civil servant)

have expressed their dissatisfaction with what they considered as poor usage or outright ignorance of the EPRs by the Economic Survey teams.[191]

6.3 Consequences on policies

Finally, if the EPR process and the release of the report can generate impacts at the level of individuals, and engender collective processes of dialogue and argumentation, we must still ask what are the ensuing effects on policies. These policy consequences may be new decisions and actions; new shared understandings; enhanced or decreased legitimacy of actors and policies; creation of networks and strengthening of policies; diffusion of policies beyond the reviewed country; and changes on the political agenda in the reviewed country. The reviews may also have the perverse side effect of engendering resistance and rejection.

The research approach adopted in the present study did not allow an examination of the long-term impacts of the reviews on policies, since this would have required monitoring of the details of policy developments in the reviewed countries over a longer period of time. Therefore, the conclusions in this section rely on the assumptions concerning the expected consequences through the above-described mechanisms at the individual and interpersonal levels, and on the subjective assessments made by the interviewees.

6.3.1 New decisions and actions

As mentioned above, an EPR seldom directly leads to changes in policies or provides genuinely new information or insights. Many of the interviewed government representatives underlined that it is the prerogative of the government to set the direction of the country's environmental policies. The OECD recommendations may be used to support this work, but can by no means have a leading role in the setting of priorities.[192] At some rare occasions, the reviews have been used rather

---

"The recommendations are considered as general policy orientations that to a large extent come from the OECD in general. For instance, the polluter pays principle and the user pays principle have been arguments frequently used in debates concerning transport policies in Lisbon." (a civil servant from the Portuguese environment ministry)

[191] Likewise, prior to the review mission, the OECD Environment Directorate systematically provide the review team members with copies of at least the summaries of the other relevant peer reviews carried out by the OECD on the country's policies over the preceding years (especially the Economic Surveys and IEA Energy Policy Reviews).

[192] *We do our policy; it is also a question of sovereignty. At the technical level, the review can have an influence, butwe don't take the recommendations as a point of departure when preparing our own policies."* (a Portuguese civil servant)
"At the political level, the review is not used in order to say: 'The OECD recommends us to do this – let us therefore implement these policies'." (a civil servant from the Portuguese environment ministry)

directly. In Austria, a recommendation concerning the environmental component of the country's official development assistance came as a surprise to the environmental authorities and prompted action. Likewise, the recommendation to ratify the Bonn Convention on the conservation of migratory species triggered the ratification process already before the report was officially released, since the recommendation figured already in the draft report. In Sweden, the review provided a major stimulus for a re-examination of the country's environmental inspection system.[193]

6.3.2 New shared understandings

The capacity to foster understanding around issues of common concern is a major type of potential conceptual use – learning – that the EPRs could generate. Shared understandings can arise, on the one hand, among the national policy players, and on the other, among the WPEP delegates. They may originate from the review process or from the publication of the review results. Four possible cases can be identified (Table 6.1). First, as was seen above, the EPRs are used to persuade and legitimise (upper left-hand quadrant, 1), but their failure to generate broad and long-standing public debate and interest compromises the potential of the review findings in this regard.[194]

The *review process* is often a more important source of influence in the reviewed country (2). For many participants in the reviewed country, the process of preparing the review is likely to be the principal source of learning, contributing to 'socialisation' and the adoption of the OECD ways of thinking in the member countries. Interaction among the different stakeholders can stimulate the **development of professional networks**. In general, the indirect impacts on networking are likely to be considerable, given the large number (typically more than 100) of individuals participating in the review process. This impact seems most important in countries where these networks of cooperation

---

[193] Personal communication, Anna Carin Thomér, Swedish WPEP delegate, 24 January 2005. The OECD recommended Sweden to "review, and revise as needed, state, regional and local inspection and enforcement roles, improving the monitoring and evaluation of environmental inspections, focusing enforcement on the areas with the greatest compliance problems and strengthening administrative and judicial sanctions" (OECD 2004c, 18). The criticism towards the Swedish inspection system was one of the key themes of debate during the OECD mission, and was highlighted by the environment minister as one of the key findings once the report was finalised (see e.g. http://social.regeringen.se/sb/d/1277/a/31076;jsessionid=apoo7yGYzrGb).

[194] *"The whole event went almost without generating any debate whatsoever, even though [the country's] number one daily newspaper […] had it on its first side."* (a WPEP delegate on his country's first cycle review)
"There was little public debate, even though some nice articles were published in the newspapers. We made a great effort in the ministry to brief the journalists and produce press releases (and the conclusions and recommendations of the review) in [our language], which helped enhance interest among journalists in the EPR." (a WPEP delegate)
"The environment is not a very popular subject today, and the issues dealt with in the reviews were quickly pushed off the agenda of public debate by other issues, such as those concerning safety, economic problems, etc. Perhaps the range of topics dealt with in the reviews was too broad. However, the media coverage was about what could be expected, given the nature of the report: no new findings were announced in the EPR." (a WPEP delegate)

between different sectors are not strongly developed. A prime example was Mexico, where especially the first cycle review enhanced cooperation both across sectors and within the environmental administration. The EPR provided an opportunity for interaction between actors in the various agencies merged together shortly before the review to form the new Ministry of the Environment. By contrast, the Japanese, and most Dutch and French interviewees did not consider network creation as a major benefit, arguing that networks to a large extent existed already prior to the review. Some French authorities from the non-environmental sectors noted, however, that the process allowed them to build contacts with those environment ministry officials that they usually did not have regular contacts with.

In Portugal, networking was mentioned by several interviewees, who nevertheless pointed out that it greatly benefits persons that have newly entered their position, but has at most minor benefits for those already having a longer experience and an established contact network already. Some interviewees thought the OECD review is not needed as a catalyst, since in a small country, all people working in a specific issue area know each other already.

An interesting example was mentioned by the Canadian environmental authorities who saw the review as essential in promoting interministerial cooperation by bringing together for the first time the numerous authorities and stakeholders around specific environmental themes, notably in the area of water management. This was somewhat surprising given the advanced level of environmental policies in the country. An explanation to this 'anomaly' is probably the decentralised character of the Canadian politico-administrative system, a large number of actors operating in each specific policy area.

Many interviewees mentioned the value of the review process as an occasion for the authorities to stop and reflect on their own work and evaluate its objectives and results. Thereby the reviews may *strengthen policies* by helping the government officials prioritise and focus on the essential and see their work in a broader context.[195] The review report also gives the greater public and other

---

[195] *"The OECD review helped us by pointing out the most critical and problematic aspects of our policies and by helping us perhaps see more clearly the priorities."* (a civil servant from a non-environmental sector)
"The process obliges us to constitute a comprehensive picture of our environmental policies, and forces in particular the non-environmental ministries to reflect on question concerning the environment. Usually, this happens only for specific issues – not for the whole panoply of environmental themes." (a civil servant from a Ministry of the Environment)

interested parties an opportunity to obtain a broader picture of environmental policies of the country. Such a process may *'boost moral'* among the civil servants working on environmental issues. One of the delegates to the Portuguese delegation in the peer review meeting argued that the good performance of the State Secretary in leading the delegation gave a positive image of the country, and thereby improved the image of the ministry in the eyes of the members of the Portuguese delegation.

For the WPEP delegates the peer review meetings are a forum for building professional networks, learning evaluation skills and 'evaluative' thinking, developing a common language, identifying issues of common concern, starting to use the same kind of reasoning when discussing these issues, learning from 'best practices' and, at a deeper level, even developing a common frame of reference and a common worldview (3). The interviewees were, however, often somewhat sceptical concerning the real value of the meetings in promoting such learning, in particular because of the lack of genuine dialogue. Moreover, it is unclear to what extent this learning at the individual level gets transmitted to the member countries and thus stimulates collective learning. Probably the link between WPEP work and the home government is often rather loose, most delegates not receiving instructions from their capitals, and reporting back to their superiors and colleagues only by writing standard meeting memoranda.[196] The need to strengthen the link between the WPEP and the EPOC as well as the general need to make the EPRs better known by various audiences have been brought up time and again in the discussions within the WPEP group. The participation of 'country experts' in the reviews of other countries likewise contributes to creating shared understandings within the WPEP and strengthening the ownership of WPEP delegates over the programme.

Finally, the records and experiences from previous EPRs constitute a body of WPEP 'case law', which helps the working party decide upon controversial issues and maintain a certain consistency and uniformity of treatment across countries (4). In the peer review meetings reference is constantly made to the commonly agreed upon OECD/WPEP principles as well as the conclusions and

---

The benefit of the EPRs providing a broad picture of the country's environmental policies was likewise mentioned by a Finnish environmental policy researcher, who complained about the Ministry of the Environment seemingly having lost sight of the 'broad picture'.

[196] While no systematic survey on countries' reporting systems was carried out, it is safe to assume that the control from the capitals in the EPRs is at least not stronger than in the politically more important Economic Surveys, on which one member of the Economic Survey group mentioned: *"In many countries, the delegates to the EDRC are given free hands to say what they want in the OECD meetings (i.e. without instructions from the capital), and sometimes they don't even report on the results of the meetings back to the capital. In these cases it is of course much less likely that the lessons reach the government more generally."*

recommendations from previous reports, in an effort to ensure an equal treatment across countries and to enhance the transparency with regard to the evaluation criteria.

Table 6.1. The type and degree of the impact from the EPR process on the creation of shared understandings.

		ORIGIN OF SHARED UNDERSTANDINGS	
		Report	**Process**
ACTORS INVOLVED	NATIONAL PLAYERS	*Weak/medium* Debate following the release of the report.	*High/medium* 'Process use' during the review.
	WPEP	*High* WPEP 'case law'	*Low/medium* Peer review meetings. Country experts' participation.

As argued above, the review team often learns the most from the reviews. While learning about the evaluation process certainly helps the experts better carry out their evaluation tasks, the learning that concerns the substance matter – environmental policies of the member countries – is at present not optimally utilised and translated into learning at the collective level. Yet, there would be significant potential for collective learning, notably through theory-building within the secretariat. Several informants noted that the few 'synthesis reports' drawing together the experiences and lessons from the EPRs (OECD 1996; 1998b; 2001d; 2003f) had been useful.[197] Largely because of the lack of resources within the Environment Directorate, synthesis reports are published only seldom.

At the most general level, the role of the OECD as a source of a general reference framework, a mode of thinking, and a forum for discussions on policy approaches and practical policy applications was underlined time and again by the interviewees. This impact included the introduction of environmental thinking into transport and agricultural policies, promoting environmental policy integration.[198] When the reviews provide a 'baseline' of environmental

---

[197] Moreover, in the recent discussions within the WPEP concerning the future of the programme, the option of combining the production of the ordinary country reports with synthesis reports on specific themes or issues received considerable support.

[198] *"The influence of the OECD stems principally from its capacity to introduce new questions and models of thinking (e.g. environmental issues in agriculture), better analytical models, etc. It provides new frames of reference. At the political level, it has less impact."* (a civil servant from a non-environmental sector)
"The OECD is useful as a meeting place, a forum for discussing policies with experts working in the same field, developing policy approaches (e.g. emissions trading) and practical policy applications, carrying out applied studies […] OECD documents are also good reference works. By contrast, the political weight of the OECD is much more modest." (a civil servant from a non-environmental sector)

policies they can also be seen to foster common understanding by providing a more rational basis for policy debates.[199] In Japan, especially the first *ad hoc* review carried out in 1978 and the first regular review (1992) gave new impetus to environmental policies in the country, including in the first case the introduction of the concept of 'amenity', hitherto unknown in the country. It is probably safe to say that, as a rule, the reviews create new, shared understandings mainly among authorities, whereas the participation of other stakeholders – both during the process and after the release of the report – is insufficient to extend major impact beyond the government. Likely the most important ways of enhancing common understandings are the EPRs' functions of 'infiltrating' economic thinking to the environmental administrations and environmental policy in general,[200] and environmental thinking to the economically oriented sectors of administration.[201] A major role of the EPRs is to promote the development of a results-based management culture in the member countries, in particular by encouraging the country to establish clear policy objectives against which performance could be evaluated. Setting more concrete objectives and targets is among the recurrent themes especially in the recommendations concerning sector integration.[202]

6.3.3 Impacts on the legitimacy of actors and policies

The above discussion has frequently pointed towards what most see as the main function of the reviews: to lend environmental issues and environmental authorities of the reviewed country increased legitimacy in the eyes of the public, the other sectors of government, and other

---

"For me, the OECD serves above all as a laboratory in which new concepts are being invented (e.g. the polluter pays principle) and a forum for the exchange of information and opinions on policies, in an environment that is more relaxed than in the context of the international negotiations." (a civil servant from the Ministry of the Environment)

[199] *"The main value of the review from the point of view of dialogue is that it allows a debate founded on a rational basis."* (a civil servant from a Ministry of the Environment)

[200] *"Traditionally, the dominant way of thought in the Ministry of the Environment has been that of the 'environmentalists', who have had a certain hostility towards more economic approaches in environmental policy. The creation of [our unit] can be seen as a sign of recognition of this [more economically oriented] approach, and a more general acceptance of the OECD, previously perceived as somewhat suspect by the 'environmentalist phalange' of the Ministry."* (a civil servant from a Ministry of the Environment)

"For us, the OECD represents an ally, an organisation with a certain reputation, which supports us and provides us studies and arguments for our efforts to promote greater use of economic instruments in environmental policy." (a civil servant from the Ministry of Finance)

[201] *"The fact that the OECD is an economic organisation is an undisputable benefit in the 'mainstreaming' work; the opinions of the OECD are taken much more seriously by actors in the non-environmental sectors than those of the Ministry of the Environment, for instance."* (a WPEP delegate)

"As the proverb has it: 'home saints don't do miracles'. To prompt concrete action, it is necessary to have an external organisation to say what we have repeated a million times." (a civil servant from a Ministry of the Environment)

[202] To take but two examples, for Austria, the OECD recommended to "establish quantified environmental goals for the forestry sector and monitor their achievement, possibly as part of the forthcoming National Forest Programme" (OECD 2003g, 27). The Spanish review, in turn, recommends the country to "pursue efforts to integrate environmental concerns in the tourism sector by establishing a national strategy for sustainable tourism development; introduce quantitative and qualitative targets" (OEDC 2004d, 27).

stakeholders. In this sense, the reviews resemble 'empowerment' or advocacy evaluation. Legitimisation stems both from the review process, which allows the environmental authorities to put forward their views and gain support for environmental issues, and from the findings and recommendations issued by a prestigious international organisation.

Opinions, particularly criticism, from the outside exert strong pressure on the environmental authorities especially in 'laggard' countries.[203] In some cases, the report was seen as an instrument to improve the status of certain types of environmental policies in the country, thus changing the order of priority between different environmental objectives and themes within the environmental administration.[204] However, since the reports do not explicitly prioritise between the different chapters and recommendations, the degree to which they can be used to improve the status of a particular area of environmental policy mostly depends on the active efforts undertaken by the authorities in that particular area. A number of persons responsible for environmental issues in non-environmental sectors welcomed the support of the OECD in helping to legitimise integration of environmental concerns in their sectors.

6.3.4 Impacts of the EPRs beyond the reviewed country

The use of the EPRs of other countries as tools of learning was mentioned by some, but is probably too unsystematic to constitute a major vehicle for policy imitation and diffusion.[205] During the review process, especially in the WPEP meetings, dialogue remains on a too general level to prompt direct imitation. The major imitation impacts are likely to be indirect, member countries adopting policies developed in 'forerunner' countries, subsequently adopted as OECD policy principles.

Diffusion of the idea and methodology of peer review has probably much more potential. In a number of interviews, as well as in the WPEP meetings, it was stressed that, indirectly, the impact

---

[203] *"For a country like Portugal, which is a latecomer compared to other countries, this type of review is useful, as it provides information on where we stand, and exerts external pressure."* (a Portuguese civil servant)

[204] *"International affairs were not a priority in our ministry. The review helped increase the weight and the visibility of these issues"* (a civil servant from the Ministry of the Environment).

[205] *"I am not familiar with [the reports of other countries], but I have used the Portuguese energy policy review [made by the IEA]."* (a civil servant from the Portuguese transport ministry)

"OECD reviews and OECD studies in general are used as reference works that provide insights notably into other countries' policies." (a Dutch civil servant from a non-environmental sector)

"I used the Portuguese EPR report when commenting on work by my colleagues who participated in the Portuguese Economic Survey as reviewing country experts." (a WPEP delegate)

"In the beginning of the first cycle, when I was a head of a department in the Ministry, I wrote short articles on the recommendations of other countries reviews (perhaps the first 10-12 ones) in the Ministry's official journal. Once I moved to other duties, I quit writing these articles." (a WPEP delegate)

of the EPRs actually reaches beyond the OECD countries, since the EPR framework has generated interest in, or is already being adopted, by other international organisations. The UN Economic Commission for Europe (UN-ECE) launched already several years ago a review programme – closely imitating the EPR model – for its European non-OECD countries, and the UN Economic Commission for Latin America and the Caribbean (ECLAC) is starting – in collaboration with the OECD – a review programme for its member countries. Reviews have already been carried out in Russia (together with UN-ECE), and Chile (together with ECLAC), while reviews of China and Argentina are under preparation. Programme co-operation is also underway with the Asian Development Bank (in the greater Mekong Region) and the World Bank. (Lorentsen 2004; OECD 2004a.) These ideas of 'exporting' the EPR concept have gained plenty of support in several member countries. One interviewee suggested, for instance, that Japan, together with the OECD, could use the EPRs more actively in the Asian developing countries. Finally, the EPR methodology has been tested and adapted to the local level in the framework of the 'Prosud' partnership project where nine European cities collaborate to develop ways of assessing and promoting sustainable development at city level.[206] In the future, this impact beyond the OECD is likely to gain further importance, and it has been suggested already that the peer review logic be applied to the evaluation of multinational enterprises, linking the peer reviews to the OECD work in this field.[207]

6.3.5 Agenda-setting

Even though the EPRs usually do not introduce completely new issues to the political agenda of the reviewed country, they often either strengthen the status of issues that have already entered the public agenda, but have not really taken off yet (see section 6.2.2 above), or revive old issues that have been abandoned at an earlier stage (e.g. section 6.1). In both cases, the impact relies on the prestige of the OECD: in the former, it enhances the legitimacy of a new issue, and in the latter, it promotes an old idea in order to see whether it might take off in a new political situation.

6.3.6 Impact on attitudes: resistance and evaluation fatigue?

The reviews have not escaped the problem of engendering *defensive attitudes and resistance* within the governments. This was clearly the case in Germany and in Finland. In the former, the organisers of the second cycle review did not seize the opportunity to use the review as a source of political

---

[206] Personal communication from Johannes Dictus, City of Vienna Environmental Department, 16 May 2003.
[207] Personal communication, an official from the OECD Environment Directorate.

pressure to improve environmental policies and strengthen their legitimacy. In Finland, the desire to postpone the review by several years, the country's upcoming EU presidency (still several years ahead) used as the official reason, clearly testifies to the reluctance of authorities to undergo a review process.[208] In Portugal, the fact that some interviewees referred to the facility with which the review team could be 'cheated' by concealing the most problematic aspects of the country's environmental policy, also shows certain resistance, even though such behaviour can be considered as perfectly logical and 'human' for a civil servant, whose work is being scrutinised. In France, resistance was not seen as a significant factor, even though it was acknowledged that the organisation of the review required substantial effort and resources. Some interviewees in the sectoral ministries admitted that, frankly speaking, all that the review meant for them was a lot of work and a risk of the civil society using the report to criticise the ministry's work.[209] Several participants have underlined that the preparation of the review requires a substantial amount of work. The Dutch argued that the review did not give rise to particular resistance, because open criticism and debate are part and parcel of the country's political culture.

6.4 Summary of the types of influence from the EPRs

This study has no pretension to objectively measure the average influence of an EPR in the different OECD countries; the variations between the reviewed countries' situations render generalisations concerning the overall impact of EPRs difficult and rather irrelevant for the future of the programme. Instead, this chapter has illustrated the pathways through which the reviews seem to exert their influence in the OECD countries. It seems that often the influence from the reviews has been modest, given the frequently rather disappointed comments concerning the usefulness of the reviews. Among the countries where the informants saw the review as influential were Norway, Mexico, and Austria, whereas Hungary, Japan, Portugal, and the Slovak Republic saw the influence as minor, yet not totally insignificant.[210] In France, general expectations concerning the usefulness of the review were rather positive soon after the peer review meeting, but information is not yet

---

[208] In informal discussions, two of the ministry's middle-level civil servants referred to defensive attitudes on the part of the ministry's top management.

[209] *"As for the usefulness of the Environmental Performance Reviews, I would say, as a civil servant of our ministry, that it would be better if the reviews did not exist: they significantly increase our workload and thereby prevent us from concentrating on our daily, more urgent tasks. They also represent the danger that pressure groups (associations, local politicians, etc.) seize the opportunity to use the review in order to criticise us and render our work more difficult."* (a civil servant from a non-environmental sector)

[210] Of course, the estimation concerning Austria, Hungary, Mexico, Norway, and Slovakia is uncertain to the extent that the information is essentially based on the views of a single individual for each country (i.e. the countries' WPEP delegates).

available about the subsequent consequences. The most pessimistic and disappointed views were those of the Dutch interviewees, none of whom argued that the review would have had a major impact.[211] A further indication of the overall influence of the reviews is provided by the annual activity ranking exercise within the OECD Environmental Policy Committee (EPOC). In the latest ranking exercise, somewhat surprisingly in view of the interviews, Portugal gave the reviews the lowest possible score, whereas the Netherlands, for instance, ranked it in the medium range. Noteworthy is the call by the U.S. for a reduced frequency of reviews in the OECD member countries, in order to make room for reviews of developing countries.

With reference to the typology presented in chapter 4 (table 4.1), we can now identify the main types of influence from the EPRs.

---

[211] Given this background, it was not surprising that the country's delegate at the EPOC meeting in November 2004 argued against the secretariat's proposal to target the required budget cuts within the Environment Directorate at the work on the environmental aspects of transport, and suggested cutting the financing of the EPR programme instead.

Table 6.2. Summary of the main types of influence from the EPRs.

	CHANGES AT INDIVIDUAL LEVEL	PROCESSES OF DIALOGUE AND ARGUMENTATION	OUTCOMES AT COLLECTIVE LEVEL
PROCESS	***Transformed beliefs*** Salience Elaboration	***Debates at national level*** Criticism Defence ***Debates between OECD team and national experts*** One-way exchange of information	***New shared understandings*** Socialisation (OECD ideas and ways of thinking transmitted to the country) Synthesis, winner, deadlock (when preparing the country's positions) ***Legitimisation*** Of environmental policies and actors ***Networking***
RECOMMENDATIONS	***Confirmed beliefs***	***Persuasion*** Environmental authorities towards other sectors ***Legitimisation*** Environmental authorities try to legitimise their policies	***Decisions and actions*** Recommendations adopted as such or with minor modifications ***New shared understandings*** Socialisation into OECD policy doctrine ***Legitimacy*** 'Empowerment' of environmental authorities and legitimisation of environmental policies 'Infiltration' of OECD thinking into environmental administrations
REPORT AS A SOURCE OF BACKGROUND INFORMATION	***Confirmed beliefs*** Environmental authorities ***Transformed beliefs*** (non-environmental authorities) Learning on the country's environmental policy and the country's position in relation to its peers	***Background information*** to the debate from a 'neutral' and reliable source	***Socialisation*** to OECD way of thinking (e.g. use of indicators) ***Legitimacy*** of environmental policies and authorities in the face of other authorities and the public

The review process 'forces' individuals to elaborate more in detail on the themes addressed in the reviews, and may lead to changes in the order of priority the individuals attach to different topics. The debate among national stakeholders, in particular when preparing the country's positions with a view to the OECD review mission, is likely to involve attempts by the different participants to persuade others, legitimise their own positions, criticise those of others, and defend themselves against criticism from others. The dialogue between the OECD team and the national counterparts consists mainly of one-way exchange of information, but was considered by some interviewees as an opportunity to learn from new perspectives and opinions brought forward by the OECD secretariat and country experts. Finally, the interviews indicated that synthesis and compromise are more likely in the process of preparing country positions for the EPR than in day-to-day policymaking, thanks to the prestige of the OECD. The processes of dialogue and argumentation probably contribute to socialisation, by disseminating the OECD policy approach. Networking effects are important where networks remain weakly developed.

The recommendations of the report mostly confirm the existing beliefs, rather than changing them, because the reviews seldom contain novelties. The environmental authorities in the first place use the recommendations to persuade other ministries to adopt certain policies, and to legitimise environmental policies both within the government and among the broader public. Criticism and defence do take place, but more seldom, because the recommendations are rather 'soft' and the public debate following the release of the report is relatively limited. In most cases, the government accepts the recommendations as such, or with minor modifications, and integrates them into the existing policy streams. The recommendations probably contribute to the socialisation processes, by repeating the central themes of the OECD policy doctrine. The main results from the dialogue concerning the review recommendations are increased legitimacy of environmental policies and authorities on the one hand, and the infiltration of OECD frameworks of thought into environmental administrations.

When the reviews are used as background material – in drafting policy documents, programmes, and strategies, for instance – they contribute to confirming and supporting the beliefs of the environmental authorities, and may transform the beliefs of individuals beyond the environmental authorities. The EPRs are frequently used as a source of background information to provide a basis for more rational debates within the government and in public arenas. This way, the reports enhance

the adoption of OECD frameworks of analysis, and help enhance the legitimacy of environmental policies and authorities in the face of the public and the other authorities.

The analysis in this chapter has shown that often the impacts from the review process are as important as those that take place after the review has been published. In particular, the preparation in the reviewed country for the visit of the review team often plays an important role in enhancing the visibility of environmental issues, and promoting environmental policy integration. The reviews seldom provide genuinely new ideas to authorities working on environmental policies, whereas civil servants from other sectors and researchers may learn more by reading the reports. Finally, the processes of argumentation during the review process and after the release of the report serve above all to enhance the legitimacy and visibility of environmental policies and authorities, as well as to infiltrate OECD-type thinking into the environmental bureaucracies.

The main types of impacts from the reviews having been exposed, the focus in the next chapter will be on the 'why' questions – the reasons behind the variations in the types and degree of review influence.

Chapter 7

Factors affecting the influence of the EPRs

The pathway and the degree of influence that evaluations have on policies largely depend on the context within which the evaluations are carried out. Relying on the concepts presented in section 4.4, let us now proceed to examining three groups of contextual factors in order to understand the variations of influence in different countries and under different conditions.[212] While the previous chapter tried to identify the ways in which the EPRs exert their influence, this chapter focuses on the judgements of the various participants concerning the factors that enable or prevent the reviews to have an impact, in order to come up with a reliable picture of how the EPRs influence policies.

The first issue that will be addressed is the way in which the reviews are designed and organised, a theme that has obviously attracted plenty of attention in the debates within the OECD WPEP group. The views of the participants concerning the significance of the review design on review influence often consisted of dichotomous choices between, for example, 'hard' and 'soft' recommendations, or comparability and country-specificity. Combining these dichotomies with the insights from the previous chapter concerning the types of review influence, it was possible to construct a typology of four 'ideal types' of pathways of influence, which will be presented in the end of the first section of this chapter.

To find out to what extent the diverging opinions concerning the review design and the preferred pathway of influence varied in function of the actor-specific characteristics, a second cluster of

---

[212] The factors are here grouped under three headings instead of the four identified in chapter six, because of the role of the evaluator (the OECD) as an inseparable part of the actor networks.

issues was looked at – the networks and power relations among the different actors involved in the reviews, as well as the various 'repertoires' of these actors. The perspective of 'epistemic communities' made it possible to classify the participants into two main rivalling but partially overlapping communities of 'environmentalists' and 'economists', each trying to exploit the EPRs for its own purposes. Such a distinction is largely based on the interpretations made by the interviewees, the majority of which can be placed in the 'environmentalist' community. The frequent reference to the existence of the two communities is therefore taken as a proof of the fact that such a division at least structures the thoughts of the major 'environmentalist' participants, and therefore has an impact on review influence. To make sense of the variations within these communities, the concept of 'repertoires' turned out useful, in drawing attention not only to the distinctions between the groups, but also to their internal divergences.

The role – especially the perceived credibility and independence – of the OECD as evaluator came up frequently in the interviews and discussions within the WPEP as a factor conditioning review influence. In the EPRs, the OECD tries to balance between the different epistemic communities by building links between the various repertoires, while at the same time operating within these communities and having its own repertoires. The way in which the OECD does this balancing is of utmost importance for the influence of the reviews, in particular through the extent to which it is capable of mobilising the potential 'change agents' indispensable for review influence. Identifying the different repertoires and the potential change agents helped place the choices concerning the pathways of review influence into the context of actors interests, beliefs and expectations.

The epistemic communities and their 'repertoires' are, however, not independent of the broader country-specific institutional context. Indeed, the variations in the country-specific conditions constitute a crucial challenge for OECD work in general, and for its environmental country reviews in particular. Therefore, the third set of questions dealt with in this chapter relates to the institutional context. Three among the several factors identified in literature on the influence of evaluations and research knowledge on policies stood out in the case of the EPRs: the national policy style and policymaking structures, socio-economic conditions, and changes in the governing coalitions and administrative structures. While far-reaching conclusions cannot be drawn on the influence of such factors in country-specific situations, a number of tentative hypotheses concerning the impact of the country-specific context are suggested for future testing. The chapter concludes by presenting three 'mini-case studies', seeking to explain the reasons for and types of review influence in France, the

Netherlands and Portugal – the countries where interviews were carried out not only among the WPEP delegates, but also among other individuals having participated in the reviews.

7.1 Review design[213]

It is natural to first seek reasons for the lack of debate in the review design: too little emphasis on explaining policy outcomes; too little cross-country comparison; too long a review cycle; too 'kind' and general recommendations; insufficient dialogue; poor dissemination of the reviews; and the absence or weakness of follow-up of the implementation of review recommendations have all been evoked as shortcomings in the EPR design.[214] However, opinions within the WPEP group as to the relative importance of these factors are rather contrasted, and consensus prevails only on a few of the improvements that would be needed, such as reducing the length of the review cycle, and improving the dissemination and the follow-up (e.g. OECD 2003i). The following analysis seeks to present the key issues and give a rough idea of the preferences of the different participants concerning the review design and the kind of impacts reviews would be expected to generate.

7.1.1 Criteria of evaluation[215]

Beyond the lack of cross-country comparison, the participants involved in the reviews have hardly questioned the set of criteria used to evaluate policies in the EPRs, and instead rather endorse the general focus on goal-achievement. A common criticism has been that goal-achievement alone is not enough, but should be complemented by other criteria, focusing in particular on better explaining policy outcomes (e.g. OECD 2003i, 5).

However, the view of the former Finnish OECD Ambassador Jorma Julin (2003) that "rankings and benchmarking may entail political sensitivities, but are usually welcomed by governments as tools for keeping focus on improving performance", reflects an opinion shared by some WPEP delegates, who argue that increasing comparability between countries would strengthen the impact of the EPRs. Yet, while greater comparability was still one of the objectives mentioned in the work plan

---

[213] See section 4.4.4 for a discussion on the importance of review design in conditioning evaluation influence.

[214] These issues were prominent already in the survey carried out in 1997 on the experiences from the 'first cycle' of EPRs (OECD 1997b).

[215] See section 4.4.4.4.

for the second cycle of reviews (OECD 1998a, 4), the interest seems to have faded since then.[216] By contrast, especially many non-European countries have explicitly rejected the idea of more comparison.[217] Business and industry call for transparent criteria for indicator selection, and care in making comparisons, so as to avoid indicators being used in abstraction from the context "presuming priorities that may vary from country to country in terms of relative importance" (BIAC 2004, 2).

7.1.2 Competence of the review team, the quality of the reports and their recommendations[218]

In general, the interviewees considered the review team members as highly qualified. Only in few cases was the competence of an individual expert in the team called in question, and the 'heterogeneity' of country experts seen to compromise the credibility of the reviews. The problem of the country experts being mislead by the reviewed country officials (see section 6.2.1 above) is easy to explain by the asymmetry of information favouring the national environmental authorities. This difficulty concerns the whole review team, but in particular the country experts who usually do not have time for intensive data and information search. An aggravating factor is that the review team normally does not have the opportunity to come back with supplementary questions to the authorities after having heard the usually more critical views from the NGOs, the researchers and the industry representatives.[219] Some informants also claimed the OECD experts did not always have sufficient knowledge about the country-specific context, such as the relations between the different levels of government in federal countries or even the subject matter. Many emphasised the importance of such contextual understanding for the quality and influence of the reviews, and the need to ensure the facts presented in the report are correct.[220]

---

[216] In the discussions on the future of the programme, on 18 May 2005, the Dutch delegate was the only one who spoke in favour of what he called a 'benchmarking' approach. Of course, cross-country comparisons are always a delicate issue, and countries' preferences in the matter tend to fluctuate. This was illustrated by the fact that the plan to develop a single index of sustainability, to be possibly used in the OECD Economic Surveys, was blocked by two small European countries that found they were unfairly treated in the proposed indicator framework (Personal communication, Finland's OECD ambassador Jorma Julin, 14 January 2005).

[217] A civil servant from a non-European country expressed the dilemma in the following statement: *"One of our main issues of concern is the extent to which the OECD, with its likely enlargement to the new EU countries, will increasingly drive a 'European vision' of environmental policy, failing to take adequately into account our specificities. This is a concern we share with some other non-European countries."*

[218] See section 4.4.4.1.

[219] During the review mission, the review team meets the national counterparts in a predefined order: first the national authorities (2-3 days), then the NGOs, industry, researchers and local authorities.

[220] *"[The review team] would need to better understand the cultural factors."* (an environmental policy researcher) *"For instance, Italy was an impossible case for many OECD team members who really did not understand what was going on."* (a WPEP delegate)

Whether the *language* of the EPR reports should be 'sharpened' and made more 'digestible' for the general public has aroused some debate. The former Finnish OECD Ambassador Jorma Julin (2003) criticises the OECD's work in general for excessive prudence, noting that "too much OECD output remains dry and shy" and that it would have more influence if "tough realities" were put on the table with honesty and rigour. With regard to the EPRs, the main questions concern the 'technicality' of the language in the main report on the one hand, and the 'sharpness' and the level of detail in the recommendations, on the other.

Some considered the language of the reports too technical, 'austere', and hardly accessible to the public at large. While the excessive length of the reports has often been evoked as a problem, it was also pointed out that very few people would read the whole report, and that the main users would anyhow be experts interested in a specific policy issue (an individual chapter), would be perfectly capable of interpreting the technical and diplomatic language of the OECD. Recently, it has been proposed that a clearer division be made between, on the one hand, the document containing the conclusions and recommendations, aimed for broad audience and therefore having to be easily 'digestible', and on the other, the main report, which would be addressed to specialised readership, and could employ more technical language (e.g. OECD 2005a, 6).

Opinions concerning the 'tone' of the *recommendations* are rather contrasted, especially when it comes to the balance between 'carrot' and 'stick'. On the one hand, some are satisfied with the present, rather diplomatic tone, arguing in particular that imposition and very strong language would not fit into the OECD's role and would engender resistance among civil servants and politicians.[221] Furthermore, too detailed recommendations would risk drawing attention away from

---

"I wonder why the review team did not include more experienced specialists who would have understood the complexities involved in the internalisation of external costs in [this subject area]" (a civil servant from a non-environmental sector)

[221] A French civil servant expressed the trade-offs by stating: *"The shortcomings [of the country's policies] must be clear in the conclusions, but I'm not sure that confrontation (through very strict recommendations) would be a good way of advancing the issues, because it could easily provoke resistance. People usually better accept legally non-binding recommendations. I admit, of course, that strategies of the type 'name and shame' (list of the largest polluters published with regular intervals) – of which we have experience in France as well – as well as the publicity and public pressure have their function, too."*

"There must a balance between 'carrot and stick': an overly critical report is easily buried and forgotten; too strong recommendations are counterproductive as they generate defensive attitudes among the authorities." (a Dutch interviewee)

"The EPRs can be accused of being too 'kind' and diplomatic in their approach, and the recommendations of not going deep enough into criticising the reviewed country's policies. […] However, the OECD […] must respect the country's right to decide upon the way of carrying out its policies and organising its public administration. The OECD

the general principles, and being interpreted as violation of national sovereignty (OECD 1997b). It should be the country's prerogative to establish its policy priorities and define the ways to achieve them. These arguments get support from the observation of Leknes (2001), that the impact of evaluation is higher on technical and 'professional' rather than politically 'hot' topics (see section 4.4.2 above).

On the other hand, a great number of the persons involved in the reviews consider the recommendations too weak, 'kind', and general; lacking prioritisation; not sufficiently future-oriented; hardly amenable to quantitative follow-up; and not credible to audiences beyond the environmental authorities. A recurrent argument in the discussions within the WPEP is that the environmental authorities cannot use recommendations that are not sufficiently 'biting'.[222] Representing little more than a declaration of the general OECD principles (internalisation of externalities, polluter pays principle, integration of environmental concerns to other sectors, etc.) the recommendations do not provide much in the way of guidance as to their application in the concrete context that varies from one country to another. The reviews hardly try to answer the 'why' questions, even though there is a consensus in the WPEP group concerning the desirability of such policy analysis (OECD 2003i).[223] The main part of the reviews still consists of description of the policy measures undertaken by the government, the state and the trends in environmental pressures and conditions, with a focus on issues that are deemed problematic. Whenever reasons for policy success or failure are given, they rely on assessments and evaluations carried out by others – usually by the government or scholars in the reviewed country. Many participants pointed out that

---

can mainly just criticise the results of policies, otherwise it would create problems within the country's government." (a WPEP delegate)

[222] The same problem was evoked with respect to other peer reviews: *"in the peer meeting, the head of the delegation first thanked the secretariat for having drafted a 'beautiful' report, but noted that the report was so nice it would be practically useless for the government. Therefore, the assessment and the recommendations of the report were strengthened, as the [reviewed country's] delegation wished."* (an EDRC delegate).
"A typical example was the development assistance review of Norway, in which the minister accepted no criticism at all, but declared upfront that the country's work in this field is outstanding." (an official from a permanent OECD delegation)
"The review was not confrontational enough to have any political impact, and did not allow sufficient comparison between OECD countries." (a Dutch civil servant)
[223] *"The recommendations usually just said we should do what we were actually doing already, rather than providing new ideas on **how** to do things. The recommendations were probably relevant, but did not provide any new ideas or give directions to future work. The recommendations should be more concrete, and should in particular allow a quantitative assessment of the degree to which they have been put into practice."* (a Dutch civil servant)
"It is not enough to just check whether the country has achieved the objectives it has established for itself, but there should be more qualitative analysis and a broader perspective to the impacts of policies." (a Dutch civil servant)
"The OECD review does not reach particularly far. It is a status report that shows what is the situation, but does not contain any analysis of why several national environmental objectives will not be reached. Overall, the recommendations are rather uninformative." (comment by Mikael Karlsson, Swedish Association for Nature Conservation, in the press conference organised to release the Swedish EPR, 1 October 2004; http://justitie.regeringen.se/sb/d/3414/a/26952/m/wai)

some recommendations were actually too specific, and seemed to be chosen rather haphazardly among the numerous potentially relevant themes within the subject area, while some found it confusing that the recommendations were so heterogeneous consisting of a mix of general principles and detailed prescriptions.[224]

In conclusion, the OECD faces hard choices between sufficient, healthy criticism on the one hand, and counterproductive 'sermonising' on the other, and between the risk of trying to impose on all member countries a 'European-oriented' environmental policy doctrine on the one hand, and rendering all cross-country comparison impossible as a consequence of too much country-specific 'tailoring', on the other.

7.1.3 Character of the issues[225]

A crucial determinant of the review influence is the degree of controversy surrounding the issues treated in the review and the appropriate 'misfit' between the country's policies and the recommendations. The degree of controversy does not, of course, determine the impact of the review as a whole, but only with respect to specific topics. The question of 'misfit', by contrast, is broader, and is intimately linked with the overall tone of the report and its recommendations.

The more controversial the issue at stake, the more eagerly policy actors seize the OECD report to seek support for their own ideas and the more vigorously the country delegation defends its views in the peer review meeting. In Portugal, the question of waste management options (incineration or not) was a case in point, while the debate on the CO_2 tax provided a similar example in Germany. In the French peer review meeting, a formulation in the draft conclusions of the review concerning the internalisation of the external costs of nuclear power was debated at length before a consensus could be found. However, the fact that a particular conflict-laden issue receives a lot of attention does not necessarily increase the influence of the review as a whole. In fact, the opposite may well be the case: the focus on few 'hot' issues may deviate attention away from the general picture.

---

[224] *"It would probably be a good idea to place the recommendations into a hierarchy, and reduce their number [...] the recommendations could be divided into two categories: the general policy level recommendations and more detailed, practical ones."* (a Dutch interviewee)
[225] See section 4.4.4.2.

7.1.4 Scope and focus of the reviews[226]

The question of whether the range of issues covered should be broad or narrow involves inevitable trade-offs between salience, credibility, and legitimacy. Currently, the breadth of the themes covered in the reviews prevents a deeper analysis of the reasons for policy success or failure. Even though the coverage and the length of the report have recently been slightly reduced (see section 5.4.4), the members of the WPEP do not agree on review coverage, as some are willing to move towards thematic reports, while others already regret the loss of broad coverage.[227] The persons most in favour of a more focused approach tend to come from countries considered 'frontrunners' in environmental policies. Yet, even these individuals often acknowledge the potential usefulness of broad reviews for the less advanced countries, and propose to choose the breadth of focus case-by-case.[228]

The choice between broad and narrow scope of the review largely depends on the views concerning the OECD's 'comparative advantage'. Those stressing the OECD's role in promoting learning argue for deeper dialogue, which is only possible on a limited range of issues. Others consider, in turn, that narrow, sector-based or even cross-sectoral but theme-oriented reviews are already carried out by various organisations and independent consultants, whereas the strength of the OECD would lie in its capacity to cover the countries' environmental policy across the board.[229]

7.1.5 Organisation of the review, dissemination, and follow-up230

The role of the review process in promoting networking, policy integration and creation of shared understandings, as well as the dilemmas related to the character of the peer review meeting were already addressed in the previous chapter, and will not be repeated here. Three more issues concerning the organisation of the review are of relevance here.

---

[226] See section 4.4.4.3.

[227] In the WPEP meeting, 18 May 2005, many delegates pointed out that the length of the report as such was not the main problem, and some considered that shortening the report and dropping out specific themes (e.g. waste management) had in some cases compromised the quality of the reviews.

[228] *"On the question of broad vs. narrow scope, I would suggest a combination: for us, a more in-depth review would have been more useful, whereas the broad coverage may be beneficial for some other countries whose national institutions don't produce reports with broad coverage."* (a Dutch civil servant)

[229] *"…we tend to prefer broad coverage. That is where the OECD could have its comparative advantage, since different kinds of sectoral reports are frequently produced by different research institutes and government bodies."* (a WPEP delegate)
"The first part [of the report] is already much closer to the assessments made by the EU, for example. By contrast, the EU has nothing that would be reminiscent of the parts II and III." (a civil servant from a Ministry of the Environment)

[230] See section 4.4.4.6.

First, the participants should be well prepared and understand the objectives and the 'logic' of the review. For instance, it is sometimes difficult for the authorities not familiar with the reviews to understand or accept that the review team principally looks for information on the performance of past policies, and not so much on the policy objectives and programmes planned for the future. The coordinators of the process in national environmental administrations carry the main responsibility for explaining the review logic and phases of the process both to the governmental and non-governmental participants.[231]

Second, there is a broad consensus in the WPEP on the need to *shorten the review cycle*, in order to 'keep up the momentum' created by a review. In practice, however, this has so far proven impossible, and the budgetary resources available for the review programme have diminished rather than increased over the past ten years. While some have proposed to drastically shorten the review cycle, the consensus seems to be that a cycle shorter than about five years might not be desirable, taking into account the long-term character of environmental issues, and the excessive resource demands that an accelerated review cycle would impose on the reviewed country government.[232]

Third, the *length of an individual review process* is currently about a year and a half, from the first visit of the head of the OECD EPI Division to the country, until the release of the final review report. Obviously, this is long enough to make some of the review material outdated by the time the report is released. For most countries, this does not seem to represent a major issue of concern, however. Moreover, the 'timelessness' of the reviews was mentioned as a positive aspect by some interviewees, who argued that the rather general nature of the recommendations has the advantage that it allows well-informed participants – civil servants in particular – to use them long after the release of the report, once the political environment has become favourable to their adoption.[233] Such a 'timelessness' is a major benefit, given that the timeliness has frequently been identified as

---

[231] Sometimes especially the NGO representatives have been called to the event at the last moment, which has compromised the quality of their contribution to the process (personal communication, Heikki Sisula, Finland's former WPEP delegate; participated in the review teams of Italy, Mexico, the Netherlands, Russia, and Slovakia).

[232] *"The review cycle is, indeed, very long, but the review process, as it is designed at present, is very resource-demanding. Shortening the cycle would therefore also require changes to the process, because conducting an exercise of this level of resource demand each three years, for instance, would mean that a significant number of our staff would be permanently managing the reviews."* (a civil servant from a Ministry of the Environment)

[233] *"Certain issues in the EPR are taken up by different actors in [our country] still 2-3 years after the release of the report. The recommendation to eliminate environmentally harmful subsidies constitutes such an example. People who are aware of the existence of the EPR report are likely to use it still well after its release."* (an OECD official)
"An advantage with the OECD review is that it has a certain 'longevity': the messages contained in the report do not become obsolete overnight, which may be the case for certain evaluations carried out by the EU, for example." (a French civil servant)

one of the key conditions for the usefulness of evaluations (e.g. Cousins and Leithwood 1986; Lang 2001).

Many informants considered poor *dissemination* as a major weakness, and only specialists are said to be even aware of the existence of the reviews. While the reviewed country has the main responsibility for dissemination, the OECD secretariat has recently paid greater attention to the issue. The OECD has been criticised for not distributing free copies of the reviews in sufficient numbers, thereby making it difficult in particular for the NGOs to access the reports. In Austria, a considerable effort was made in order to attract media interest, for example through press briefings, as well as by providing as much as possible German-language information to the press. From the point of view of the OECD, the release event represents a major occasion to make publicity both for the EPRs and the OECD work in general.

The importance of regular *follow-up of the implementation* of the review recommendations has been emphasised time and again (e.g. OECD 2002c, 10; OECD 2005a, 4). Many have underlined the importance of follow-up as a mechanism for keeping up the 'momentum' created by the review, strengthening the 'peer pressure', 'responsibilising' actors involved, and ensuring government accountability towards citizens between two reviews. The industry or the NGOs could, in principle, also organise seminars to support the follow-up and make the reviews known by their reference groups, but presumably this has not been done in any OECD country. More rigorous follow-up has been mentioned as an indispensable 'second-best option' to shortening the review cycle.

In reality, the follow-up reports do not seem to have really delivered on their promises. Either the reports have generated rather modest interest both among the stakeholders and in the OECD, further reducing the motivation to produce such reports (such as in Finland, Japan, and the Netherlands), or there has been no follow-up whatsoever (e.g. Canada and Portugal after the first review).[234] A number of interviewees evoked the absence of quantitative targets and/or deadlines for the implementation of the recommendations as a major weakness that renders rigorous follow-up difficult. Moreover, the motivation for both dissemination and follow-up should be seen in the context: if the review itself is perceived as 'waste of time' and its recommendations as useless, and

---

[234] France seems to have been an exception, as the results from the follow-up, drawn up by an independent consultant, were deemed positive: *"I firmly believe in the vital usefulness of the follow-up report […] It allowed us to be reminded about the recommendations, as well as to initiate the new civil servants [that had entered the ministry since the release of the first review report], and those who had not participated in the review, to the themes addressed."*

if the first follow-up report has generated almost no reactions at all, it is hardly surprising if enthusiasm for drafting such a report is modest.[235]

7.1.6 Participation of stakeholders in the review process[236]

The direct involvement of stakeholders in the EPRs is limited to hearings held during the review mission, and commenting the report in the review release event in the reviewed country. However, only the former is 'obligatory', imposed by the review programme, whereas the reviewed country government decides whether or not to solicit comments from the stakeholders once the report has been released. Therefore, the stakeholders do not have great 'ownership' over the review.

Greater involvement of the intended beneficiary groups in the review process was identified already in the survey on the experiences of the first cycle in 1997 as a major means of improving the "marketability, credibility and quality of the reviews" (OECD 1997b). This objective has been reinstated several times since then in the discussions and documents of the WPEP (OECD 1998a; OECD 2003i). Some WPEP delegates have also spoken in favour of involving stakeholders in the peer review meeting.[237] Lately, the interest in stakeholder participation in the process seems to have faded, and replaced by concerns for, on the one hand, strengthening cooperation with other review programmes within and beyond the OECD, and on the other hand, involving stakeholders (notably ministers and high-level officials) at the time the report is being released. In practice, not much has been done to increase participation, and the process remains highly government-driven.

Three among the problems related to participation identified above in section 4.4.4.7 are of major concern in the EPRs. Firstly, in view of the limited budgetary and personnel resources available for the reviews at present, the review approach would have to be significantly modified to achieve genuine participation.[238] The second concern stems from the fact that the conditions of stakeholder

---

[235] *"After the first review, we produced a follow-up report which was sent officially by our State Secretary to the OECD secretariat, but it was never discussed and treated in any way. There will probably not be a follow-up report on the second cycle review. The recommendations are difficult to follow-up, because they are too vague."* (a civil servant from a Ministry of the Environment)
A Dutch interviewee lamented the fact that while the release of a follow-up report was announced in a WPEP meeting, and the report was even officially sent to the OECD secretariat, it was not further discussed at the WPEP.
[236] See section 4.4.4.7.
[237] The 27-head delegation of the Czech Republic included one NGO representative in country's second cycle review on 19 May 2005. This was probably the first time an NGO representative participated in a reviewed country's delegation.
[238] Nevertheless, some participants have considered the lack of resources as merely an excuse for not changing anything; a better prioritisation, focusing and allocation of the existing resources would make significant changes possible.

participation in the review process are largely dictated by the OECD and the administration-oriented code of representation inherent in the reviews. OECD work in general has been accused of supporting an authoritarian, non-inclusive policy approach.[239] While this critique has its relevance, notably for the 'model-weak' actors, and it tends to limit the range of issues that can legitimately be discussed,[240] it probably has much less importance for established and well-organised environmental NGOs or industry representatives.

The third problem relates to the credibility of the review process and the OECD as an evaluator. Key dilemmas facing the OECD concern the trade-offs between the salience, credibility and legitimacy of the EPRs in the eyes of its key 'clients' – the environmental administrations – and the broader range of stakeholders. To ensure salience, credibility, and legitimacy to the environmental authorities the reviews need to avoid excessively extending participation. This choice is coherent with the OECD's character as an intergovernmental organisation. However, restricting participation notably compromises the credibility and the legitimacy of the EPRs in the eyes of the other stakeholders.[241] From the perspective of deliberative democracy, the frequently expressed wish to better integrate various stakeholders in the review process is therefore a double-edged sword: it would probably increase the willingness of the stakeholders to participate in the debate on the review findings, but might ultimately threaten the image of the OECD as a source of reliable, independent, non-political evaluations – an image that the peer review logic crucially depends upon.

The Japanese example illustrates some of the complexities related to participation in the EPRs. The influence of the review was probably enhanced by the fact that in the highly administration-driven Japanese system consultations with other stakeholders are rare.[242] Policy debates are carried out inside the government, which means that the 'participation-deficit' of the EPRs rather strengthened the influence of the reviews by helping it focus on issues that the government perceived as essential and salient. The Netherlands represents perhaps the other end of the spectrum, a country where decisions are prepared in close collaboration between the government and stakeholders. In such a

---

[239] This criticism was manifest in particular in connection with the failed negotiations over the Multilateral Agreement on Investments (MAI).

[240] For example, the review team is unlikely to take seriously arguments calling into question the desirability of economic growth as a basic policy objective.

[241] For example, a Portuguese civil servant noted that *"the industry used [the report] very little: it says that the review gives too positive an image of the administration, and that the OECD was mislead by the Portuguese civil servants."*

[242] This 'administration-driven' policy style was brought up by the Japanese interviewees as a specificity of the country's political system.

context, a government-centred evaluation has less chances of gaining sufficient legitimacy and salience to generate debate.[243]

The EPRs have adopted some of the measures proposed above in section 4.4.4.7 to mitigate the problems related to participation, in particular by aiming at the broadest possible debate after the review, and enhancing legitimacy by providing at least a symbolic possibility for stakeholder participation during the review process. However, the EPRs are not characterised by broad participation in the early stages of evaluation, as recommended by Baron (1999, 141-145). The scoping of the review – the choice of the topics to be covered – is at present largely determined by the OECD review framework, and to the extent that there is still room for choice,[244] it is being exercised jointly by a very limited number of OECD officials and civil servants in the environmental administration of the reviewed country. During the review process, 'communicative rationality' is limited, because of the large number of participants and issues on the agenda. In contrast with the recommendation of Baron (1999, 141-145), participation at this stage is broad, but lacks the depth required for a process that would genuinely integrate stakeholders in formulating the conclusions and recommendations. For instance, the quality of the NGO and industry contributions varies greatly from one country to another: sometimes meetings with the non-governmental representatives are considered highly useful, as they provide a perspective different from that of the government, whereas at other occasions the representatives have not been 'briefed' well enough to have the capacity to contribute significantly, or the review team may no longer be particularly motivated if it already has enough 'hard data' to write the analysis. The way participation has been organised seems to have strengthened, at least in the Dutch review, the phenomenon of 'parapraxis' described by Sager (2001, 207-208): the environmental authorities in the country were first frustrated by the limited influence they had on the review design, and by the lack of communicative rationality in the review process. Many perceived both the sessions with the review team and the final peer review meeting as uninspiring events of one-way communication instead of occasions for real dialogue. Consequently, these actors were not very content with the review findings either (lack of instrumental rationality). The potential 'change agents' lost their interest in using and disseminating the review report, and the capacity of the review to promote dialogue (communicative rationality) remained under-exploited.

---

[243] Against this background, it is somewhat surprising that – according to an expert participating in the Dutch review as an OECD consultant – the stakeholders in general, and the NGOs in particular, were particularly poorly integrated to the review process.
[244] This concerns mainly the choice of the 'special theme' to be dealt with in the second part of the review.

7.1.7 Summary: how much does the review design count?

Many changes could indeed be made to the EPR framework to render the reviews more influential: the recommendations could be sharper and more to the point, the explanatory and comparative aspects in the reviews could be strengthened, the results could be disseminated more broadly, and the follow-up of the review recommendations could be more focused and systematic. However, only on few issues do the WPEP delegates have sufficient agreement and determination required to drive through changes to the review framework. Likewise, some delegates have perhaps not fully understood the need to strike an appropriate balance between often contradictory objectives, such as deepening the causal analysis on selected topics, increasing cross-country comparability, and shortening the review cycle. Making these choices requires a clear understanding of the overall purpose of the reviews. The decisions concerning review design often involve dichotomous choices as those that we see in Table 7.1. All these choices correspond to distinct perceptions of the ways in which the reviews should influence policies. Should the reviews primarily foster accountability of the government towards its constituencies, function as 'advocacy' evaluations advancing the environmental 'cause', or aim at policy improvement through learning and dialogue? In the final analysis, the questions boil down to the choice between – and the respective weights of – learning and accountability as the main purposes of the review. At present, the EPRs try to achieve both, partly because of the contrasting expectations of the different stakeholders.

Table 7.1 Dichotomous choices concerning the review design.

Recommendations	Critical, specific	'Soft', general
Approach to cross-country comparability	Comparative	Country-specific
Thematic focus	Focus on 'hot' issues	Treat all issues in a balanced manner
Breadth of the review	Narrow focus on few themes	Broad coverage
Focus of analysis	Explanation of causal effects	Checking goal-achievement
Participation of stakeholders	Broad participation	Administration-oriented EPRs

On the basis of the above distinctions and the types of influence identified in chapter 6, we can now distinguish four 'ideal types' of pathways of review influence (Figure 7.1). These pathways operate, on the one hand, through *political pressure* inside the government (A) or among a broader range of

stakeholders and the public (B), and, on the other hand, through ***learning*** concerning the evaluated policies (C) or consisting of more general, long-term socialisation, enlightenment, modification of mental models and frameworks of thought (D).

Figure 7.1. The four pathways of influence from the EPRs.

	Accountability/advocacy		Learning	
	A	B	C	D
TYPE OF INFLUENCE	Interministerial 'Low profile'	Public debate Argumentation	Dialogue Best Practices	Socialisation Long-term frameworks of thought
KEY ACTORS	Change agents in the government	NGOs Public pressure	Experts in the government	All participants but especially civil servants
FOCUS OF INTEREST	Cost-effectiveness Environmental policy integration	Goal-achievement	Why-questions Expanation Understanding	Agenda-setting Frameworks of thought Status of environment in sectoral policies

Pathway A – *'advocacy through intragovernmental debate'* – to a large extent corresponds to the expectations of many civil servants in the reviewed countries, and represents 'soft coercion' in its purest sense. Gathering together, around environmental issues, a large number of civil servants from different sectors of administration, the review process and the recommendations from a respected intergovernmental economic organisation give more visibility and weight to environmental policies and authorities in their attempts to improve the legitimacy of environmental policies, thereby helping them in the inter-ministerial negotiations. The review reports and the recommendations are

appreciated since they apply diplomatic language, stick to the facts, avoid confrontation, and have the capacity to appeal to actors that would be unlikely to take seriously pure 'environmentalist' arguments. Policy implementation, cost-effectiveness of policies, and sectoral integration are among the likely themes of interest.

Pathway B – *'advocacy/accountability through public debate'* – in turn, represents a 'tougher' version of advocacy/accountability, since the main influence stems from public debate and argumentation among a broad range of stakeholders – the media and the NGOs playing a key role. Rather than procedural issues and administrative effectiveness, the main focus of interest among the stakeholders is goal-achievement in the spirit of accountability evaluation: while cost-effectiveness of policies is a part of the government accountability towards its citizens as taxpayers, the public at large and the NGOs are likely to be more interested in whether the government has delivered on its promises.

To be effective, pathways A and B both require that the OECD be perceived as an organisation, whose message carries political weight and has the capacity to persuade through public or intra-governmental pressure. Pathway C – *'learning from best practice'* – by contrast, does not assume such a political role for the OECD, but sees it mainly as a forum for dialogue, facilitating mutual learning on rather technical issues of policy implementation. As opposed to the accountability/advocacy pathways, where the idea was to exert pressure on the government to ensure the attainment of policy objectives and the establishment of cost-effective implementation mechanisms, here the reviews engender dialogue on why policies have produced the observed results, why certain policies have failed and others succeeded, and in this way promote innovation and learning from best practices.

The EPR programme has been designed to exploit all of the three pathways mentioned above. However, for a number of reasons, which we shall explore more in detail further down, it is not uncommon that none of the three really operates in an optimal manner. This calls attention to the pathway D – *'long-term conceptual learning'* – where the reviews gradually shape the frameworks of thought, and the relative importance of different policy issues on the policy agenda, notably by strengthening the status of environmental issues in sectoral policies and rational, results-oriented, efficiency-oriented thinking in environmental policy. While the motivation of 'change agents' to carry forward the OECD message certainly facilitates this type of impact, its absence is not fatal, unlike in the three other pathways. Already the participation of a large number of stakeholders in

the review process – participation that is rather intensive for those most immediately involved – cannot pass without shaping the ways of thought. Yet, often the review serves to corroborate the existing mental models rather than radically changing them.

The above discussion has already shown that different actors have varying preferences concerning the preferred pathways of review influence. In fact, the experience from the WPEP and the interviews showed that the opinions vary not only in function of the actor group and the person involved, but that many individuals in fact underlined the importance of several of the above pathways. While a number of interviewees and participants in the EPR programme clearly favour one or the other of the pathways, a typical comment concerning the choice between the purposes of learning and accountability went along the lines: *"I would not make such a distinction – both are important and essential elements of the reviews."* To a large extent the absence of clear preferences reflects the fact that the character of the EPRs as a combination of the two purposes is seen as one of the strengths of the programme. However, the inability to choose also reflects the power relations between the different policy communities, and the mutually contradicting beliefs, interests, and expectations of different actor groups and even individual actors. In the following, the preferences of different actor groups and actors are analysed in the light of our discussion on 'repertoires' and 'epistemic communities' in order to better explain the preferences and arguments behind the different review designs and expected pathways of influence.

7.2 Actors with their networks and expectations[245]

Among the key actor groups influencing the EPRs are the following:

- OECD Environment Directorate (in particular, its Environmental Performance and Information Division – EPI), in charge of the practical management of the review programme;
- reviewed country government (especially its environment ministry), as the main 'client' of the reviews;
- OECD Working Party on Environmental Performance (WPEP), as the body that assembles for the peer review meetings to discuss and approve the conclusions and recommendations

---

[245] See section 4.4.2.

for individual country reviews, as well as to plan and decide upon the management of the review programme;

- OECD Environmental Policy Committee (EPOC), as the body supervising the WPEP and deciding on the practical organisation of OECD work in the field of the environment;

- stakeholders in the reviewed country: industry, environmental NGOs, environmental policy researchers and, ultimately, the public at large.

In addition, the Economics Department and the EDRC group affect the EPRs indirectly, through their central position in the OECD and the sustainable development work in the Economic Surveys.

The potential main beneficiaries of the reviews are the environmental authorities of the country and the officials working with environmental issues in other sectors. The role of the non-governmental stakeholders is much more peripheral. To understand why this should be so, the following section looks more closely at the power constellations between the policy communities within which the different stakeholders operate. While the questions concerning the allocation of budget and personnel resources merit attention as framework conditions affecting the reviews, the main focus in the following will be on two less visible forms of power: the positions of the actors in the networks, and the symbolic, ideational power.

7.2.1 Environmentalist and economist policy communities

We can distinguish two main policy communities around the EPRs. Firstly, the 'environmentalist' community is organised around the ministries of the environment, and encompasses, apart from the environmental administration, the OECD Environment Directorate and at least a large part of the environmental NGOs and environmental researchers. Secondly, the 'economically-oriented' community is congregated around the ministries of finance, and includes as its main non-governmental stakeholders a large proportion of the economic scholars, as well as the OECD Economics Department. The latter community looks at environmental policies primarily from the point of view of economic efficiency, and draws its policy arguments ultimately from economic theory. The environment is seen as one among the many factors affecting economic welfare.[246] For

---

[246] Two arguments can be put forward to prove the existence of such a community. First, the assumption is supported by comments by persons involved in the Economic Surveys, such as: *"I wouldn't say the reviewed country government or its finance ministry have excessive influence on the outcome of the Survey, but the impact comes indirectly, through the general philosophy and framework of thinking that is shared by most of the EDRC delegates, finance ministries in the member countries, and the OECD Economics Department"* or: *"The influence of the Ministry of Finance rather stems*

the environmentalist community, by contrast, environmental policy objectives have a priority over the economic ones, and the epistemic basis for the arguments ultimately draws upon natural sciences and their definition of environmental problems.[247] These communities are 'epistemic' to the extent that they share a common policy enterprise, values and epistemic beliefs. For both communities, the OECD represents a potentially useful instrument in helping gain more influence and resources.

Within the OECD secretariat, the struggle between the two epistemic communities manifests itself in the tensions between the Economics Department and the Environment Directorate, the former representing the hegemonic, ideological core of the OECD. From the perspective of the Public Choice theories, the relationship can be seen as a competition over the distribution of resource power, each community attempting to maximise its tasks and resources.[248] However, while the assumption about maximising the tasks and resources may explain a lot of the behaviour of the Environment Directorate, which continuously has to fight for its place within the organisation, the Economics Department can be much more confident about its position, and therefore does not need to fight for each and every individual work opportunity. On the contrary, the relative reluctance of the EDRC to take on the task of analysing sustainable development in the Economic Surveys supports the assumption that the Economics Department actually tries to keep its hands off issues that fall outside its specific sphere of competence. However, the example also suggests that along with the desire to expand budgets, the international bureaucracies do indeed have their own preferences concerning the substance matter. The Environment Directorate and the Economics Department differ in their fundamental approach to environmental policy, and these differences are much easier to explain by 'epistemic' considerations, notably the different educational and professional background of the persons working in the respective directorates, than by explanations referring exclusively to their self-interest in terms of expanded power and resources. While the

---

from the fact that the economists working in the ministry share largely the same framework of thought as their counterparts at the OECD Economics Department. One can thus talk about very homogeneous global markets for economists – a market which is greatly influenced by academic research in economics." Second, even if many members of the economic community did not recognise that such a community would exist, this would not change the fact that systematic reference was made in the interviews and in the discussions within the WPEP to the environmentalist-economist distinction, and that such communities exist at least in the consciousness of many of the actors. More important than the existence in 'reality' of such communities is that the participants involved believe in their existence and behave accordingly.

[247] While economics obviously play an important role in the EPRs as well, these reviews rely on the scientifically defined environmental problems and the politically established policy objectives, rather than economic considerations, as the basis for assessment.

[248] A simple comparison of the number of permanent expert-level staff working on the Economic Surveys (30) with that of the Environment Directorate officials managing the EPRs (4), gives an indication of the relations of power between the two directorates, measured in terms of resources.

Economics Department, in drafting the sustainable development sections of the Economic Surveys, mainly looks at the cost-effectiveness of environmental policies (e.g. OECD 2004a),[249] the EPRs have a broader set of criteria, with environmental performance as the first concern, and cost-effectiveness only one of the 'secondary' criteria. An alternative explanation for the lack of interest of the EDRC in sustainable development draws on more identity-related factors: the committee was described by one interviewee closely involved in the Surveys as 'fiercely independent', and therefore hostile towards any attempts by 'outsiders' to dictate what kind of work the group should undertake.

It would seem that the Environment Directorate has not been very successful in defending its budgets and resources, especially for the EPRs, since the personnel resources available for the programme have not increased over the 13 years of the programme's existence.[250] By contrast, the tasks have increased, with the entry of new members in the OECD, and the carrying out of EPRs in some non-member countries.[251] To defend its budget, the Environmental Performance and Information Division of the OECD must remain in the 'mainstream' line of thinking of the OECD,[252] but above all needs to obtain the support from the member country delegates, notably in the EPOC. Active lobbying[253] has indeed brought good results within the EPOC, and the continuation of the EPR programme has been secured so far.

The EPRs can be seen as one of the tools that the WPEP, through the Environment Directorate, uses in order to influence policymakers, by strengthening and legitimising the less powerful 'environmentalist epistemic community' in relation to its more powerful economic counterpart. The reviews also seek common ground between the two rivalling epistemic communities. As we saw in

---

[249] *"In the beginning, when the framework for analysing sustainable development in the Economic Surveys was being developed, the intention was to look at the issue from a broad efficiency perspective allowing the comparison of social, economic and environmental policies and objectives on a common scale of measurement. This idea was soon abandoned, however, because it would have been politically too difficult. We wanted to be more modest as to what the reviews could achieve, since we did not have the tools for making such efficiency comparisons."* (a person involved in the Economic Surveys)

[250] The budget of the Directorate has declined over the past ten years by about 30 % - more than the reductions in the overall budget of the OECD. The main part of the decline has, however, been compensated by increased voluntary contributions from the member countries. (Personal communication, Christian Avérous, head of the Environmental Performance and Information Division, 26 January 2005).

[251] Bulgaria (1996), Belarus (1997) and Russia (1999) were reviewed in cooperation with the UN-ECE, and Chile (2005) together with UN-ECLAC.

[252] This idea – the need to avoid being marginalized by deviating too much from the mainstream policy doctrine represented by the Economics Department – has come up in many of the informal discussions with the Environment Directorate officials.

[253] Such lobbying may take the form of the WPEP chairman urging the delegates to remind their superiors about the high quality of the programme and talk about the resource needs; official letters from the chairman and the secretariat to

the previous chapter, strengthening the legitimacy of the environmental policies and policy actors in the reviewed countries is one of the main functions that national environmental authorities and policy actors attribute to the EPRs. The reviews 'empower' the members of the 'environmentalist epistemic community' in the reviewed countries and may indirectly strengthen the position of the OECD Environment Directorate within the OECD secretariat in relation to its more powerful economist counterpart. On the other hand, the EPRs serve to infiltrate 'economist' ideas to the national environmental administrations, thereby strengthening the 'economist faction' within these administrations.[254] Indeed, the fear of a 'contamination' of the environmental policy discourse by an economic one may be one of the main reasons for at least some of the NGOs having doubts about the OECD approach.[255] None of the interviewees denied, nonetheless, that policy recommendations coming from an economic organisation tend to carry much more weight in policy debates than those emanating from the 'environmentalist epistemic community', but many regretted the fact that the public is seldom aware of the OECD environment work.

The dilemma facing the 'environmentalist policy community' is how to balance between the need to safeguard the integrity of the environmentalist discourse on the one hand, and to influence the dominant discourse on the other. The fact that the Environment Directorate is willing to consider the views of the Economics Department, up to a point of sometimes approaching self-censorship when drafting the EPR reports, can be seen as one consequence of the unequal distribution of power within the organisation.[256] On the other hand, the 'softness' of recommendations referred to above may partly stem from the need to ensure the support from the environment ministries: the OECD secretariat fears that too harsh a critique might diminish their interest in the reviews.[257]

---

the governments requesting voluntary contributions; or informal discussions with delegates at the margins of the OECD meetings.

[254] A civil servant working in the economic field alluded to such an impact saying: *"One function of such independent reviews might be their capacity to reveal the 'stupidity' of certain policies, a task which cannot be achieved simply through analyses made by national actors and institutions."* Civil servants working in the environment ministry with economic matters, and those in the finance ministry responsible for environmental issues recognised the usefulness of the EPRs and the OECD work in general, as tools for strengthening economic reasoning among environmental authorities.

[255] *"I have a positive attitude, but the NGOs sometimes may think the OECD has an attitude biased in favour of economic interests."* (an NGO representative)

[256] A typical example from the EPRs is when the OECD general policy line of eliminating subsidies is taken to its logical extreme, the EPRs calling in question even subsidies designed to strengthen the competitiveness of environmentally sound technologies or energy sources, as was the case in the review of Sweden, for example (see OECD 2004c, 22).

[257] One can suspect that the need to treat kindly the main contributor to the organisation's budget partly explained the cordial attitude and the willingness of the WPEP to accept almost all of the comments from the reviewed country in the review of the U.S., in May 2005.

Looking at the question from the opposite angle, one could ask to what extent the EPRs are able to influence the general OECD policy line, and affect the hegemonic position of the 'economist' community in the organisation. This influence is likely to be modest, considering that while the draft EPR reports are habitually sent for comments to the Economics Department, until recently, the reverse was not true: comments on the Economic Surveys were seldom solicited from the Environment Directorate. According to an OECD secretariat official, the situation has changed, as the Economics Department regularly uses the EPRs in drafting sustainable development sections in the Economic Surveys, and comments on the draft Economic Surveys are regularly sent for comments to all OECD directorates. To what extent the comments are taken into account, is uncertain, however.[258]

7.2.2 Expectations and 'repertoires' of actors[259]

Above, we have already touched upon the question of symbolic power embedded in epistemic communities. The concept of 'repertoire' can be used to give the issue more content, and analyse the epistemic dimensions in relation to the EPRs. We can distinguish the following key repertoires (table 7.2), which cut across the epistemic communities described above.

---

[258] The willingness of the Economics Department to take into account comments from the Environment Directorate has been called into question by WPEP delegates who have at several occasions regretted the obvious lack of coordination between the EPRs and the sustainable development sections of the Economic Surveys.

[259] See sections 4.4.2.2 and 4.4.2.3.

Table 7.2 Repertoires of the main actors.

Ministry of the Environment
OECD-oriented 'environmental economics' repertoire
• Environmental objectives have priority over the economic ones • Endorsement of the OECD policy doctrine and advocacy for the use of market-based instruments • The main function of the EPRs is to 'sell' the idea of results-oriented, more market-based environmental policy approach within the Ministry of the Environment and the sectoral ministries • OECD is highly credible and useful as an organisation promoting economically more rational environmental policies • The OECD 'quality label' is useful in political advocacy – i.e. in advancing the legitimacy of environmental issues within the government
Mainstream, 'floor-level' repertoire
• Environmental objectives have priority over the economic ones • A balanced mix of instruments should be used in environmental policies; scepticism towards excessive emphasis on economic instruments • OECD is a highly credible and important organisation, but not necessarily very relevant in political terms • EPRs are not very useful, because they do not tell us anything we would not have known already • EPRs require a lot of work, with very uncertain returns; in fact, the potential criticism in the reviews may threaten our legitimacy and complicate our work
Sectoral ministries
• OECD/environmentally oriented repertoire • Primacy of sectoral policy objectives • EPRs help gain support, visibility and legitimacy for our own work on environmental issues • EPRs provide valuable lessons on the broad picture of our country's environmental policy • OECD work in general and the EPRs in particular allow to maintain environmental policies on a rational path, and help avoid 'stupidities'
Mainstream, floor-level repertoire
• Primacy of sectoral policy objectives • EPRs, as well as evaluations in general, represent an additional burden whose usefulness is often difficult to see • Our task is to 'get the facts right' and defend our ministry against unjustified attacks from 'environmentalists' • EPRs may complicate matters, especially if civil society actors seize them to attack our policies
Environmental NGOs
• Primacy of environmental objectives • OECD has high credibility, but as an economic organisation, its policy messages are somewhat biased • OECD could in principle be used to advocate for the environment, but we do not know the organisation's work very well; its reports are technical and not easily accessible • EPRs do not contain politically 'hot' messages, they lack criticism, and are written in a somewhat technical language, which reduces their appeal to an NGO
OECD Environment Directorate
• Primacy of environmental objectives • To continue work, need to remain in the OECD 'mainstream', but also ensure support from environment ministries, especially those of the most powerful countries[260] • Have an interest in maintaining the 'status quo' in relation to the way the EPRs are carried out[261] • The role of the EPRs is to improve the legitimacy of environmental policies, and instil (economic) rationality to environmental policies through accountability evaluations carried out by a prestigious international organisation • The power and the influence of the EPRs stem from their credibility and independence; therefore it is vital to ensure the facts are correct and give the impression that the review team is fully independent from the national

---

[260] As voluntary contributions cover a growing share of the programme costs, it has become increasingly important for the secretariat to ensure continuous support from the countries supporting the EPRs financially or through 'secondments' of personnel.

[261] It has been suggested that, over the years of WPEP work, the secretariat has been somewhat reluctant to accept major changes proposed to the EPR framework (see section 7.2.6 below).

stakeholders

The environmental administrations can be seen to have the strategic interest of gaining legitimacy for their policies, and the cognitive interest to learn about 'best practices' and other countries' experiences. Different actors attribute varying degrees of importance to these two objectives. The legitimisation objective certainly dominates,[262] but learning was very strongly put forward for instance in the Netherlands.[263]

Often, the EPR is, however, mainly perceived as an obligation that has to be fulfilled, resulting in a desire to hide negative aspects, and to get through the process with as little trouble as possible. This type of thinking is probably frequent in sectoral ministries, which may fear critique from the OECD and potential attacks by the NGOs.[264] However, as was seen in section 6.3.6, there seems to be resistance also among the environment ministry officials.

The environmental NGOs, in turn, are likewise motivated by the desire to learn and particularly to educate the citizens, but as pressure groups their main interest is political: to use the EPR as a source of criticism against the government policies or, when appropriate, to support environmental authorities against resistance from other sectors. The interest of the media is similar, i.e. they are likely to seek in the reviews 'hot' revelations.

---

[262] *"Learning results from the work of the other groups in the OECD, rather than from the [Environmental Performance] Reviews whose function is primarily to exert pressure."* (a civil servant from a Ministry of the Environment)

"For the Ministry of the Environment, which is still a young ministry in its present form, the review and its recommendations will certainly be useful thanks to the support they can give to our own demands." (a civil servant from a Ministry of the Environment)

[263] The following statements clearly illustrate the Dutch perspective:

"The main benefit we could extract from the OECD work is that of dialogue on policy problems so as to promote learning; how have other countries dealt with the problems we are struggling with and how do OECD experts see these questions? The accountability function of the EPRs is not at all important for us."

"The OECD is in the Netherlands perceived as a research organisation, instrument capable of promoting learning, rather than being a political 'heavy weight' player. We tried to use its political power during the first review and succeeded, because the conditions were favourable: we were among the first countries to be reviewed, environmental issues in the OECD were relatively new, the EPR programme was a new invention, the environment was high on the policy agenda in the Western countries. […] OECD can be in the political 'heavy weight' class as an innovator, but once an issue becomes 'mainstreamed', the organisation's role becomes more technical, while other actors (e.g. the EU) overtake the roles of political importance."

"We don't need the OECD as a source of political pressure: we have active NGOs and research institutes (especially RIVM) that produce their own reports and criticism towards the government."

[264] *"For the Ministry of the Environment the review is probably useful as a 'weapon' in the intersectoral fights. For us, the review does not have this utility, because the critique it contains is directed at us."* (a civil servant from a sector ministry)

"Other ministries had a very hard line in the preparations for the [peer review] meeting. These preparations included various phases of negotiation." (a civil servant from a Ministry of the Environment)

Apart from the distinction between the 'environmentalist' and the 'economist' elements in the repertoires, following the division between the epistemic communities, we can therefore classify repertoires into the 'bureaucratic' and 'political' ones. In the former are the different ministries and state bureaucracies, focused on policy effectiveness, using rather 'dry' and technical language, and often stressing the learning aspects of the reviews. The media and the environmental NGOs, in turn, fall within the 'political' repertoire, seeking in the reviews mainly outspoken criticism, and clear messages on 'hot' issues. The bureaucratic repertoire can be considered more sensitive to criticism, more accustomed to deciphering in the EPR reports the meanings behind what is considered by many other actors as 'dry' language, replete with 'empty words', and more eager to stress the benefits of interministerial debates carried out in a relaxed and confidential atmosphere.[265] This distinction probably also explains the diverging views among the Dutch participants on the tone of the EPRs and the reports prepared by the country's environmental research institute, RIVM (see section 6.2.2): the argument considering the EPRs too consensual saw their potential mainly in terms of evaluations useful in generating public debate, whereas the argument that the RIVM reports lacked the balanced carrot-and-stick approach of the OECD reflected the view that the reviews should primarily operate through 'soft' coercion within the government. The ministers typically straddle between different repertoires, in many cases falling within the conflict-laden, political one, while at other occasions adhering to the bureaucratic repertoire, and yet in some situations focusing on the need to achieve consensus among different sectors of government and/or actors involved.[266]

The EPRs could enhance learning across different groups by challenging different repertoires and helping actors understand each other. As van der Meer (1999, 390-392) notes, learning often takes place as a response to reactions to evaluation by groups representing different repertoires, rather than by responses to the evaluation itself. However, for such learning to take place, there needs to be a forum for and interest in debate. At present, the EPRs mainly appeal to the bureaucratic

---

[265] *"In fact, it may be rather an advantage that the OECD is not very well known by the broader public, because this way the organisation does not become a forum for 'hot' negotiations. We need an organisation, where politically hot issues can be discussed calmly, sheltered from publicity and the accompanying agitation."* (a civil servant from a Ministry of the Environment)
"The advantage of an exercise such as that of the OECD review is that it provides an 'outsider' view of questions that are often marked by conflicts and stalemates, thereby helping treat these questions from a more 'objective' and dispassionate perspective." (a civil servant from a Ministry of the Environment)
[266] An example of a minister acting within the 'political' repertoire was when the Portuguese minister defended the government's nature conservation policy. The Swedish minister's decision to launch a reform of the country's environmental inspection and enforcement system falls rather within the bureaucratic repertoire: the OECD had detected in the country's environmental policy system a 'technical' problem, which the minister decided to address. Finally, the attempt to achieve consensus among different actors – while present in the political repertoire as well – was clear in the efforts of the German and the French ministers to gather support for green tax reforms.

repertoire, without generating any substantial reactions by actors within the more conflict-oriented political repertoires. Thereby they clearly prioritise the objectives of finding common ground between the two bureaucratic repertoires, strengthening the two minorities, i.e. the 'OECD-minded' repertoire in the environment ministries and the OECD/environmentalist repertoire in the sectoral ministries.

The team tries to exploit the dominant repertoires to increase review influence, and particular attention is given to meetings with high-level officials and ministers.[267] However, this attempt may have the disadvantage of producing overly consensual reports, as noted above.

7.2.3 Change agents, position of the WPEP delegate and 'support from the top'

The experience from the EPRs suggests that the existence of 'change agents' motivated and capable to carry forward the review message in the national administration is a necessary requirement for their influence. The country's WPEP delegate has an important role in this respect, as a permanent link between the review programme and policymaking at home. It is unlikely that the WPEP delegate could obtain sufficient resources for the dissemination of the report and the implementation and follow-up of recommendations without good access to the ministry's leading officials.[268] In Portugal, it was suggested that a change of the WPEP delegate's position decisively affected the capacity to disseminate the review. In a small country, the change in the position of an individual expert may have profound consequences, given that specific themes often become personified in the only individual in charge of the issue. In France, one unit of the Ministry of the Environment, committed to carrying the 'OECD message' forward, has the potential to function as a change agent.

The somewhat disappointing experience from the inclusion of a sustainable development section in the Economic Surveys[269] likewise reminds of the importance of motivation: evaluation influence feeds on the motivation of the key persons involved. The lack of enthusiasm for the review within the administration certainly reduced the impact of the EPR in the Netherlands, whereas obvious

---

[267] These discussions are often carried out as lunch meetings with no more than 2-3 review team members participating.

[268] The lack of direct access to and influence on the ministry's top management by the persons organising the EPR was evoked as a possible reason for the modest impact of the first EPR of Finland.

[269] None of the interviewees considered the experiment as an incontestable success, whereas the most critical views within the WPEP group saw it even harmful for the 'sustainability agenda'. The lack of motivation and expertise on sustainable development within the EDRC, and the almost exclusive focus on cost-effectiveness of environmental and

resistance against the review and reluctance to take on further evaluation duties has in some cases contributed to the postponement of the review process (in Finland).[270] The reasons for lacking motivation may be numerous, the belief that the review would be 'useless' being perhaps the most common one.[271] Others include the tiredness of constantly being forced to participate in evaluations and fulfil the numerous reporting obligations,[272] as well as the lack of familiarity with the review programme, resulting from changes in the government personnel, for instance. Fear of being criticised was also mentioned as a possible explanation.[273] Resistance and lack of motivation can therefore be interpreted as a result of a failure of the EPRs to link with the dominant repertoires, notably those of the potential key change agents.

A factor reducing the potential of the reviews to have an influence is that the information often does not reach beyond a small number of individuals participating in the process. Environmental issues in sector ministries are often the responsibility of one single individual. The influence of the EPRs therefore crucially depends on the position of this person in her ministry, as well as on her eagerness to bring forward the lessons learned. In one country, the relative lack of influence of the review was attributed to the excessive cautiousness of the organisers in the face of potential criticism.[274] Knowledge of the OECD working methods and culture is therefore an essential condition without which an actor is unlikely to function as a change agent, regardless of whether she operates within the Ministry of the Environment or in other sectors of administration. The sheer lack of awareness of and interest in the OECD activities among politicians, civil servants, and the public may compromise the influence of the review.[275] On the other hand, the EPR process itself is

---

social policies were among the main shortcomings evoked by the interviewees. By contrast, all informants were content with the synthesis report on the experiences of the sustainable development sections.

[270] The official reason for the postponement was that the EPR process would have disturbed preparations for the Finnish EU presidency, scheduled for the second half of 2006.

[271] *"The effort required to organise the review was perceived as excessive in relation to the expected results. Probably it was not even totally clear what the result would be. The ministry did not consider the EPR important enough to justify real investment in the process."* (an environmental policy researcher)
"I already knew more or less what to expect, on the basis of my experience from the first review […]. The OECD is an important organisation, but I knew already that our policy would not be greatly influenced by the EPR." (a civil servant from a non-environmental ministry)

[272] *"All kinds of monitoring and reporting obligations for UN Conventions, the EU, indicators, etc. require a lot of work. Intergovernmental organisations should cooperate more closely so as to coordinate their action."* (a civil servant in the environmental sector)

[273] *"It might also be that the ministry had something to hide, given that the progress over the preceding few years in the environmental policies had not been startling."* (a Dutch interviewee)

[274] *"The person in charge of coordinating the review did not understand its political character and potential to create pressure and get more resources for the Ministry of the Environment; she was far too shy and cautious, trying to avoid any criticism against the government policies. The Deputy Director General of the ministry understood the political role of the EPRs, but once she was moved to another position, the whole set-up fell apart."* (an OECD official)

[275] *"In my view, [our country's] environmental administration clearly overlooks the OECD, and does not take full advantage of the services the organisation could provide. This is an opinion we have made clear to the top management of the Ministry of the Environment."* (an official from a permanent delegation to the OECD)

a way of making publicity for the non-economic activities of the OECD within the environmental administration and other sectors.

Finally, for a change agent to fulfil her 'duty', backing from the top level of the organisation is primordial. The lack of top-level support clearly diminished the potential of the reviews to have an influence in Portugal and the Netherlands, whereas the strong engagement of the ministers seems to have enhanced review influence in France, Mexico and Sweden.[276] Again, such support largely depends on the degree to which top managers and ministers are familiar with and appreciate the OECD working methods.[277]

7.2.4 Role of the OECD as evaluator[278]

The above discussion has already demonstrated the importance of 'repertoires', epistemic communities and principal-agent relationships in conditioning the influence of the reviews. The OECD is not, however, an evaluator just like any other, but its prestige and reputation are crucial determinants of the impact of the EPRs. Let us now look at the role of the EPRs in the OECD's work as a creator of norms and identities, their place in the 'norm life cycle', and the degree to which the OECD plays – through the EPRs – one or several of the four roles described by Marcussen (2001) as typical for the organisation.

Perhaps the single most important factor shaping the influence of the EPRs is the OECD's image and perceived independence and credibility as an external evaluator. All but one of the interviewed persons underlined the high prestige the OECD enjoys among the stakeholders, as well as its independence and competence especially in producing reliable analytical information in various fields of public policy. However, the OECD's role is changing as a result of transformations in the international policy environment, and the organisation has been going through somewhat of an 'identity crisis', as was evoked above in chapter 5.

---

[276] *"The support from above (Ministers Lepage in 1996 and Lepeltier in 2004-2005) was essential for the success of the review. In consequence, the civil servants were ready to cooperate, either because they were more motivated thanks to ministerial support, or because they felt obliged to do so."* (a French civil servant)

[277] *"Our vice-ministers, for instance [vice-minister X], use the EPR recommendations to give more weight to their arguments. This is certainly partly because he knows the OECD work, has a long experience from high-level posts at the ministry, and has therefore acquired a sufficient 'political eye' to know how to sell his ideas."* (a WPEP delegate)
By contrast, the impact of the first Finnish EPR was said to suffer from the environment minister's lack of experience and interest in OECD work.

[278] See sections 4.4.1, 5.1.2 and 5.1.2.

The independence of the OECD in carrying out EPRs is seldom called in question by the 'insiders'. By contrast, especially the independence of the Economic Surveys has sometimes been questioned (e.g. Ingen risk för…1998; Vi måste satsa...1998; IMF:n ja OECD:n Suomi-kuvat…2000), even though this might stem simply from the higher frequency and political status of the Economic Surveys. The environmental administration of the reviewed country does indeed influence the contents of the reviews at least indirectly, often by reminding the review group throughout the process about certain key questions, but this is seldom seen as a serious problem for the 'integrity' of the EPRs. The review team seeks to be sensitive to the wishes of the reviewed country,[279] and the chairman of the WPEP routinely explains in the peer meetings that the aim is to render the review as useful as possible to the member country government. However, since most of the informants for the present study consist of 'insiders', i.e. people more or less closely associated with the EPRs, the conclusion concerning the perceived independence of the reviews must be taken with a grain of salt.[280]

The OECD hardly plays adequately any of the roles in Marcussen's (2001) typology described above. In the light of the interviews, which suggested the OECD is perceived as highly independent, the EPRs should have a good potential of being an example of OECD as an *'ideational artist'*, operating as a 'norm entrepreneur'. However, as we argued in the previous chapter, the reviews very seldom provide genuinely new ideas, as this role would imply.[281] The EPR programme as a whole was certainly in its beginning an innovation representing the OECD in its role as an

---

[279] The informants expressed few complaints on this matter. One civil servant from an environment ministry lamented the fact that *"[t]he OECD secretariat proved to be very inflexible in adjusting the recommendations according to our wishes: we would have liked to modify the recommendations [in the draft report sent to the country prior to the peer review meeting] in order to render them more useful for our own work."* Another interviewee from a non-environmental sector noted with satisfaction that there was *"a big difference between the first and the second review. In the first review, the team arrived with a predefined idea of the policies of [our country]. In the second, they showed greater openness in their thinking, greater understanding, and more willingness to learn."*

[280] Indeed, perceptions may be quite different even among presumably well-informed persons who look at the programme from some distance. In an informal discussion, an official from another OECD Directorate cast a doubt on the EPRs' independence by remarking that the review team seems to write in the report "precisely what the reviewed country wants". Furthermore, an official from a permanent OECD delegation expressed his concern for the OECD peer reviews in general by stating: *"What I'm worried about is the integrity of the OECD in giving its judgements and recommendations in the peer reviews […] It is of utmost importance that the work of the organisation be perceived as independent and reliable, that the results of peer reviews, for instance, not be considered as manipulated by the reviewed country. At present, I think this integrity is not always clearly protected. There are probably a great number of worrying examples of situations where the OECD has been too flexible. In addition, the fact that certain country reviews are partly financed directly by member countries is a growing problem, and can gradually weaken for instance the confidence of the media [in the reviews]"*

[281] In fact, some informants criticised not only the EPRs, but the entire environment work in the OECD for the incapacity to come up with new ideas. Similar critique has been expressed concerning other OECD work. For instance, the publication of the OECD review on Finland's competition and regulatory policies was criticised for not providing any new perspectives beyond those that are already common currency in its member countries, and thus failing to generate public debate (Englannin kieli…2003).

ideational artist, but has now become 'mainstreamed', and the idea is being taken over by other actors. There are signs that the interest of many 'core' OECD countries in the EPRs as a politically powerful instrument would be declining somewhat, while other international actors are becoming interested in the concept. This could be interpreted as the OECD playing the role of an 'ideational artist' in relation to the outside world.[282]

The peer review logic is fundamentally based on the idea of the OECD operating as an *'ideational arbitrator'*, promoting 'norm cascade' and learning through the processes of socialisation, imitation, and coercion. While this role was emphatically endorsed by many informants, who saw the EPRs as a potential forum for learning through dialogue, this was also among the areas where the reviews were seen to fail the most blatantly, as shown by the criticism concerning the lack of dialogue throughout the process, and in the peer review meeting in particular. The informal discussions at the margins of the meetings certainly help forge friendships, and possibly lead to new initiatives at home, but the concrete policy impacts largely depend on the way in which the country delegates behave back home, what is their position in the national administration, etc.[283] The other function attributed to an 'ideational arbitrator' – the potential of the EPRs to exert 'soft coercion' over the reviewed country – fails to satisfy many participants, as shown by the criticism concerning the excessive 'kindness' of recommendations, and the overly diplomatic tone of the reviews. Again, these critical views must be qualified, because even many among those who regretted the absence of coercive power and sanctions as a weakness in the OECD work, at the same time underlined the positive side of the OECD as a 'nice' organisation, whose influence and appeal rests precisely on its image as a non-coercive actor, capable of promoting policy learning through dialogue – even on sensitive issues – in a non-adversarial atmosphere (see section 7.2.2). The key question for the EPRs is therefore their capacity to balance between praise and criticism in promoting socialisation and imitation. This study has brought to the fore in particular the importance of legitimisation as a

---

[282] In the previous chapter, we already saw that the United States has recently suggested that the EPR programme should focus more on reviewing developing countries, and reduce the frequency of OECD country reviews. Moreover, in countries such as Mexico (especially in the first cycle review) and Chile, the political importance of the EPRs has been great, and several ministers have participated in the release events. Persons involved in the Economic Surveys suggested a similar development has taken place in their field. One interviewee pointed out, for example, the high status of the Surveys in the new member countries by noting that *"Slovakia […] had – during the two-year interval separating its first review from the second – implemented virtually all OECD recommendations, which would be quite unheard of in the case of the other countries."*

[283] *"The civil servants and experts that participate in OECD meetings learn a lot in the process, and take the lessons with them back home. […] Of course, the degree to which the experts take the lessons back home depends on many factors such as the personal characteristics and the position in his/her ministry of the delegate."* (a civil servant involved in the Economic Surveys)

key function and means of socialisation. The EPRs' capacity to promote 'complex learning' through the provision of new approaches and perspectives to national policies – a function mentioned by many as a major contribution of OECD work in general – was not seen to play a major role in the EPRs.[284] In this perspective, the EPRs serve as a 'reminder' of the well-known OECD policy principles rather than as a source of new ones.

In the EPRs, the OECD hardly plays the role of an *'ideational agent'*. Even though it is clear that the secretariat devotes more time and effort to reviewing large, economically and politically important countries than the smaller ones, the interests and ideas of the most powerful OECD countries do not seem to have had any disproportionate impact on the programme so far, and the U.S.'s new 'unilateralism' has not yet influenced the EPRs to any appreciable extent. However, the differences between the countries may rather appear in the care taken by the team to formulate the recommendations so as to ensure that the most powerful countries do not lose their interest in the reviews.[285] In other words, the balance between praise and criticism leans more towards the former when the most powerful countries are being reviewed.

By contrast, as an *'ideational agency'* the OECD indeed tries to 'smell' the general policy atmosphere in the member countries, and pick up new ideas to the review framework. Yet, many participants have deemed the rate of adoption of novel ideas frustratingly slow, and many have also seen the capacity of the reviews to respond to 'emerging issues' as inadequate. The EPRs seldom innovate on their own, but rather pick up issues as they arise on the general OECD policy agenda (e.g. indicators of 'decoupling' and total material flows). An area where the EPRs clearly could play a more active role is in 'repackaging' the general principles. The criticism against the reviews for their failure to provide anything more concrete than recommendations urging countries to apply the general OECD principles shows that the demand for an 'ideational agency' exists.

In terms of the discussion on the 'norm life cycle' (4.4.4.1), in principle, the EPRs operate at the level of 'norm cascade', pushing for the generalisation of certain international norms through the mechanism of peer pressure. Often, however, the EPRs role is rather that of paving the way for the

---

"The national experts learn immensely in the various OECD meetings; it is really a training organisation for the civil servants participating in the work. However, it is less certain that the lessons would be transmitted from civil servants to the government." (a civil servant involved in the Economic Surveys)

[284] *Learning results rather from the work of the other OECD groups, whereas the reviews are there to exert pressure.* (a civil servant from a Ministry of the Environment)

[285] The review of the U.S. can be considered exceptionally consensual, in particular in view of the numerous potential themes of conflict.

internalisation of environmental policy norms by the governments, thereby participating in the battles between the epistemic communities. The reviews influence, therefore, rather policy selection and policy persistence than policy innovation. Instead of operating as a 'norm entrepreneur', stimulating policy innovation, the EPR programme helps consolidate the status of the already existing OECD policy doctrine in the member countries, by taking up such principles and assessing the degree to which countries actually implement them.

To the extent that the EPRs have a *credibility problem*, it stems partly from their perceived lack of independence, but also from their inability to contribute with new ideas. From this perspective, the OECD is a club of like-minded countries – a description in fact eagerly used by the organisation itself – rather than a genuine challenger and source of new ideas to its member countries. During the cold war, the OECD played an important role in helping Western countries to co-operate and open themselves up towards the rest of the world, providing a 'bridge to the West' for countries such as Finland, Japan and Portugal (Englannin kieli ei…2003).[286] Now, the cold war has ended, Portugal has become a democracy, the European Union has grown to become a major player in all fields of policy in its 25 member countries, and Japan has developed a host of contacts with the rest of the developed world. The *influence from other policy fields* (see section 4.4.3.2 above) therefore plays a key role, the new member countries from the former Soviet bloc representing a typical example. The interviewees suggested that especially in economic policy, the OECD peer reviews have had a great impact as a means for the governments on the one hand to gain a 'label of approval' for their policies helping them in the negotiations on the membership in the European Union, and on the other hand to justify to the citizens the need for the planned policy reforms.[287]

Finally, as the discussion on 'epistemic communities' and 'repertoires' of actors suggested, talking about the independence of the OECD without reference to the power relations among the actors involved is a very incomplete perspective. The EPRs are unavoidably an element in the power struggles between epistemic communities and the OECD can be seen to play a role of ideational

---

[286] The Portuguese interviewees explained the importance of external factors in general and the OECD in particular by the fact that during the military dictatorship (until 1974) the OECD was one of the few international organisations that accepted the country as its member. Likewise, the invitation to participate in the Conference of Stockholm in 1972 was an important victory for the country in terms of prestige. In Japan, *"[t]he OECD played the role of providing this bridge to outside influence. Japan is a relative newcomer among developed countries, and partly for this reason takes very seriously influences from the outside, trying to imitate its older 'peers'."* (a civil servant from the Ministry of the Environment)

[287] *"For the countries of Eastern Europe, the Economic Surveys are extremely important in guaranteeing the credibility of these countries in the eyes of the large financial institutions of the European Union. One proof of this importance was the participation of the Czech minister in the peer review meeting – it was the first time that a minister attended these meetings."*

agent or agency operating in favour of certain epistemic communities rather than particular member countries.

7.2.5 Change agents, repertoires, and pathways of influence

Let us now look at the potential change agents and their expectations concerning the EPRs so as to understand which of the four pathways of influence described above in section 7.1.7 (table 7.1) these groups might prefer. Table 7.3 summarises the repertoire-based requirements for credibility, salience and legitimacy for different actors, suggesting which groups are the most likely change agents for the different pathways of influence. It also shows the purpose that the 'difficult' stakeholders – those that seek to prevent the reviews from having an impact – attribute to the reviews.

Table 7.3. Requirements for credibility, legitimacy and salience of EPRs and the purpose of the reviews as perceived by different participants.

	MoE$_{ENV}$	MoE$_{STRAT}$	MoE$_{ECO}$	ECO$_{ENV}$	ECO$_{ECO}$	NGOs
Credibility	Narrow scope. Team with experience in environmental policy practice. Team independent in relation to OECD 'mainstream'.	Team with broad experience in environmental policy practice. Team independent in relation to operational level civil servants in all sectors.	Team with practical competence in environmental-economic policy. Team independent in relation to operational level environmental authorities.	Team independent in relation to environmental authorities. Team competent in sectoral policy issues.	Team independent in relation to environmental authorities. Limited participation of NGOs.	Team independent in relation to authorities. Sufficient NGO participation. Team competent in environmental matters. Critical tone of the reviews.
Legitimacy	Broad participation of floor-level environmental authorities. Care with cross-country comparisons. Tone of the review sufficiently balanced.	Broad participation of high-level authorities in all sectors. Respect country-specificities in cross-country comparisons. Tone of the review sufficiently balanced.	Tone of the review sufficiently balanced. Country-specificities taken into account in cross-country comparisons.	Tone of the review balanced, but critical. Country-specificities taken into account in cross-country comparisons.	Tone of the review sufficiently balanced. Country-specificities taken into account in cross-country comparisons.	Broad participation of NGOs. Clear cross-country comparisons.
Salience	Broad participation of non-environmental authorities in the review process. Focus on institutional issues and explaining policy outcomes.	Balanced focus on goal-achievement and institutions. Broad review scope. Sufficiently critical approach to help mobilise action.	Focus on cost-effectiveness, use of market-based instruments and explaining policy outcomes.	Focus on goal-achievement and integration of environmental concerns into sectoral policies.	Focus 'objectively' on priority environmental problems.	Focus on goal-achievement, and possibly explain the reasons for shortcomings in environmental performance.
Purpose of EPRs	The review process, but to some extent also the report, enhances learning about technical issues of environmental policy and visibility of environment in other sectors.	The reviews provide 'ammunition' for intra-governmental battles, offer a broad picture of where we stand in relation to other countries, and help in setting policy priorities.	Reports and the process provide legitimacy to a more market-based approach to environmental policy, both within the environmental administration and in other sectors.	Review process and reports enhance legitimacy of environmental issues in sectoral ministries and support for market-based approaches to environmental policy.	The reports can serve to provide an 'objective' picture of the real environmental problems, so as to help us defend our sector from attacks by civil society.	The reports from a respected expert organisation provide legitimacy to our demands for stronger environmental policy.

MoE$_{ENV}$ = Operational level civil servants at the Ministry of the Environment. (Mainstream environmentalist bureaucratic repertoire)

MoE$_{STRAT}$ = Strategic management level at the Ministry of the Environment. (Environmentalist political-bureaucratic repertoire)

MoE$_{ECO}$ = Economically oriented civil servants at the Ministry of the Environment. (OECD-minded environmental economist bureaucratic repertoire).

ECO$_{ENV}$ = Sectoral authorities responsible for environmental issues in their sector. (Environmentally-oriented economist bureaucratic repertoire)

ECO$_{ECO}$ = Sectoral authorities. (Mainstream sectoral bureaucratic repertoire)

NGOs = Environmental NGOs. (Environmentalist political activist repertoire)

The *operational level civil servants* in the environmental administration seem to prefer learning on technical issues of policy implementation, but may be equally interested in the accountability/advocacy type of influence, as long as it does not involve direct criticism against their own field of action.

While the 'floor-level' civil servants are seldom sufficiently well placed in power structures to take on the function of genuine change agents, the *strategic level managers* in the environmental administration are much more likely to do so. These actors expect the EPRs to exert political pressure on other sectors and help establish priorities in environmental policy.

A third group in the environmental administration consists of the *civil servants already favourable to the market- and results-oriented environmental policy approach* advocated in the EPRs. Often these individuals are not only sympathetic to the OECD message, but are also familiar with the OECD working methods, which helps them interpret and use the review recommendations to promote their own agenda. Like the strategic management level, this group sees the EPRs as a tool in the attempts to advocate an OECD-type policy style. Some civil servants emphasise a non-adversarial approach believing it is more likely to appeal to conflict-averse civil servants, while others consider it necessary to exert tougher, more political pressure to make civil servants 'move'.

The *persons responsible for environmental matters in the sectoral ministries* see the EPRs as a way of enhancing the legitimacy of these issues within their own sectors, thereby improving the chances for better policy integration. The civil servants responsible for environmental policies in the finance and economy ministries see the EPRs primarily as a means of 'selling' the message of 'economic logic' in environmental policy and as a way to legitimise their own work and status within their own organisation. Such a legitimacy comes on the one hand through 'soft coercion', but also through learning, since the reviews help sectoral authorities obtain a broad picture of the country's environmental policies. These individuals can operate as change agents within their own sectors, but only to a limited extent, since the environmental administration carries the main responsibility for the reviews, and because environmental issues still have only a marginal place in most sectors of administration.

The *'mainstream' sectoral authorities* focus on 'minimising the damage', i.e. making sure the possible criticism against their work (lack of integration of environmental concerns in the sector)

does not weaken their legitimacy. To the extent that they find in the report reason to highlight good environmental performance of their sector, they may use the findings to gather support, but in most cases they are likely to prefer the politically less risky learning function for the reviews.

Finally, the *environmental NGOs* seek in the EPR report criticism against the government and therefore support for their own arguments from a prestigious international organisation.

Figure 7.2 gives an illustration of the interplay and relationships between the repertoires, change agents and review design. In view of our discussion above on the epistemic communities, it was easy to distinguish the repertoires first in relation to the primary area of interests – environment or the economy – representing the horizontal axis in the figure. The other basic divide goes between those who prefer that the EPRs remain a mainly government-oriented affair and those who argue that the reviews have no significance unless they involve a broader range of participants in a public debate, where the media play a crucial role. This distinction is presented as the vertical axis. It is in the nature of the concept of 'repertoire' that individuals or groups of individuals do not necessarily fall within any single repertoire – any of the four quadrants in figure 7.2, but that they typically operate, depending on the context, within a range of different repertoires. This became clear also in the interviews, in which many interviewees attributed to the EPRs many different functions, without always clearly prioritising any of them.

Figure 7.2. Pathways of influence and actors' repertoires.
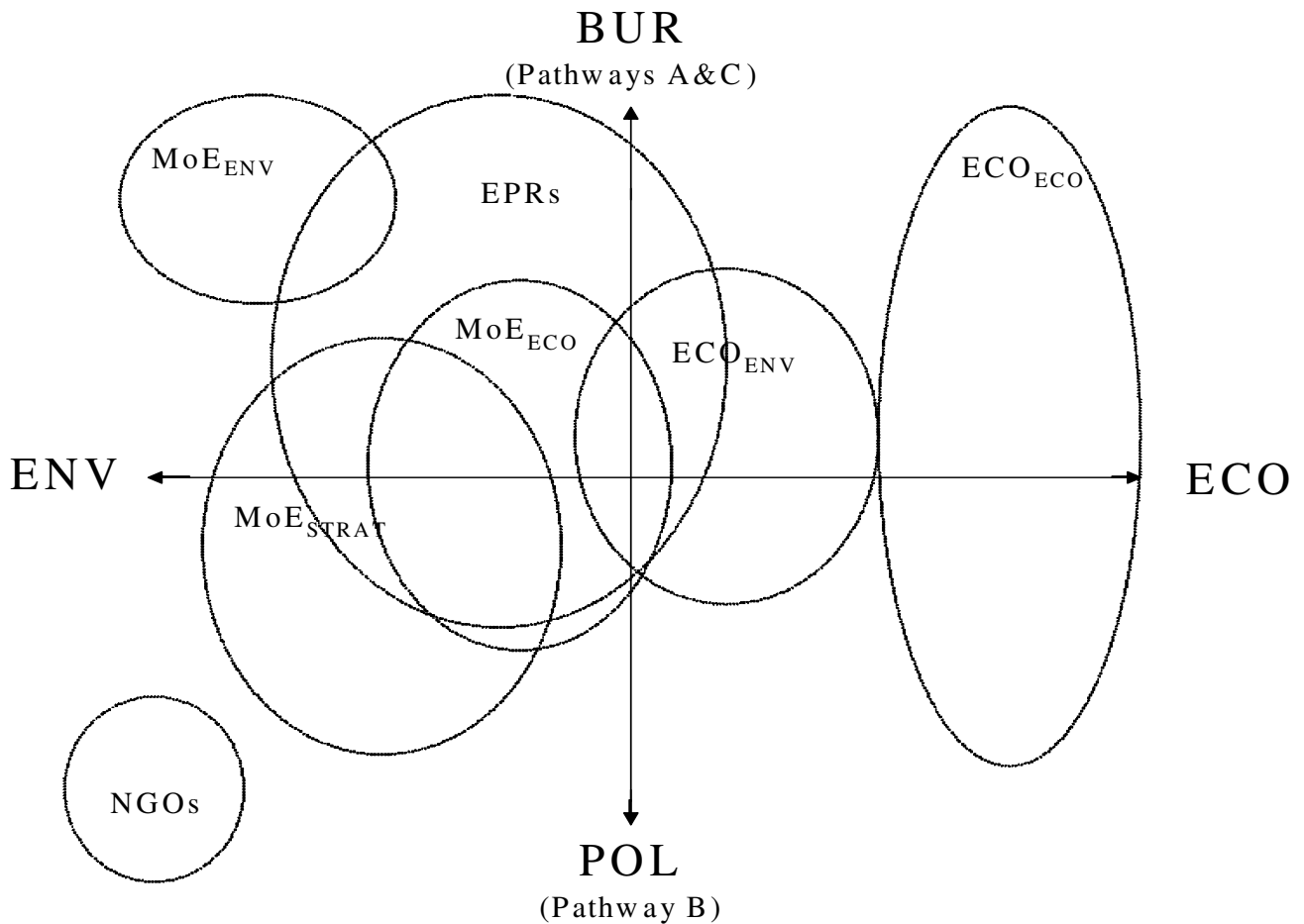
The pathways of influence A ('advocacy through intragovernmental debate') and C ('learning from best practice') fall clearly within the bureaucratic repertoires, whereas the pathway B ('advocacy/accountability through public debate') is in line with the 'political' repertoires. Pathway D is not represented in the figure, because it may operate, depending on the conditions and 'windows of opportunity', within any of the four quadrants. The exact positions of different stakeholders along the two axes obviously vary from one country to another.

The reviews remain to a large extent within the bureaucratic repertoire. Many civil servants, also in the environmental administration, prefer the politically less explosive, diplomatic type of pressure. One could say that the authorities opt for a 'risk-avoidance' strategy: given that the report may contain criticism against any of the governmental players involved in environmental policy, to stay on the safe side, it is better if the reports remain cordial and do not attract too much publicity.

A number of actors, especially the strategic management of the environment ministries, call for a more 'political' approach, implying a shift towards the lower left-hand quadrant in figure 7.2. Despite its attractiveness, this type of EPR use also frightens several participants, since it might engender resistance and lead to rejection of the reports, and would go against the OECD's role as a forum for dispassionate discussions even on 'hot' policy issues. The pathway of learning through dialogue (C) is preferred by three main groups of actors: those who are afraid of potential critique; those who do not see the OECD as a force politically strong enough; and those who have been frustrated by the excessively general nature of the recommendations and see them more or less useless, anyway. Pathway D is recognised by many as a key type of influence from the OECD work in general, but is not seen as a specific strength of the EPRs. Moreover, it cannot alone justify the existence of the reviews, since their fundamental *'raison d'être'* lies in the more immediate impacts.

The review influence can be expected to be greatest in situations where most key actors are located in the same quadrant, in particular where the strategic level managers and the economically-oriented actors in the environmental administration prefer the bureaucratic, rather than the political repertoire, where the civil servants responsible for environmental issues in sector ministries are sufficiently motivated to advocate for the use of the EPRs and are in strong enough a position in their ministries, and where the 'mainstream' of environmental administration is relatively open to the OECD type policy approach.

7.2.6 Why are the ERPs what they are?

The basic approach of the EPR programme has not significantly changed over the years, despite numerous reform proposals by the WPEP delegates. Clues as to why this should be so can be sought in the principal-agent analysis. Despite the common objective of improving environmental policy, shared by the Environment Directorate and the environment ministries of the member countries, some informants suggested that the secretariat has, in fact, skilfully exploited the disagreements between the WPEP delegates in order to maximise its own margin of manoeuvre.[288] The secretariat would have used the lack of consensus among the member countries on the one hand, and its information advantage, on the other, in order to resist changes in the programme, thereby pursuing interests of its own, rather than those of the organisation as a whole. It is impossible to say to what extent such 'shirking' might have exceeded the limits of what can be

considered as legitimate defence of one's own interests. Moreover, rather than being an egoistic attempt to only look after one's own interests, the resistance may simply derive from a strong conviction about what the secretariat believes to be the 'best' way of maximising the impact of the programme. What is clear, however, is that the absence or slowness of progress in modifying the programme has frustrated some of the participants, and reduced their motivation to contribute to its further development. Some participants have pointed out that there was more enthusiasm among the WPEP members during the early years of the EPR programme, which at the time provided plenty of opportunities to learn both about the review process and the substance matter, whereas the lack of innovation in the EPRs has lead to disillusionment as the programme is seen to produce nothing but just 'more of the same'.

While the consensus rule in decision-making can be blamed for a lot of the stagnation in OECD work, it has at least one significant advantage for peer reviews. Since all member countries must agree upon the recommendations, the reviewed country cannot simply snap its fingers at the conclusions and recommendations saying they only represent the views of the OECD secretariat or some member countries.[289]

On the other hand, the internal dynamics of the WPEP group certainly play a role in preventing significant changes in the programme, as exemplified in the comment by a person replacing her colleague in one WPEP meeting:

"I had the impression that WPEP group is to a large extent made up of people participating in the process since a long time already, and therefore having a feeling they produce good quality work, but not getting the recognition they consider they deserve. Also, the group seems to be constantly on the defensive, fearing future budget cuts, and focusing on convincing others of the quality of their job, instead of getting down to business and doing something to improve the quality of their work."

---

[288] This observation attests to the existence of 'shirking' or 'bureaucratic slippage' (Verbeek 2001, 7).

[289] A civil servant involved in the Economic Surveys evoked the advantages of consensus in the following comparison with the IMF Article IV reviews: *"Countries can ignore the IMF reports simply arguing they only reflect the opinion of the IMF secretariat. In the case of the OECD EDRC reports they cannot do so because they explicitly agree due to the consensus rule. Secondly, while the consensus requirement may sometimes water-down the recommendations, it may work in the other direction as well […]. Thirdly, and most importantly, the consensus rule prevents countries from 'hiding' behaviour: when the report is published they cannot say that they don't agree and ignore the recommendations."*

7.3 Country-specific institutional context[290]

So far we have dealt with the EPRs in abstraction from the country-specific political context in which the reviews operate. However, all along the discussions concerning the future of the reviews, it has been clear that the OECD faces tremendous challenges in dealing with the great variations between its member countries – with respect to their economic, political, geographic, ecological, cultural and social conditions, as well as when it comes to their expectations concerning the EPRs. In the following, we shall see how the contextual, country-specific factors condition the EPRs' influence. The factors that stand out the most have to do with the national policy style, socio-economic conditions, changes in the politico-administrative structures, and the position of the potential change agents in the national policymaking structures.

7.3.1 National policy style and policymaking structures[291]

The interviews suggest that the attitudes in the country concerning the opinions from beyond the country's borders significantly affect the capacity of the reviews to have an impact. The cases of Japan and the Netherlands represent two extremes: Japan has a long-standing tradition to pay considerable attention to foreign opinions – the country has, indeed, throughout the history, absorbed outside (largely Western) influences, but adapted them to its own needs and conditions. The OECD was, in the 1970s and early 1980s, one of the few arenas where Japan could actively communicate with other industrialised countries. The Dutch society, by contrast, seems more self-sufficient with respect to ideas, and more indifferent with regard to outside opinions.[292] In Japan, the influence of the EPRs was probably strengthened by the policymaking style that involves continuous inter-ministerial negotiations, in which the OECD recommendations, among others, are used as means of persuading other authorities. In Portugal, in turn, outside opinions have usually carried considerable weight – indeed, the country's environmental policy agenda is seen by some observers as excessively determined by outside forces, such as the EU – but a hypothesis can be put forward that the country's political culture allows a much more flexible and opportunistic use of the outside views than is the case in countries with a less legalistic, and more implementation-oriented governing style. Chronic problems of poor policy implementation and the ability of the country, by

---

[290] See section 4.4.3.
[291] See section 4.4.3.1.
[292] While this assumption was contested by some Dutch interviewees, there are signs that a certain 'national arrogance' may play a role, as shown by a comment by a Dutch environmental policy researcher: *"I am afraid it is just a matter of the old 'belief' that we are doing better than most others and don't need advice from outside."*

virtue of its long history, to maintain social cohesion even in the absence of strong formal institutions, were evoked as typical characteristics of the Portuguese society (Soromenho-Marques 1997; 1999; 2001). In such a context, since formal institutional rules are easier to change than the informal ones (e.g. Hukkinen 1999), the EPRs face even greater challenge in trying to make a difference.

Another factor shaping the influence of the reviews is *the degree of cooperation and networking* in the country's policymaking institutions. The Dutch, the Japanese, and many of the Portuguese authorities explained that the review process did little to improve networking, because the networks already existed, and were fully exploited. Only for newcomers in the policy area did the EPR process provide an opportunity to learn to know other actors in the area and develop professional networks. In Mexico, such networking function was seemingly much more important, especially during the first review. The newly established environment ministry saw in the EPR an opportunity to enhance cooperation and coordination among the previously disparate agencies now gathered under the auspices of one single organisation. As for the impacts from the review *report*, the question is more complex, because highly integrated policymaking structures also mean that the country has greater capacity to diffuse the review findings both within and beyond the administration.

A crucial question is whether the country's *policy style* is *adversarial or consensual.* In the former, one might suspect the report to be used through pathway B ('advocacy/accountability through public debate'), since the policymaking process in the country is characterised by criticism and conflict. However, the interviews point towards a different interpretation as well: the diplomatic tone of the reports is particularly appreciated, because it provides a rare opportunity for dispassionate discussions in an otherwise conflict-laden environment. Conversely, actors in a country with a consensual policy style may precisely seek harsher critique and new perspectives from beyond the country's borders in order to compensate for the lack of confrontation in the country's policy debates. A parallel can be drawn to the *character of the issues* addressed in the review: highly controversial issues tend to attract a lot of attention, and therefore have the potential of generating impacts, but they may also have the adverse impact of stealing attention from the broader picture (see section 7.1.3 above). To some extent such a focus on controversial topics might be useful for the advocacy and accountability functions, while compromising the potential to generate learning.

The EPR process seems to be more influential in new OECD countries or in countries whose environmental policy institutions are still in a formative phase, while the most advanced member countries find it harder to see what value-added the review could bring. The cause-effect relationships involved are, nevertheless, highly complex, because the 'laggard' countries usually also lack the resources and even the know-how needed to fully exploit the potential of the EPRs. Moreover, the use of the report for advocacy/accountability is more difficult in a highly advanced country, not least because the review team may have a hard time coming up with strong enough criticism in a 'frontrunner' country. What is clear from the discussions in the WPEP, by contrast, are the differences of priorities between countries in function of the level of development of their environmental institutions: the 'frontrunners' are most interested in the issues related to sustainable development and integration of environmental concerns into other sectors (Part II of the reviews), whereas the less developed countries stress the importance of the 'traditional' environmental issues (Part I of the reviews: air, water, nature, waste).

Differences in the *discussion cultures* between 'Latin' and Anglo-Saxon countries may explain why the overall diplomatic and cautious tone of the EPRs seems to please some countries, while being strongly criticised by others. A Portuguese interviewee suggested that the 'Latin' countries are likely to be more accustomed to interpreting diplomatic, hidden messages, whereas the Anglo-Saxons would require more direct language. The desire of the Dutch delegates both in the WPEP and the EDRC group to drive reforms allowing more cross-country comparison and more rigorously formulated recommendations in the environmental and economic reviews lends some support to this argument. Anglo-Saxon countries were in general pushing for more radical changes to the review framework during the preparations for the second cycle of the EPRs.[293] On the other hand, these Anglo-Saxon countries seem to take reviews more seriously, whereas the political culture in the 'Latin' countries is probably more permissive in relation to lack of policy implementation. The case of the Netherlands – where the government's response that it had already taken or was about to take the measures recommended in the EPR – suggests that, contrary to the argument of Gambetta (1998), the reaction of the type "Obvious! I knew it all along!" may not be confined to the 'Latin' cultures with 'indexical' beliefs about knowledge. The fact that the 'Latin' countries tend to be more satisfied with the diplomatic tone of the reviews can be explained by the differences in beliefs about knowledge: relatively vague formulations allow the countries with such 'indexical' beliefs to interpret more freely the review results, and thus convince their populations, whereas the governments in the 'analytical' countries believe citizens would prefer more direct language and

arguments amenable to empirical testing (see Gambetta 1998, 29). Again, since the 'laggard-frontrunner' divide in environmental policies largely coincides with the 'Latin'/Anglo-Saxon division, disentangling the respective impact of different factors is not easy.

The degree of policy integration and institutionalisation of policies (see section 4.4.3.1 above) partly seems to dictate the facility with which review findings can be absorbed by the politico-administrative system. The lacking institutional structure was felt as one impediment preventing the chapters on the environmental-social interface from prompting reactions in the reviewed countries.[294] Since none of the ministries or government agencies is clearly responsible for issues at the interface, often no government body takes the responsibility for the recommendations either.[295] In other areas, such as agriculture or transport, the OECD countries have already at least a basic preparedness to take into account environmental issues, and the reviews have been perceived as useful precisely in promoting policy integration and improving the status of environmental issues in sectoral policies.

Weiss (1999) has assumed the existence of a *tradition of policy evaluation* in the country to enhance the chances of evaluations being used. However, in the case of the EPRs the impact seems to be rather the opposite, i.e. it is precisely the existence of national evaluation institutions, especially in the larger countries, that reduces the interest in the EPRs, which are not perceived to contribute anything new to the national evaluation systems.[296] Moreover, the multiplication of different

---

[293] Among the most active countries were Australia, Canada, Finland, and the Netherlands.

[294] The majority of persons involved in the reviews have considered the chapter on environmental-social interface as important, notably because the social 'pillar' of sustainable development is seen as the least developed one, and because, it has been argued, no other OECD body than the WPEP has the interest in working on the issue. However, the way the environmental-social interface has been dealt with and the attention the chapter has received in the reviewed countries, were considered by the interviewees as disappointing, and new ways are being sought to keep the issue alive (e.g. OECD 2005a). Comments on the environmental-social interface chapter included the following:
"Normally, the social chapter is not really 'social', but just a compilation of facts that previously were in the economy chapter." (a WPEP delegate)
"The social dimension is important, but the way it is treated in the EPRs is not very useful. Mixing procedural and substance issues together does not produce a good result: issues relating to democracy should not be dealt with under the theme 'social' and alongside with distributive and other social matters. What would be useful is to look at the situations where the environmental and the economic interests clash, and what is the proportion of cases in which either one of them wins." (a WPEP delegate)
[295] *"The problem with these cross-sectoral issues is that they are easily forgotten, since none of the sectors of administration feel responsible for keeping the issue alive and implementing the recommendations. Clear environmental issues are simpler: it is there we can use the recommendations, because we have the power and responsibility for implementation."* (a civil servant from a Ministry of the Environment)
"The social chapter did not have a lot of visibility. The trade unions should take more responsibility for social and environmental questions. However, in [our country] the trade unions are still very conservative." (a civil servant from a Ministry of the Environment)
[296] *"[b]ig countries usually have their own 'think tanks' and institutions producing similar assessments, and therefore the OECD does not provide them too many new insights. For example in the U.S., the EDRC surveys rarely get any*

evaluations and reporting obligations was sometimes mentioned as a source of the 'evaluation fatigue' referred to above in section 6.3.6. By contrast, the EPRs may help to introduce an evaluation culture' to countries where such a tradition is absent.

7.3.2 Socio-economic conditions: economic conjuncture and the status of the environment on the national policy agenda[297]

Economic downturn is generally believed to compromise the potential of the EPRs to have an influence, since in such a context it is difficult to argue for additional investments in environmental policies. Since most OECD countries faced economic difficulties and budget cuts over the past few years, the comments referring to the unfavourable economic context, especially in Portugal and the Netherlands, were not surprising.[298] However, OECD recommendations could actually be more useful in times of economic downturn, thanks to their emphasis on cost-effectiveness.[299] The improvement of environmental conditions in most OECD countries was brought up by one interviewee as a major reason for the lesser public interest in the EPRs.[300]

The general status of the environment in the public consciousness was also evoked as a factor affecting review influence. A Norwegian WPEP delegate argued, for instance, that the EPR was relatively influential, partly because environmental issues always tend to attract public interest in the country.[301]

---

publicity in the media." (a civil servant involved in the Economic Surveys) The perception that the OECD is very poorly known in the U.S. and therefore the reviews may not easily reach the media agenda was confirmed in informal discussions with a number of U.S. delegates during the country's EPR in May 2005.

[297] See section 4.4.3.2.

[298] *"The environment is not a very popular subject today, and the issues dealt with in the reviews were quickly pushed off the agenda of public debate by other issues, such as those concerning safety, economic problems, etc."* (a civil servant from a Ministry of the Environment)
"In the present situation of budget cuts (especially in the Ministry of the Environment) it is difficult to implement any new policies or instruments." (a civil servant from a Ministry of Finance)

[299] *"Issues that were high on the policy agenda at the time of the review were the tightening of the budgets and the need to improve the cost-effectiveness of environmental policies (e.g. by removing environmentally harmful subsidies). The OECD revived this discussion, yet without really contributing to it with new ideas."* (a Dutch civil servant)

[300] *"Probably the reason for the environmental issues being lower than earlier on the policy agenda is the dramatic improvement in the quality of the environment in the Netherlands over the past few decades, and the resurgence of social and economic problems that have taken over the priority."* (a Dutch interviewee)

[301] Personal communication, Gunnar Farestveit, 19 May 2005.

7.3.3 Changes in the governing coalitions and administrative structures[302]

Frequent changes in the government structures during the review are likely to reduce the influence of the reviews, as a result of a loss of 'institutional memory' and lack of continuity if the persons responsible for the review are removed to other tasks. In Portugal, changes of government, top management of the environment ministry, and the position of the WPEP delegate created instability that was considered harmful for the impact of the review. Likewise, a new minister often needs to be 'trained' before he understands the value of external critique coming from the OECD or from the NGOs as an asset in the ministry's struggle for greater influence, rather than as a threat to its legitimacy.[303] On the other hand, a change of government or environment minister at a suitable moment may open up new 'windows of opportunity'. This was mentioned as one of the key reasons for impact in Norway and Mexico, where the new environment ministers were eager to use the review as a benchmark against which the new policies could be assessed, and a source of legitimacy for the new government's environmental policy objectives, but perhaps also to prove the superiority of their own government in comparison to the previous ones. Often, attempts are made to use the report in connection with parliamentary elections or the formation of a new government. Experience from attempts to exploit the EPR in connection with the election campaign have been less successful, probably because of the modest weight of environmental issues on the general political agenda.

It is difficult to generalise on the importance of the political orientation of the government on the left-right dimension. In Portugal, the interviewees argued that the left-wing governments had clearly conducted more 'environmentally-friendly' policies and thus were more likely to take more interest in the EPRs, whereas in the Netherlands such a distinction was not seen as relevant.[304]

7.3.4 Conclusions on the country-specific institutional context

The country-specific factors pose further difficulties to the EPRs, given the heterogeneity of the OECD member countries. A large country, with well-developed political and administrative structures in the field of the environment, a broad range of research institutes and 'think tanks'

---

[302] Section 4.4.3.2.
[303] *"The ministers, when they take office, do not understand the importance of the NGOs as 'allies' of the Ministry of the Environment, and therefore want to cut the financial support granted to these organisations. Usually, after a few months the attitudes change."* (a civil servant from a Ministry of the Environment)

carrying out policy assessments and evaluations and with an open political discussion culture would have to be reviewed through a substantially different approach than a small 'laggard' country, with a more closed discussion culture, and where stakeholders play a limited role in policymaking. Changes in government just before or during the review process have sometimes been favourable to review influence, while stability in the administrative structures would be essential to guarantee sufficient continuity in the process and make sure that the 'change agents' do not disappear in the middle of the process.

The material available for this study did not allow us to draw generalisable conclusions concerning the importance of the country-specific institutional context on the influence of the reviews. Moreover, far-reaching generalisations across countries would run against the basic assumptions underlying this study, notably the need to take the context seriously. However, a number of tentative hypotheses can be put forward for eventual future testing. Table 7.4 presents the key aspects concerning the country-specific institutional context dealt with in this chapter, and provides tentative suggestions as to their impact on the various potential pathways of influence.

Regardless of the pathway, the reviews are likely to have the highest impact in small countries that are sensitive to outside opinions, have recently joined the OECD, have a relatively modest level of development of environmental policy institutions and a weak tradition of environmental policy evaluation, in a situation with the maximum possible stability in the administrative structures. A more direct, 'Anglo-Saxon' discussion culture favours the 'political' pathway B, and 'learning from best practices' (C), whereas pathway A, 'advocacy through intragovernmental debate', tends to be too 'soft' and 'diplomatic' to make a difference. Economic downturn may either increase or reduce review influence, since it may on the one hand increase the interest of governments in cost-effectiveness arguments typical for the OECD, but on the other, makes more unlikely any additional investments in environmental policy.

Beyond these general observations, the contextual factors affect differently the capacity of the reviews to influence policies through the different pathways. Pathway A, 'advocacy through intragovernmental debate', is most likely to be used by the government in a country with a 'Latin' discussion culture, in a situation where a new government has recently taken office. By contrast, the more 'political' pathway B, 'advocacy/accountability through public debate', operates best in a

---

[304] *"The government was right-wing, but this was not decisive: at times, right-wing governments have in the past been very progressive in their environmental policies."* (a Dutch interviewee)

country where environmental issues are high on the policy and media agenda, and a new government has recently taken office, wishing to exploit the review to provide a platform for its policies. Pathway C, 'learning from best practice', is likely to work best in a country with a consensual policy style, little networking across sectors of administration, and a high level of stability within the administration during and after the review process. In contrast with what was said above about the 'laggard' countries being more likely to benefit from the reviews in general, however, at least the Dutch experience suggests that it is precisely the 'frontrunners' who might prefer the learning pathway, since they do not see the EPRs and the OECD as sources of political pressure.

Finally, the significance of the different contextual factors to the pathway D, 'long-term conceptual learning', is the most difficult to estimate. It seems plausible to assume that also this pathway would work better in a country sensitive to outside opinions, but it is impossible, within the present study, to provide more in-depth analysis into the significance of different factors for long term conceptual learning.

Table 7.4. The impact of contextual factors on the likelihood of the reviews to bring about effects through the three first of the pathways.

		A	B	C
Attitudes to outside opinions	Sensitive	+	+	+
	Autarchic	-	-	-
Degree of networking in government	High			-
	Low			+
Political style	Adversarial	+/-	+/-	
	Consensual	+/-	+/-	+
Discussion culture	Anglo-Saxon	-	+	+(?)
	Latin	+/-	+/-	-
Size of the country	Large	-	-	
	Small	+(?)	+	+(?)
Degree of environmental policy development	Frontrunner	-	-	-/+
	Laggard	+	+	+
'Evaluation culture' & routine	Strong	-	-	-/+
	Weak	+	+	+
Status of environmental issues on public policy agenda	High	+	+	+
	Low	-	-	-
Changes in politico-administrative structures	New government	+	+	+
	Administrative changes during the process	-	-	-
Economic downturn		+/-	+/-	+/-

\+ : enhances potential impact
\- : lessens potential impact

7.4 Interplay between review design, actor relationships, and country-specific institutional context: illustration through examples from France, the Netherlands, and Portugal

To illustrate the way in which the different factors interact and to explain differences in review influence in different country-specific situations, we can now take a closer look at the experiences from the EPRs in France, the Netherlands, and Portugal – the countries in which interviews were carried out among a broader range of participants than just the WPEP delegates and other 'insiders'.

In **France**, the peer review meeting took place recently (January 2005), and only the conclusions and recommendations had been released by the time of the interviews for this study were conducted, which limits the lessons that can be drawn on the review influence. However, the interviews and discussions in the WPEP indicated that, overall, the representatives of the environmental administration appreciated the review process and expected the review report to be useful in their future work. The process was seen both as an opportunity to reflect upon the country's policy priorities and approaches in the light of the novel perspectives from outside experts coming from a different political and cultural environment, and as an opportunity to enhance the political visibility of environmental issues. While some non-environmental representatives had a reserved attitude towards the exercise, and some representatives of sector ministries criticised the review process for not providing enough opportunities to get their message across, others pointed out the value of the process as an opportunity to interact with environment ministry officials they normally would not have interacted with, and as an occasion to learn about the country's environmental policies. The Ministry of Ecology and Sustainable Development expects the recommendations to be used 'politically' and wishes to mobilise the civil society to exert pressure on the government. The interviewees therefore expected the review to operate mainly through pathway B,[305] but also to help create pressure and enhance legitimacy of environmental policies and authorities in cross-sectoral work (pathway A). Learning (pathway C) was evoked by some interviewees, but only as a secondary source of review influence.

The key explanations for the largely positive picture in France can summarised as follows:

---

[305] Of course, this expectation may prove to be based on wishful thinking, based on an unrealistic understanding of the potential of the review to have an influence.

- France is a large and influential country in the OECD, and has traditionally had a relatively 'autarchic' policy style – factors that tend to reduce potential review influence and engender defensive attitudes – yet there are signs that the policy style is changing, and that OECD work is taken rather seriously in the country.[306]

- The country's discussion culture can be described as 'Latin', which increases the likelihood of the rather diplomatic EPRs being received positively by the government.

- While not a 'laggard', neither is France a 'frontrunner' in environmental policies, and therefore may be more open to outside influences and ideas than the more advanced countries.

- Evaluation of public policies is widespread in France, but less so in the field of environmental policies. In particular, the country does not have a large number of independent 'think tanks' producing evaluations and assessments. Therefore, the EPRs have an added value, and there is perhaps less 'evaluation fatigue' than in some other countries.

- The environment ministry, in its present structure, is relatively young and has a modest status within the government. The EPR provided it with much-needed political support from an international organisation and an opportunity to consolidate its position.

- During the review process, there were no significant changes in the administration and in the organisation in charge of the review during the process, which enhanced the potential of the reviews to produce an impact.[307] However, the change of the Minister of the Environment in early June 2005 may affect the way in which the EPR will be used.

- Top level support: The minister himself headed the delegation in the peer review meeting, and wanted, for instance, to publish the conclusions and recommendations of the review already three weeks after the peer review meeting.

- Change agents: there was, in the Ministry of Ecology and Sustainable Development, a group of persons knowledgeable of and sympathetic to OECD work, as well as capable and willing to make advocacy for the review.

---

[306] Defensive attitudes were seen in the difficult discussions during the peer review meeting on the wordings concerning nuclear energy. The changing policy style, in turn, was referred to by one interviewee as follows: *"France has evolved a lot; some ten years ago, a minister wanted our country to leave the OECD, because he argued that foreigners should not judge French policies. Today, the ministers are more open to criticism and France is becoming used to external evaluations. The times of 'sovereignty' are therefore already behind us."*

[307] One interviewee also considered that it was easier to deal with the present right wing government, because it had a more coherent political platform than its multi-party, leftwing predecessor.

- The OECD headquarters being located in Paris, there was probably a livelier dialogue between the secretariat and the French counterparts, which improved the quality of the review.

In *the Netherlands*, the picture was rather different from that of France, since all interviewed persons considered the review had very modest influence, if any. The use of the recommendations was limited to the Secretary of State for the Environment referring to the report in a number of speeches during the weeks following the release of the review, and sporadic reference made to the report by environment ministry officials trying to convince their colleagues in the sector ministries to adopt measures recommended by the OECD. The NGOs hardly used the review, and even environmental policy researchers did not seem to be aware of its existence.[308] The Secretary of State also sent a communication to Parliament explaining what the government intended to do to implement the recommendations, with essentially one message: the government is already doing, or is about to do, everything that the OECD recommends. The review did revive some policy initiatives that had already been rejected during the mandate period of the previous government (e.g. per-km tax, the elimination of certain environmentally harmful subsidies), but the interviewees were pessimistic about the prospects of these initiatives being adopted in the new situation. Finally, many civil servants involved in the process perceived it as a waste of time, since adequate networks of cooperation across sectors were seen to exist already, and because they did not see the process as dialogical enough to promote genuine learning. The government did not envisage drafting a follow-up report, because of the rather disappointing experience from publishing such a report after the first cycle review. The little influence that there was, primarily operated through pathway A, through the impact that the EPR recommendations may have had in the intra-governmental debates and negotiations.

The factors explaining the predominantly disappointing experience in the Netherlands include:

- Even though the Netherlands is a small country, it can be considered to be rather 'autarchic' at least with respect to environmental policies. The OECD is not considered as a 'heavy weight' player in international environmental policies; therefore, the coercive effect it could have on the country's policies was expected to be minimal.

---

[308] *"I have hardly ever heard anything substantial about [the Dutch EPR's] impact here […] It indeed seems as if the review was considered irrelevant and (largely) ignored."* (Duncan Liefferink, personal communication, 4 March 2005)

- Not only is the Dutch policy style highly participatory, and the limited degree of stakeholder participation in the review process compromised the legitimacy of the review, but the highly networked policymaking structures reduced the potential of the review process to have an impact.

- The political and discussion culture in the Netherlands require a more direct straight-to-the-point approach than the very diplomatic OECD reviews can offer; the 'ceremonial' aspects of policymaking do not have the same importance as in some other countries.

- The country is a traditional 'frontrunner' in environmental policy, with its highly advanced research institutes, 'think tanks' and active NGOs; the OECD would not be needed to exert political pressure on the government.

- The status of environmental issues on public agenda was in a strong decline, the country was undergoing an economic downturn, and the new government wanted to drive a 'minimalist' line in its international environmental policy, doing only what was necessary to fulfil the country's international commitments and abandoning the ambition to continue as a 'frontrunner'; therefore, the OECD work in general had a low priority.

- Lack of enthusiasm and motivation to work for dissemination, advocacy and follow-up: most persons involved in the process expected the review would be more or less a repetition of the country's first EPR, and would therefore not bring about any significant impacts.

- The environmental administration considered the quality of the review report as mediocre: the recommendations were too general and did not suggest *how* the country should go about in its environmental policies; the report did not contain enough cross-country comparisons.

- Since the government had significantly lowered the level of ambition of environmental policies over the previous years, it was not eager to be criticised in an international assessment.

- The Dutch were frustrated as a result of their failure to drive through the reforms in the EPR framework they had forcefully advocated in the course of the preparations for the second cycle of reviews.

Finally, the experience of **Portugal** lies probably somewhere in between that of France and the Netherlands. Overall, the influence of the review could be deemed as relatively modest. The process was considered as useful in so far as it forced the civil servants to stop and reflect upon their policy priorities and performance, and provided an opportunity for networking for those that had recently taken up their duties. The minister used the review report in a couple of occasions to justify his

policy choices and to make publicity for the government's allegedly excellent record in nature conservation. Otherwise, the review process was deemed more important than the report as a source of influence. Therefore, in addition to the sporadic use of the EPR to give legitimacy to environmental policies through public debate – notably by the environment minister (pathway B) – and to 'mainstream' environment in sectoral policies through legitimisation (A), the review process in some cases enhanced network building (C).

Factors affecting the influence of the Portuguese review can be summarised as follows:

- External influences have traditionally had a strong influence on policies in Portugal. The OECD still has a high prestige in the country, which makes it a potentially useful source of justification for policy arguments.
- Environmental policymaking in Portugal is relatively well networked, and most actors already know each other. This limited the networking effects from the EPRs to the 'newcomers' in the administration.
- The country's political style is consensual and elitist, and its discussion culture 'Latin', enhancing the potential impact of a cordial, administration-driven evaluation.
- Portugal is a small country, where the review is less likely to generate defensive attitudes than in the larger ones.
- Portugal can be considered as a 'laggard' rather than a 'frontrunner' in environmental policies; therefore, critique does not have to be very harsh to have an impact.
- The tradition of policy evaluation is weak, there are few if any 'think tanks', and therefore impulses for policy changes tend to come from abroad.
- The numerous changes of personnel responsible for the review during the review process hampered the continuity of efforts.
- Dissemination of the review was clearly deficient (no public seminar, no broad distribution of the report; no follow-up report), largely because the ministry's top management did not sufficiently support the process and understand the value and the character of OECD work.
- The Minister of the Environment used the report with tactical skill, yet did not have the strategic eye for a more long-term utilisation.

7.5 Conclusion: Key factors conditioning EPR influence

The numerous factors that condition the degree and the type of influence of the EPRs were above divided into those concerning the way in which the reviews are conducted (review design), the relationships between the actors each having their repertoires, worldviews and expectations, and the country-specific framework conditions within which the reviews are being carried out. The four pathways of review influence differ with respect to the purpose of the reviews – learning or accountability – both of which can be further divided into two types. The repertoires of the key stakeholders, in turn, largely determine the preferred choices between the different pathways, and therefore the desired type of review design.

The relationships between the different parties involved have a crucial importance for the influence of the reviews. While opinions diverge greatly as to the relative importance of the objectives of reviews – learning and accountability – there is a consensus among the persons involved in the reviews at least on the need to safeguard the character of the OECD reviews as a combination of carrot and stick – policy dialogue and soft persuasion. This is seen as the best way to ensure success of the EPRs in straddling between the two 'epistemic communities' – the 'economist' and the 'environmentalist' one – and serving as a tool for the latter to strengthen its position in relation to its more powerful 'economist' counterpart. The challenge for the EPRs is, therefore, to locate stakeholders within the range of repertoires, identify their positions in the relations of power, and try to link with the dominant repertoires while remaining sufficiently close to the OECD mainstream, in order not to compromise their credibility in the internal struggle for power within the organisation. The question is above all that of symbolic, ideational power – practically the only power resource available to the OECD when trying to influence policies.

The different 'repertoires' of the actors involved in the programme reveal a more complex picture, the two 'epistemic communities' being cut into different fractions along the dimensions of 'bureaucratic' vs. 'political', and 'OECD-minded' vs. 'traditional' repertoires. The EPRs not only serve as a tool for the 'environmentalists' to strengthen their position, but also as a vehicle for the OECD to advocate its view of good environmental policy within the environmental administrations. To serve their double function of environmental policy advocacy and the promotion of the OECD policy doctrine, the EPRs need to motivate the potential 'change agents' by linking with the dominant repertoires in the country, yet without losing their integrity and credibility among the

'environmentalist' community. To some extent, the EPRs have failed in this task, partly because of the inability of the programme to innovate and renew itself, but also because the reviews at present do not seem to optimally serve any stakeholder group. Moreover, through the EPRs, the organisation does not clearly play any of the four roles of ideational artist, arbitrator, agent, or agency. In principle, the OECD operates, through the EPRs, as a norm and identity creator, generating peer pressure that operates through legitimacy, esteem and conformity of government stakeholders, and leads to the internalisation of norms. In practice, the incapacity of the OECD to function as a norm entrepreneur, in the service of policy innovation, tends to reduce the enthusiasm and motivation among the potential change agents – especially those who have participated in the reviews long enough to see their potential and limits. Socialisation, in turn, is reduced by the failure of the EPRs to have significant coercive impact, because of the highly diplomatic tone of the reviews and the low political status of the OECD in the member countries, and because of insufficient policy dialogue.

Finally, the EPRs must seek compromises, in order to accommodate the different country-specific contexts. Not all of the pathways are equally feasible in all contexts, hence the need to be flexible and attentive to the wishes of each country. However, to safeguard sufficient comparability and continuity, the reviews cannot be too flexible either. And, finally, listening to the desires of governments always entails the risk of falling victim to opportunistic political power-play, in which the OECD cannot but lose legitimacy.

This chapter, together with the previous one, has shown that the EPRs do influence policies, even if often not to the extent that the parties involved would like them to have. We have likewise seen that the influence of the reviews as well as the expectations concerning the reviews tend to vary according to the relationships between the actors; that the role of the OECD as an independent, external evaluator is an indispensable asset for the reviews' influence; and that the type and the degree of review influence varies significantly from one country to another. However, the discussion has so far focused on the relatively direct impacts of the reviews, through the three first pathways identified above in figure 7.1. Nevertheless, the fourth pathway of influence – the one that operates through the long-term effects on 'mental models' and frameworks of thought – may often be the most significant. Such an assumption relies not only on the lessons from the evaluation research (section 4.3) stressing the importance of the indirect types of influence, such as 'enlightenment', but also on the material for this study. Even the interviewees that had a positive attitude towards the EPRs and believed in their usefulness were often incapable to cite concrete

examples of evaluation use. In addition, a number of interviewees were simply disappointed by what they considered as reviews without any significant impact. Yet others pointed out that OECD work in general tends to influence policies, not so much directly, but rather through indirect, long-term effects on frameworks of thought. Analysing such impacts from our 'institutionalist' perspective to sustainable development will be the task in the following chapter.

Chapter 8

Should the EPRs have influence? The review programme from the perspective of institutional economics

This chapter applies the criteria developed in chapters 2-4 and summarised in the end of section 4.2.4 for an 'institutionalist' evaluation approach in order to analyse to what extent the OECD Environmental Performance Reviews conform to such criteria. Moreover, the typology of the types of democracy, civil society and impacts of evaluation on democracy presented in section 4.3.7 is employed for analysing the EPRs' capacity to enhance deliberative democracy.

Chapter 6 concluded that the EPR programme does have influence, in particular by legitimising the environmental policies and actors in the reviewed countries, and by helping to 'infiltrate' OECD ideas into national environmental administrations. Yet, often the direct impacts are hardly visible, and the main effects may be of conceptual character. This leads us to the question concerning the nature and the desirability of such an influence: what is the substance of the learning assumed to take place through indirect mechanisms of 'enlightenment' and long-term learning? What kind of message do the EPRs transmit and what kind of 'mental models' do they promote? Is the OECD message in line with what in this study is understood by sustainable development? Is the review process such that it promotes sustainable development in general, and deliberative democracy in particular? One might expect that – as an organisation having economic growth and expansion as its main objectives – the OECD would not be the best placed to promote the kind of 'institutionalist' view on sustainable development that underpins the present study.

In order to provide a background for the subsequent analysis of the EPR approach, this chapter starts by describing what are the more or less explicit key principles of good environmental policy in the EPRs, as well as the main sustainable development themes addressed in the reviews. It then moves on to the analysis, focusing on four main issues. First, we look at the extent to which the EPRs have integrated to the review framework the idea of the biophysical limits to human activities – that is, the substantive, or 'survival' side of sustainable development. Second, we shall see to what extent the conceptualisation of sustainable development in the EPRs embraces the procedural side of sustainability, manifest in concepts such as coevolution, institutional change, complexity, and methodological pluralism, and whether the EPRs' conceptualisation enhances the understanding of the links between the different dimensions of sustainable development. The third bundle of questions relates to measurement of sustainable development that the EPRs advocate: is the measurement ideal grounded in the ideas of incommensurability of values, balance between qualitative and quantitative information, and appropriate level of aggregation in the spirit of broad concept of internalisation? Finally, the entire EPR process will be analysed in relation to its capacity to enhance deliberative democracy and critical reflexivity.

The analysis relies on a hermeneutic research approach, with an attempt to seize the essence of the OECD message through a method where I interpreted the message of the EPR reports with the help of the general contextual understanding I had gained throughout the WPEP work, and from discussions with persons involved in the reviews. The interpretation of the EPR reports involved moving between the specific passages – sentences and words – and the whole of the text, in an attempt to understand the message in its context. Examples are given to illustrate the arguments and justify the assessment made.

8.1 Key elements of good environmental policy and sustainable development in the EPRs

The structure of the report and the character of the recommendations follow a very similar pattern from one review to another, based on the general review guidelines (OECD 1997b), and the background document outlining the contents of the review and approved jointly by the OECD secretariat and the reviewed country's government. To provide a basis for the subsequent analysis, the principles implicit in the EPR policy recommendations are summarised in box 8.1.

Box 8.1. Principles and criteria of good environmental policy according to the EPRs.

Effective
- Does the country comply with the government's policy objectives achieved and international commitments?

Rational
- Does the government have strategies and plans for the key areas and aspects of environmental policy?
- Policy coherence: are strategies, plans, and actions coherent across sectors?
- Results-based policy: are targets, monitoring and evaluation systems in place to allow the follow-up of the effectiveness of the policies?
- Implementation and enforcement: Are adequate mechanisms in place to ensure the strategies and plans are translated into concrete action (including sufficient financial and personnel resources, legal sanctions, etc.)?

Cost-effective
- Are adequate mechanisms (economic evaluation, cost-benefit analysis) in place to ensure the objectives are achieved at the minimum possible costs to the society?
- Are the external costs of policies on the environment internalised into the price mechanism (polluter pays and user pays principles, elimination of environmentally harmful subsidies, environmental taxes, etc.)?

Transparent and democratic
- Is adequate environmental information made available to the public?
- Do citizens have access to decision-making and the justice system in environmental matters?

Equitable
- Are the costs and benefits of environmental policy (financial burden from policies, health effects from pollution, benefits from amenities, etc.) equally distributed in society, or are some groups particularly exposed?

Precautionary
- Are environmental policies based on a precautionary approach so as to deal with uncertainties?

Balanced
- Are environmental, economic and social considerations taken into account in a balanced way in the design and implementation of environmental policies?

The key for understanding what the EPRs mean by *sustainable development* lies in the second part of the report, dealing with the integration of environmental concerns into economic policies on the one hand, and the interaction between environmental policies and social issues on the other. The reviews take the 'three-pillar' approach to sustainable development as the basis for analysis,

nevertheless focusing on the environmental 'pillar', and seeking to understand its linkages, or 'interfaces' with the two other dimensions of sustainability.

The main topics that the EPRs address when analysing the *environmental-economic interface* can be summarised as:

- *Decoupling* of environmental pressures from economic growth
- Policy integration at the level of *formal institutions* (strategies, action programmes, and councils of sustainable development; regulatory, voluntary, planning and other non-economic policy instruments; Local Agenda 21, etc.)
- *Coherence*, at the different levels of government, of sectoral goals, targets and instruments with those of the government's environmental policy (including e.g. the elimination of environmentally harmful subsidies; mechanisms of cooperation between sectoral authorities)
- *Market-based integration* through the internalisation of external environmental costs through the market mechanism (pollution pays and user pays principles, ecological tax reform, etc.)
- Environmental *expenditure and financing*

The *environmental-social interface*, in turn, consists of four main areas of analysis:

- *Environmental democracy* (access to environmental information by the citizens; mechanisms of participatory planning and decision-making in environmental matters; level of environmental awareness)
- *Distribution* of environmental goods and 'bads' among different population groups (including the incidence of environmental taxes and costs of environmental policy)
- Impact of environmental policies on *employment* (including employment in 'eco-industries' and services)
- *Health* impacts (positive and negative) from environmental degradation and amenities.

The environmental-social interface is treated in a pragmatic *ad hoc* manner, relying on a checklist of issues to be examined, without a theoretical framework based on notions such as 'capabilities' or social capital. The OECD has done some work on environmental-social issues and developed frameworks for their measurement (e.g. OECD 2001e), including work on social and human capital

(OECD 2001a; 2001b), but lessons from this work are yet to be incorporated to the EPRs. The organisation also launched work on analysing the international effects of national policies, including social aspects (OECD 2003c).

Combining the philosophical and theoretical principles underlying the present study (chapter 2), the basic principles of sustainable development seen from the perspective of institutional economics (chapter 3), and the criteria for 'institutionalist' policy evaluation (section 4.2.4), the requirements for environmental policy evaluation pursuing to sustainable development were classified into four categories (Box 4.1):

- recognition of the existence of planetary limits for human activity;
- methodological and procedural aspects related to the definition and conceptualisation of sustainable development;
- measurement of sustainable development; and
- capacity to improve deliberative democracy and critical reflexivity.

The analysis of the EPR approach below will be structured around these four categories.

8.2 Recognition of the 'limits to growth'

As was argued in chapter 3, it is impossible to identify at the global scale the limits to the maximum scale of human activity that would not entail a risk of irreversible damage to vital ecological life-support systems. Such limits are always flexible and depend on the structure of society, in particular on the type of technology applied, population size, consumption level and patterns, etc. Nevertheless, it is evident that ecosystem 'carrying capacity' has been and is being exceeded in a number of local and regional circumstances, that man is affecting the functioning of the global atmospheric system, that the stocks of many non-renewable resources are being exhausted, and that overexploitation is causing irreversible damage to renewable resources in a number of instances across the globe. Our inability to quantify risks involved does not imply that nothing could or should be done against them. Therefore, the first criterion for the OECD reviews from the sustainability perspective is the extent to which they recognise what can be called the questions of 'survival', i.e. the existence of limits posed by the capacity of the biophysical and ecological life-support mechanisms to withstand pressures from human activities. Is the existence of such limits

recognised, and if so, what is the approach adopted? Do the reviews refer to concepts such as critical natural capital, ecosystem functions, resilience, and irreversible, synergistic and cumulative effects? Can they identify clearly unsustainable situations and establish 'ecological guardrails' (Hinterberger et al. 2000)?

The EPRs primarily attempt to evaluate performance defined as the degree of goal-achievement, measured against the country's own policy objectives and international commitments. However, reference is frequently made to 'objective' external criteria to identify clearly unsustainable situations, at least when a fair amount of scientific and/or political consensus exists concerning such limits. Such references are most frequent in the chapters on environmental management, i.e. water, air and nature management, where information on the limits of ecosystem carrying capacities is more easily available than in matters of environment-economy interface, for instance. Hence, frequently mentioned problems include excessive natural resource exploitation (e.g. fish stocks), nitrogen surplus in agriculture, or soil neutralisation capacity (in relation to acidification). In the area of air management, health-related criteria, such as ambient air quality standards, are often used as a starting point. While the reviews mainly focus on trends in performance and therefore give credit to countries for any advances in reducing environmental pressures, they nevertheless refrain from calling policies 'sustainable'. On the contrary, the 'carrot-and-stick' approach applied in the EPRs entails a division of each section in the conclusions and recommendations of the review into two parts – the first one listing the main achievements and advances in the policy area in question, and the second enumerating the problematic aspects. This second part typically starts by sentences such as: "[d]espite this progress, water management in Spain is still *far from sustainable…*" (OECD 2004d, 20); "[n]evertheless, water quality problems are far from being solved…" (OECD 2004d, 18); "[n]evertheless, the state of biodiversity is still declining…" (OECD 2003g, 21); or simply: "[t]hese successes notwithstanding, much remains to be done" (ibid., 20). The reviews do not use concepts such as critical natural capital, ecosystem functions or resilience, but seek support primarily in politically declared objectives and OECD policy principles, and, in the absence of such reference points, in standards fixed by the scientific community.

A shortcoming from the perspective of global sustainability is the reviews' focus on performance at the level of nation-states without sufficient attention to broader impacts at the planetary or international level. The indicators and measures applied in the EPRs hardly touch upon the question of the impacts of the reviewed country's consumption patterns on other countries. The key element of the EPRs' sustainable development perception is 'decoupling', which is systematically assessed

for CO_2 emissions from energy use, emissions of SO_x, and NO_x, nitrogenous fertilizer use, and the generation of non-hazardous waste. Often, the analysis also includes total greenhouse gas emissions, water abstraction, and pesticide use (e.g. OECD 2003g, 99; 2004c, 96; 2004d, 105). Sometimes material flow analyses or material intensity are referred to (e.g. 2003g, 99; OECD 2004c, 97).[309] Nevertheless, the reviews do not try to relate the material flow or decoupling indicators to global limits or targets, but focus on trends, distinguishing 'weak' and 'strong' decoupling. The former denotes a situation where environmental pressures increase slower than the GDP, whereas the latter implies reduction of the environmental pressures despite economic growth. The emphasis tends to be on encouragement and praise, the countries being given credit for their success in the former, 'relative' decoupling, but the need to achieve absolute decoupling is also recognised.[310] The newly started work on material flow accounting under the OECD Annual Meeting of Sustainable Development Experts can be expected to improve the factual basis for analysing impacts on global sustainability.

8.3 The procedural side of sustainable development

The second set of criteria builds on the conviction that sustainable development is not only a question of 'limits' and survival, and that any definition of such limits would be socially constructed and dependent on the context. The first criterion for this 'process aspect' of sustainable development can be formulated as a requirement for the EPRs to:

- See sustainable development not only as an end-state to be achieved, but essentially as a *coevolutionary process* of *institutional* transformation, with an emphasis on institutions, history, and phenomena such as cumulative causation, path dependence and lock-in. Pay attention to both formal and informal institutions and their multiple functions, as well as to the pervasiveness of power and conflict.

---

[309] The EPRs do at least occasionally recognise the problems in interpreting material flow indicator trends, noting that the indicators do not take into account, for example, dissipative use or losses (e.g. copper emissions from brake linings), or the toxicity of the materials (OECD 2004c, 97-98).

[310] On this issue as on many others, the text in the main report is more explicit and critical than the conclusions and recommendations. The conclusions state: *"The United States <u>decoupled several environmental pressures from economic growth</u> over the review period. Pesticide use and emissions of some air toxics have been strongly decoupled, though most pressures (e.g. energy consumption, road traffic, water abstraction) have been weakly decoupled" (OECD 2005c, 25, 121).* In the main body of the report, in the chapter on 'environmental-economic interface' (ibid., 123), one can read: *"…the United States made progress in decoupling environmental pressures from economic development […].*

Despite their focus on performance, measured preferably in terms of quantitative results, the EPRs do pay substantial attention to policy processes. This is manifest firstly in the fact that the second part of the review is entirely devoted to the analysis of environmental policy integration, and secondly, in the great emphasis on the processes of implementation, enforcement, and evaluation of policies, on citizen participation and environmental education, and on institutional arrangements in the fields of environmental policy and sustainable development, such as those aimed at improving the coherence of policies across sectors. The EPRs manifestly attempt to support policy initiatives that are in their 'infancy'.[311]

However, the EPRs tend to largely neglect the informal institutions. Despite paying ample attention to the official regulatory framework, the reviews contain little analysis of the behavioural and cultural aspects, routines, habits, identity, or consciousness of the policy actors. Furthermore, even the analysis of the 'formal' institutions suffers from being largely descriptive, somewhat 'atheoretical' and unsystematic. The criteria of judgement are not always made explicit, and therefore it is sometimes difficult to know on which grounds government's measures are deemed 'insufficient', for instance.[312] Explicating clear criteria for analysis would make the evaluation more transparent, allow comparisons, and improve readability.

The reviews do not explicitly apply concepts such as cumulative causation, path dependence and lock-in. They primarily focus on the efficiency and effectiveness of policies and administrative arrangements, as well as the degree to which the OECD policy principles have been respected. To the extent that long-term dynamic effects are dealt with, this is done within a framework following the principles of modern mainstream economic theory, focusing on issues such as the internalisation of externalities, differences between short- and long-term elasticities, and cost-effectiveness of measures.

The EPRs portray sustainable development and environmental policy in a rather 'apolitical', technical perspective, which may serve to hide the conflicts characterising the issue. The reviews do raise issues of conflict – such as those between different uses of water or environmental amenities,

---

While most pressures were weakly decoupled, […] some were strongly decoupled […] Assuming continued strong growth in GDP, […] weak decoupling will not be sufficient to reduce environmental pressures in the long term."
[311] See section 6.2.2 above.
[312] While it is easy to understand that in the 'performance logic' dominating the EPRs 'insufficiency' of policy measures can be simply deduced from the non-attainment of policy objectives, the matters become more complicated with regard to statements such as "…there is insufficient economic information overall (e.g. on expenditure, employment, water prices) to support environmental management" (OECD 2004c, 128).

or between different sectoral policies (e.g. trade and environmental policies). However, these conflicts tend to be seen as symptoms of a lack of rationality in policies, reflecting and idea that conflicts could be overcome simply by implementing rational and coherent policies across sectors, in the spirit of a single, overarching 'public interest'. One of the manifestations is their tendency to emphasise the synergies between the different 'pillars' of sustainability – the win-win solutions – instead of highlighting the trade-offs involved. A rather frequent passage in the reports is one that takes note of the progress in environmental management in industry and the high standards of environmental regulation, and declares there not being evidence that the competitiveness of the country's industry in the world markets would have suffered from these ambitious environmental policies (e.g. OECD 2003g, 110). Another example comes from the Swedish review, where the report refers to the benefits that a reform of the permitting system would provide both for economic and environmentalist communities: "Long permit processing periods mean high transaction costs and may hinder investment that would be beneficial for both economic development and environmental protection" (OECD 2004c, 17). In general, 'hot' and controversial topics such as climate change policies, or environmental issues in problematic sectors are handled rather cautiously, and the wordings are frequently 'softened' in the peer review meeting in response to resistance from the delegation of the reviewed country.[313] Often, when the reviews address matters of conflict and obstacles to policies considered desirable by the review team, they do this discretely, without pinpointing one specific culprit. The following passage from the review of Canada illustrates this point:

> "A *number of constraints* affect greater uptake of economic instruments. *Industry* is concerned about day-to-day competitive pressures […] It has difficulty understanding how to implement new instruments such as trading. Within *governments*, economic agencies have supported economic instruments in principle, but resisted specific proposals for targeted incentives on allocative efficiency grounds. The *public* is wary of new fees and charges, and of the allocation of 'rights to pollute'. There is a general resistance to external pressure to change consumption patterns. Small but influential groups have blocked some proposals." (OECD 2004f, 124-125.)[314]

---

[313] For example, the U.S. delegation succeeded in the peer review meeting in removing from the conclusions and recommendations comparisons to other countries, indicating that the country has the highest per-capita level of energy-related CO_2 emissions in the OECD. Likewise, the review team's allusion to the modest level of ambition of and the likely future difficulties of reaching the greenhouse gas emission targets were removed from the final text as a result of the U.S. delegation's demands.

As a whole, the reviews – especially the conclusions and recommendations – are designed so as to allow the government and the press to report on the findings noting that the country is on a good path, and has advanced in its environmental policies over the review period, but that a number of problem areas still remain.

The second criterion concerning the process character of sustainable development has to do with the relations between the dimensions of sustainable development and the role of causal explanation:

- Facilitate the understanding of the *links* – synergies and trade-offs – *between the different dimensions* of sustainable development, and at different scales, as well as help distinguish essential from the non-essential; take into account the qualitative differences between the dimensions of sustainable development, therefore conceptualising *the 'social'* through values, meaning, identity, and culture, rather than merely through distributional equity. Apply *causal explanation* as a tool for facilitating deliberation, rather than assuming it as an accurate description of reality.

The three-pillar approach to sustainable development adopted by the EPRs is clear and easy to understand, but entails the disadvantages mentioned in chapter 3. It tends to corroborate the status quo rather than inviting to fresh thinking concerning the trade-offs involved in political decisions relating to sustainability; it reinforces the idea of three separate, yet interacting spheres; and it assumes the three dimensions are qualitatively equal. While the reviews implicitly establish priorities, by devoting more space to economic aspects, for example, the priorities are not made clearly explicit. Some interviewees mentioned the lack of hierarchisation of the recommendations as a shortcoming (section 7.1.2). However, the willingness of the OECD to integrate the social dimension into the review framework supports the perception that sustainable development is not – even for a predominantly economic organisation – merely a question of reconciliation between environmental and economic considerations.

The nature of the EPRs as essentially macro level 'country reviews' renders difficult the integration of local level concerns into the framework of analysis. This is probably most problematic concerning the social dimension, given that local level is where the interfaces between the dimensions become concrete, through aspects related to meaning, identities, and cultural values, for instance. Local level issues are addressed in the EPRs, but their role is clearly secondary compared

---

[314] Italics in the original text.

with the national policy issues. During the review mission, sessions with local level representatives take place in the end of the mission. In the review reports, local level events are used to provide concrete examples, but their function is more that of showing success stories or highlighting particularly difficult problem areas, than providing an illustration of relations between levels. Disparities between regions and social groups, as well as the situation of indigenous peoples are likewise addressed, but more qualitative in-depth analysis would probably be needed if the aim were to mobilise support from the civil society, for example.[315] Apart from the attention given to citizen participation, access to courts and environmental information, and environmental education and awareness, the environmental-social interface is treated mainly as a question of distributional justice, and as noted above, the aspects relating to (cultural) identity and values are hardly touched upon. The reviews make an effort to examine the most evident problems such as acute health problems generated by accidents or toxic pollutant releases, or even the status of indigenous peoples, but fail to address the more day-to-day concerns that go beyond simple economic concerns such as employment. If sustainable development is, above all, about institutional change, the conditions for and consequences of institutional changes should be at the core of the analysis of environmental-social interaction. More attention could be paid, for example, to the obstacles at the individual and social level that stand in the way of the policy reforms required for sustainability. Finally, intergenerational social equity is hardly touched upon in the chapter on environmental-social interface.

Two fundamental problems arise from the way in which the EPRs define sustainable development and the relations between its dimensions. Firstly, the requirement to demonstrate the link between 'the social' and 'the environmental' before an issue can be addressed in an EPR automatically excludes a number of questions that are highly relevant, but involve causal relationships that are complex, intractable, or difficult to demonstrate. Often the link from the social to the environmental operates through the economic dimension. This leads us to the second shortcoming, namely the tendency to analytically separate in the EPRs the three dimensions of sustainability, which abstracts environmental policy and outcomes away from the greater picture and mechanisms of globalisation. The reviews give a good picture of what the reviewed country's government is doing to integrate environmental concerns into economic and sectoral policies, as well as of the degree to which economic growth is achieved without increase in environmental pressures, but do little to help the

---

[315] An NGO representative noted: *"Perhaps it would be better if the OECD used concrete examples in order to evaluate, for instance, citizen participation, instead of using macro level facts. This would make it possible to assess not only the quantity, but also the quality of participation. For instance, in the case of [our country], the quantity of participation increased, while the quality deteriorated."*

reader understand the underlying dynamics of development, globalisation, and interaction between different levels of decision-making.

The third criterion for the procedural side of sustainable development reads as follows:

- Recognise *complexity*, *multiple rationalities* and explanatory frameworks in the spirit of *methodological pluralism*, instead of relying exclusively on a single explanatory and methodological framework based on, for example, efficiency or maximisation of individual utility. Ground policy recommendations in a careful examination of the context, instead of universally valid 'axioms'.

In calling for rationally organised, results-based environmental policy with quantifiable targets, adequate mechanisms of control, (mainly economic) incentives, and sanctions, the reviews implicitly assume that individuals behave – or at least should behave – rationally, calculating and maximising their individual utility. Considering the variety of motivations steering human behaviour might result in a more varied range of policy recommendations. Despite the efforts of the review teams to be sensitive to contextual, country-specific aspects, the final report is nevertheless marked by universalistic tendencies, with by and large the same policy message and similar recommendations given to all countries. Examples include the 'standard' recommendations such as removing environmentally harmful subsidies, internalisation of environmental externalities, introduction of an ecological tax reform, full-cost pricing of environmental (e.g. water) services, and improvement of cost-effectiveness of policies. While such recommendations can be useful general policy guidelines, they should be better adapted to the particular context.

However, the assumption of rational action has its positive side as well. Even if it tends to produce one-sided policy recommendations if taken as a realistic description of the behaviour of individuals in society, it can be helpful in allowing the citizens to evaluate government action: the authorities can obviously be expected to behave rationally, and pursue official policy objectives as effectively and efficiently as possible. From the citizen's perspective, perhaps the main role of the EPRs is precisely checking the 'instrumental rationality' of policies, by examining to what extent and at what cost the government has or has not delivered on its promises.

Despite the tendency of the EPRs to rely on the universal OECD 'axioms', the general principles of good environmental policy underpinning the EPR recommendations have some virtues of

pragmatism. They give particular attention to institutions – notably those facilitating the integration of environmental concerns into other sectors and economic policies – instead of calling for internalisation of environmental costs into market prices as the only solution to environmental problems.[316] The emphasis on implementation and enforcement also reflects a pragmatic approach and understanding of the political 'realities' of the policymaking system. Likewise, the EPRs have adopted the principle of cost-effectiveness instead of efficiency as the leading principle, hence not calling into question the government's policy objectives, but recognising the legitimacy of the politically established environmental objectives. This context-dependence is also manifest in the choice of language: even though the reviews predominantly apply commonly agreed OECD terminology, sometimes the terms used in the review are chosen so as to maximise their resonance in the reviewed country (e.g. choices between 'land use planning' and 'spatial planning'; use of the term 'environmental justice' in the U.S. review).

Nevertheless, the EPRs tend to prefer market-based solutions to other policy instruments. Alone the space allocated in the report to economic instruments greatly exceeds that given to other policy approaches. To take an example, about five pages in the Swedish review and six pages in the Canadian review were devoted to land use planning, environmental impact assessment, strategic environmental assessment, Local Agenda 21, environmental education, and citizens' access to environmental information, whereas the sections on 'market-based integration', economic instruments and subsidies covered 8-9 pages, in addition to the treatment of economic aspects in the different chapters dealing with sectoral and traditional environmental management issues. Internalisation of externalities, introduction of taxes or increasing tax rates, and the urge to respect the polluter pays and user pays principles are among the most frequent recommendations, together with those asking the government to improve or ensure the cost-effectiveness of policies.[317] The stress on the economic aspects of policies is not as such a great shortcoming, given the character of the OECD as an economic organisation, and because of the need to maintain a clear focus and

---

[316] A number of WPEP members as well as officials from the OECD Environment Directorate have criticised the sustainable development section in the Economic Surveys for being based on such overly theoretical 'textbook' approach and focusing exclusively on the economic costs of environmental and social policies, rather than integrating the three pillars of sustainable development and examining the interactions between them. For instance, they have lamented the absence of analysis of the social and environmental impact of economic policies, and the recognition of the need for *ex ante* integration of environmental concerns in economic activities and policies. Moreover, the Economic Surveys are seen to adopt a very traditional approach to environmental issues, focusing on health impacts of certain pollutants, overlooking the impact on ecosystems. (cf. OECD 2004a, 22-32.)

[317] Two examples from the French review serve to illustrate the economic emphasis: "Environmental policy objectives should be clearly stated in the context of marginal costs and benefits (e.g. level of externalities, cost-effectiveness of possible solutions) associated with achieving the objectives" (OECD 2005b, 110), and "…a more rigorous economic approach will be in order, along with development of cost-benefit analysis in river basin management" (ibid., 139).

coherence of the framework of analysis in the reviews. Yet, the EPRs could be more explicit about the reasons for the emphasis on economic aspects and economic solutions to policy problems. If the EPRs are supposed to represent a genuine attempt to look at environmental policies across the board, adopting a broad sustainable development perspective, an exclusive emphasis on economic aspects is not justified.

The principle of cost-effectiveness is not consistently followed all the way through. The government may not have objectives for some policy areas, or the review team may find the objectives lack ambition taking into account the country's potential, and therefore, the EPR may recommend them to be revised.[318] Another deviation is the recommendation to use cost-benefit analysis, which in its traditional form indeed seeks optimal combination of objectives and policies, rather than taking the objectives for granted. The review team does not carry out cost-effectiveness or cost-benefit analyses itself, but the reviews frequently recommend the country to make use of CBA. In the Swedish review, for example, the report implicitly calls into question the reasonableness of precautionary principle as the guiding principle for the country's health policy, arguing that this "may lead to greater expenditure than would be justified by strict cost-benefit analysis" (OECD 2004c, 139).

On the whole, the principles of good environmental policy implicit in the EPRs are biased in favour of market-based solutions, but taking into account the institutional context within which the EPRs operate, such a bias is understandable. Moreover, the EPRs do make a serious effort to integrate institutional analysis and a more understanding approach, but seem to lack the expertise and the resources to pursue very far such an approach.

The last criterion relating to the procedural aspects of sustainable development has to do with the fundamental premises underpinning the concept, and calls for the EPRs to:

- ***Enhance critical reflexivity*** by not taking premises such as progress, modernity, economic growth and technological development for given, as phenomena inherently desirable for sustainable development, but stimulate debate on such fundamental assumptions.

---

[318] Following from the observation that the U.S. climate change policies lack ambition, the review recommends the country to "strengthen national climate change policy, programmes and implementation…" (OECD 2005c, 32, 212).

There is no doubt about the perspective of the EPRs towards the essential precepts of modernity: the reviews are clearly situated within the ecological modernisation ideology, with their belief in 'greening of growth' through decoupling, promotion of eco-technologies, greater reliance on market solutions in environmental policy, and ultimately, the desire to make environmental policies more 'rational' through improved planning, target-setting, and the subsequent monitoring, assessment, and evaluation of goal-achievement. Beyond the analysis of decoupling, the causal relationships between economic growth and environmental pressures are not explicitly dealt with in the EPRs. Together with the decoupling objective, this may be taken as an implicit endorsement of economic growth as an overarching objective for the OECD countries. The question of the desirability of technological development is not dealt with directly, but at least issues such as the call for faster renewal of the car fleet, and the presentation of the development of 'eco-industries' in a favourable light do not attest to 'technological pessimism'. In other OECD work, of course, it is easy to find reference to the importance of economic growth and technological development for sustainable development (e.g. OECD 2001e; 2001f). The emphasis in the EPRs is on 'eco-technologies', implying the belief that technological development can be steered towards a more sustainable path. As we saw in section 8.3 above, the EPRs often seek to demonstrate the existence of win-win solutions, especially the convergence of environmental and economic objectives, in line with the idea, inherent in ecological modernisation, that 'pollution prevention pays'. To some extent, the EPRs also endorse the idea of the changing role of the state, given the emphasis on issues of environmental information, ecological tax reform, decentralisation, participation, and public-private partnerships. However, the EPRs by no means declare adherence to a belief that the state would have lost its importance, quite the contrary: the reviews essentially concentrate on government action, and therefore rather legitimise than discredit state action.

From the institutionalist perspective, the unquestioned endorsement of the blessings of modernity is obviously problematic to the extent that it hinders frank and open discussion concerning situations in which such processes actually promote the 'betterment of the humankind' or sustainable development, and when the negative consequences of modernisation might, by contrast, override the positive ones. Likewise, the *a priori* assumption of the desirability of modernisation masks the distribution of the curses and the blessings of such development behind an argument that 'on average everybody will be better off'. Two further problems with the unreservedly positive attitude towards modernisation in the EPRs are that it precludes a finer, more critical analysis of the broader globalisation processes, as noted above in the previous section, and that in promoting a rationalistic, results- and control-based environmental policy style, the EPRs risk engendering rejection by those

who see the EPRs as nothing but an disguised attempt by the OECD to 'sell' its pro-globalisation message to national environmental authorities.[319]

8.4 Measurement of sustainable development

The third set of criteria for assessing the EPR approach to sustainable development concerns the way of operationalising sustainable development through the methods of measurement that the reviews advocate. The first criterion calls for the EPRs to:

- Ground the measurement of environmental performance and sustainable development in the idea of *incommensurability of values*, instead of seeking a single measure of performance or well-being.

The numerous indicators used in the reviews (see OECD 2003e), in particular those on decoupling and sectoral environmental integration, aim at capturing some of the interactions between the dimensions of sustainable development. The EPRs do not attempt to construct a single, aggregated indicator of environmental performance or sustainable development, but instead their basic philosophy is to measure environmental pressures, the state of the environment, and the policy responses using measuring units characteristic to the matter in question. The fact that the country's own policy objectives are taken as the starting point for evaluation proves that the legitimacy of environmental values is recognised regardless of efficiency concerns. In this perspective, the EPRs are well in line with our criterion. However, the question of commensurability is not addressed explicitly, and the frequent exhortation to use cost-benefit analysis to assess policy options suggests that the idea of incommensurability may not be followed consistently.

The second criterion for the methods of measurement asks the EPRs to:

- Pay equal attention to *qualitative and quantitative* elements.

---

[319] While the OECD has frequently been portrayed as a flag-bearer of neoliberal policies (see section 5.1.1), few interviewees brought up this negative side of the OECD reputation, while the large majority in fact emphasised the advantage of having a prestigious economic organisation give recommendations on environmental policies. An NGO representative, for instance, noted: *"I have a positive attitude, but the NGOs sometimes may think the OECD has an attitude biased in favour of economic interests."* A French civil servant remarked that *"as long as the OECD is not being considered as the standard-bearer of a WTO-style ultra-liberal policy, it will not have serious problems of reputation in France…"*.

Partly because of the macro level of analysis, many non-quantifiable issues are excluded. However, perhaps more problematic and frequent than outright exclusion of the qualitative is the lack of a clear structure and framework for analysing the qualitative issues. This leads to a certain imbalance between the quantitative information on the economic aspects of policies, emissions, norm values, etc. on the one hand, and the general description of the qualitative, institutional issues, on the other. The reader's attention is easily caught by the 'hard' data in the form of tables and figures, whereas information on the institutional issues risks being seen as merely 'complementary'.

Thirdly, the measurement of sustainable development should:

- Seek appropriate level and degree of *aggregation*.

The question of an appropriate level of aggregation entails various facets. On the one hand, the approach chosen for most of the data in the EPRs is to aggregate information at the national level, seek country-level averages, and look at, for example, total numbers of pollution events per year in the whole country (see e.g. OECD 2003e). At the same time, the reviews try to identify and analyse the situation in the most problematic geographic areas and sectors, information on problems and policies related to air pollution in large cities being one of the most frequent examples. Hence, for instance the Italian press eagerly seized the issue of air pollution – pointed out in the EPR as a serious problem in the country – and noted the praise from the OECD to the authorities of the city of Rome in addressing the problem of air pollution.

The adopted level of aggregation may be appropriate, if national level policymakers are taken as the main 'clients' or 'users' of the reviews. While the focus on the national level allows cross-country comparisons, which many stakeholders and the broader public are likely to be most interested in, problems appear when one wishes to motivate a broader range of actors to take part in discussions and actively use the reports. Even at the national level average figures for the entire country are not always particularly policy relevant, since they hide regional and sectoral differences. What such aggregated figures can do, is to help the government legitimise its own work by showing it is doing well on the average. Whether this provides the most pertinent and truthful picture of where the country stands in terms of environmental performance and sustainable development, may be debatable. This is so in the light of the fact that the reviews are not very capable in analysing the indirect impacts of the reviewed country's policies or production and consumption patterns on other – notably developing – countries.

The last criterion concerning measurement calls for the EPRs to:

- Adopt a *broad conception of internalisation*, referring to the entire political process, instead of advocating internalisation simply within a framework of Pareto efficiency.

Where the reports evoke the term internalisation of external costs, they clearly refer to internalisation in the narrow sense, i.e. as monetary internalisation, ultimately based on the ideal of Pareto efficiency in resource allocation.[320] Yet, since the country's environmental policy objectives are taken as the starting point for assessment, this logic is not applied strictly. Moreover, the WPEP sometimes wishes to advocate policies that go beyond mere internalisation, e.g. the use of economic instruments for incentive purposes, regardless of whether or not the externalities in the Pareto sense have already been internalised. For instance, in the peer review meeting of France, a long discussion was carried out concerning the degree to which the external costs of nuclear power were already internalised in prices and whether it would be reasonable to require internalisation beyond the currently known and quantifiable externalities. Even though the review reports stick rather strictly to the conventional internalisation logic, this type of discussions demonstrate that the WPEP sometimes tries to adopt a broader view on the issue. On the other hand, the attention given to distributional impacts of policies, notably in the chapter on environmental-social interface, but also more generally, attests to a broader perspective recognising the incapability of a strict cost-benefit analysis to take into account the distribution of costs and benefits.

Overall, the EPR approach remains very close to the conventional view of internalisation, given the frequent recommendations to strengthen economic analysis of environmental policies, but with few exceptions (see section 8.3 above), the recommendations follow a cost-effectiveness logic instead of calling into question environmental policy objectives in the name of efficiency. Therefore, the EPRs attempt to strike a balance between, on the one hand, the conventional approach, advocated notably by the OECD Economics Department, with its desire to find the most efficient solutions by weighing against each other the costs and benefits of different sets of social, economic and

---

[320] For Sweden, the review notes "[w]hile a shift to renewable energy sources is highly desirable, all energy production involves external costs (which should be internalised), so promotion of energy conservation should be prioritised over subsidisation of even the most environment-friendly types of energy use" (OECD 2004c, 22), and for the U.S. the review recommends to "continue to apply market-based instruments to internalise the environmental costs of <u>energy</u> use in price signals and to reinforce incentives for <u>energy efficiency</u> in the transport, energy and households sectors" (OECD 2005c, 26, 120).

environmental policy objectives, and, on the other, the 'environmentalist' perspective that underlines the incommensurability of different types of values and policy objectives. The behaviour of the WPEP delegates suggests that, probably for tactical reasons, many of them are ready to accept the idea of the limited internalisation, despite its shortcomings. They see the term as useful in so far as it appeals to the economic policy community, and assume that internalisation, even if only in the sense of Pareto efficiency, would lead to higher level of environmental protection.

8.5 Deliberative democracy and critical reflexivity

In the previous sections, the EPR approach was analysed to some extent 'out of the context' – an approach at odds with the fundamental premises of institutional economics. The issues related to the 'limits', the procedural side of sustainable development, and the measurement approach of the EPRs were assessed against an 'ideal' model of institutionalism. Now it is time to come back to the 'real world', and re-examine the contribution of the EPRs taking into account the context – the possibilities and limitations imposed by the institutional setting within which the reviews are being carried out. The key question thus becomes: to what extent do the EPRs exploit the strengths and the opportunities to enhance critical reflexivity that their character as peer reviews carried out by a respected intergovernmental organisation confers them, while simultaneously minimising the disadvantages and limitations imposed by that very context?

The last set of criteria for the EPRs' sustainable development approach examines both the review process as a whole, and the review reports, so as to identify their capacity to foster deliberative democracy and critical reflexivity. The specific assessment criteria are the following:

- The evaluation process is *inclusive*, *deliberative*, and *dialogical*
- The process and the recommendations are designed so as to maximise their potential to *reduce inequalities in the distribution of power;*
- The process and the recommendations *empower civil society* and foster a bottom-up approach to policymaking;
- The process strengthens the *'mediating' type of civil society*; and
- The process and the reports help *overcome dichotomies* such as positive-normative, objective-subjective, quantitative-qualitative, etc.

'Taking evaluation to the people' (section 4.3.7.2), in the sense of truly deliberative evaluation, is not the primary objective of the EPRs, and the review process cannot be considered as particularly inclusive, participatory or dialogical. However, ever since the preparations for the second cycle began, there has been a consensus in the WPEP on the need to increase stakeholder involvement. Moreover, the reviews are inclusive to the extent that stakeholder involvement is relatively broad: a typical review process involves participation by easily a hundred civil servants and representatives of academia, business, civil society, and interest organisations. Yet, the participation does not go very deep, the main role of the participants being to provide information to the review team in the about two-hour sessions organised during the review mission. To a limited extent these meetings do involve dialogue, which is however limited by time constraints. The meetings during the review mission hardly include deliberation, given the objective of the review team to 'extract' information from the informants, without necessarily revealing the review team's opinions.[321]

Participation during the drafting of the review is very limited. The OECD secretariat stays in touch with their counterparts in the reviewed country, mostly to solicit complementary data and documents that the national authorities could not provide during the mission. However, the secretariat tries to minimise this interaction, in order to avoid excessively burdening the national authorities. In any case, this drafting phase involves no deliberation, and hardly any dialogue beyond the review team.

The principal occasion for dialogue and deliberation between the country and the OECD as evaluator starts when the draft report is sent to the country for comments, and in the review meeting in Paris about a month later. The commenting phase does not involve dialogue or deliberation: the government sends its comments to the OECD secretariat, which then classifies them in function of their relevance and degree of controversy, and prepares a new version of the draft conclusions for the peer review meeting. This meeting is perhaps the main occasion for dialogue and deliberation between the reviewed country, its peers from other OECD countries, and the OECD secretariat. These discussions are carried out on a strict government-to-government basis, even though at least once (in the Czech review), the reviewed team delegation included an NGO representative. As we have seen in chapters 6 and 7, many participants have been disappointed at the lack of genuine dialogue during the review meetings, which are often perceived as prearranged 'theatre plays',

---

[321] Officially, the review visit is not supposed to be a 'fact-finding mission', but in practice, it largely serves this function. The review team members have, to a varying degree, already studied the background documents before arriving in the country, and seek during the mission information that will allow them to fill the remaining gaps and thereby to finalise their chapter.

rather than occasions for real exchange of views. Deliberation is, however, constrained on the one hand by the OECD policy doctrine, which the reviews rely upon, and on the other, by the fact that only minor changes to the conclusions and recommendations can be made at such a late stage of the review.

The degree of inclusiveness, dialogue, and deliberation throughout the review process largely depends on the reviewed country, in particular on the way the government organises the preparatory phase and the release and follow-up of the review report. The preparatory phase typically involves extensive dialogue where different government authorities seek common agreement on the country's positions during the review team's visit to the country. A similar procedure is often repeated when preparing the comments for the draft review report. These phases of the process therefore enhance dialogue and often probably also genuine deliberation *among the authorities* involved. By contrast, depending on the country, non-governmental actors participate little or not at all at these stages.

Finally, the release and the dissemination of the review report provides a key occasion for promoting inclusive dialogue and deliberation among all relevant actors in the country. Again, as was shown in chapters 8 and 9, for a number of reasons the reviews often fail to generate broad interest. Their main impact usually remains on the governmental level, the interest of the broader range of stakeholders fading away soon after the report has been released.

Despite this somewhat pessimistic view of the EPRs' role in promoting participation, the reviews have other virtues from the point of view of deliberative democracy. Seen in the light of Hanberger's (2001b) typology of ways of 'taking evaluation to the people' (section 4.3.7.2), the reviews play a double role. On the one hand, carried out by an intergovernmental organisation, the EPRs clearly rely on the model of representative democracy, and a 'positivist-technocratic' approach. The credibility of the EPRs largely depends on their image as objective, fact-based evaluations conducted by a team of experts almost totally independent of the reviewed government. Their role is to serve the bureaucrats directly, and the public at large only indirectly. The EPRs mainly strengthen elitist-representative democracy and a libertarian civil society by adopting an objective, fact-based approach and promoting rationalistic, market-based environmental policy. Therefore, perhaps the main contribution of the EPRs to democracy is their capacity to promote reform of the state and restore the legitimacy of the government through greater transparency.

However, behind this official image of the EPRs, there is a much more pragmatic, advocacy evaluation approach, which has as its main objective to 'empower' environmental authorities through the legitimacy that an evaluation carried out by an international organisation of the OECD's standing can provide. While the process of preparing the review cannot be described as especially participatory with regard to the civil society, it indeed mobilises a large number of authorities, and therefore contributes to intra-governmental deliberation and environmental policy integration. The EPRs clearly seek to obtain a good understanding of the power constellations in the reviewed country, seen as a key requirement for maximising the political impact of the reviews. Through the 'hidden agenda' of empowering the environmental authorities, the reviews come close to the pragmatic evaluation approach, whose main objective is to serve the clients' needs. The EPRs also correspond to the 'advocacy approach' in so far as they help governments carry out self-evaluations and render policies 'evaluable' by pushing the government to establish clear targets. By virtue of the credibility arising from its reputation as an economically-oriented, highly independent international organisation, providing reliable information in practically all sectors of public policy, the OECD is particularly well-placed to carry out reviews that, ideally, have the capacity to generate public debate and deliberation on environmental and sustainability policies in the reviewed countries, and 'shake' habitual ways of thought. It does this, on the one hand, by instilling a dose of economic rationality to the 'environmentalist community' in general, and the environmental administrations in particular, and, on the other hand, in other sectors, by reminding about the importance of environmental concerns in sectoral decision-making.

The danger that an advocacy approach would end up promoting a 'communitarian' civil society by unduly strengthening one section of the society does not seem serious in the case of the EPRs, given the underprivileged position of the environmental authorities within governments in general. Rather, such empowerment and enhanced legitimacy conferred to the 'environmentalist community' can be seen to improve the prospects for better deliberation by redressing asymmetries of power. However, the combination of such a pragmatic approach and the need to safeguard the image of the review team as an objective, uncompromising source of independent assessment leads to a somewhat schizophrenic situation, with the need to retain a 'façade' of uncompromising objectivity, despite the actual review practice being close to advocacy evaluation.

The EPRs can hardly be seen as an instrument in favour of a 'mediating' type of civil society, because of their very limited capacity to engender greater engagement and responsibility on the part of ordinary citizens and voluntary organisations.

Moreover, the peer review mechanism represents a unique attempt to overcome the dichotomies between the evaluations aiming at learning on the one hand and accountability on the other, given their ambition to conduct independent, external reviews in a cordial, non-adversarial atmosphere in close cooperation with the key client groups, i.e. the environmental authorities of the reviewed countries. This position would allow the reviews to promote critical reflexivity by, on the one hand, retaining sufficient distance in relation to the evaluated country government to avoid co-optation, and on the other, by engendering sufficient mutual trust to avoid rejection and resentment.

The EPRs' survival and influence largely depends upon their ability to exploit their room for manoeuvre and deviate from the OECD mainstream repertoire on the one hand, and from the environmentalist repertoires on the other. The potential of the reviews to have an influence is greatest at the interface of the two repertoires – when they succeed in mobilising, as change agents, the 'OECD-minded' civil servants in the environmental administrations and the 'environmentally-oriented' civil servants in other sectors. While not totally absent, this potential is at present far from fully exploited, first, since the review team does not engage in a real dialogue with the national experts, but rather retains a critical distance in order to safeguard its independent image, and second, because of their incapability of engendering broad debate in the reviewed countries.

One of the main reasons evoked for the lack of debate is the rather' dry' and 'austere' tone and appearance of the review reports, which does not easily stimulate interest beyond government circles. The NGOs and the media have a hard time finding hot revelations or material for front-page issues in the reviews. Moreover, many environmental authorities complain that the reviews are simply too 'nice' and general to really provide tools that would be useful in intra-governmental negotiations. The problem is therefore not so much that the balance between advocacy and objectivity would be incorrect, but rather that the EPRs are sufficiently independent neither in relation to the fundamental philosophical and theoretical premises underlying the OECD policy doctrine, nor in relation to the sections of the government and environmental administration afraid of criticism. Building trust and confidence is certainly useful, but critical reflection requires that the reviews succeed in shaking the fundamental assumptions and habitual ways of behaviour.

Finally, a key problem preventing the reviews from stimulating reflexivity is that they fail to engender enthusiasm among the WPEP members and the environmental authorities. The dry and austere appearance of the reports may not be the fundamental reason, but rather the lack of

innovation and reform in the review programme. The degree to which the EPRs are able to mobilise the 'momentum' needed to stimulate sufficient dissemination and follow-up, for instance, in the reviewed countries, depends not only on 'rational' calculation of the benefits and disadvantages of the reviews, but on more subjective, even emotional aspects. Bureaucracies too depend on the enthusiasm of individual civil servants in order to prosper, develop, and renew themselves. At present, too many participants in the review programme seem to perceive the EPRs as 'more of the same', the changes in the review programme having been perceived as insufficient.

Summing up, while the EPRs cannot be seen as an ideal example of the 'mediating approach' to taking evaluation to the people, they do nevertheless have the capacity of enhancing deliberation. Their potential to foster deliberative democracy arises from their ability to 1) improve the *factual background* upon which public debate would be conducted; 2) help make governments *accountable* to their constituencies, thus strengthening the democratic control of those in power; and 3) *'empower'* the environmental authorities in relation to other sectors within the government; and 4) strengthen the status of environmental issues on the agenda of public policies in general, thus helping correct imbalances of power and approach an 'ideal speech situation'.

Chapter 9

Conclusions

In this thesis I have analysed the OECD Environmental Performance Review programme in order to illustrate the mechanisms through which international environmental policy evaluations can influence countries' efforts in moving towards sustainable development. More specifically, I sought to find out *what consequences, through which pathways, and under which circumstances, does country-level environmental policy evaluation, carried out by an intergovernmental organisation, have for institutional change towards sustainable development.*

To answer the question, I chose the OECD Environmental Performance Reviews (EPRs) as a case study, as one of the few attempts at the international level to evaluate industrialised countries' environmental policies in a comparative perspective, with the ultimate goal of moving towards sustainable development. The interest of such an analysis stems from the increasing use of environmental policy evaluation as a policy instrument in the new governance context of growing influence of non-state actors, including the international organisations.

The main research question was further divided into three sub-questions relating to the OECD Environmental Performance Reviews, concerning (i) the ways in which the EPRs have influenced policies in the reviewed countries, (ii) the factors that condition this influence, and (iii) the degree to which the influence of the EPRs promotes or prevents progress towards sustainable development, seen from a coevolutionary institutional economics' perspective.

The thesis was started by developing a framework of analysis, first defining the philosophical and theoretical premises underlying the study, and putting forward the arguments justifying the choice to rely on institutional economics as the theoretical foundation of the study. In order to develop a framework for analysing the impacts of the EPRs on sustainable development, I defined sustainable

development as a combination of substantive and procedural elements – consisting on the one hand of the imperatives stemming from the limits – even though flexible – to human activity posed by the biophysical and ecological life-support systems, and the deliberative institutions providing the basis for social learning. Sustainable development should be seen not only as a desirable end-state, but also, and above all, as a process of institutional transformation through democratic deliberation.

The brief survey of the evolution of evaluation research and practice showed that the 'paradigm wars' that have characterised the development of social sciences over the past decades can also be found in evaluation research. Overcoming dichotomies that have underpinned such battles between different schools of thought is a key challenge for an 'institutionalist' evaluation of environmental policy for sustainable development. Bridging the positive and the normative, quantitative and qualitative, explanation and understanding, judgement and learning, by combining the substantive and the procedural elements of sustainable development should help us move towards a more reflexive society. To do so, evaluations should have the capacity to enhance critical reflection on the fundamental values of society, be based on a clear understanding of the prevailing power relations, and perceive conflicts as a potentially positive vehicle promoting institutional change in society. They should enhance understanding on the processes behind the policy outcomes, i.e. answer the 'why' questions, while at the same time measuring goal-achievement and cost-effectiveness; and integrate in the analysis on an equal basis aspects relating to both instrumental and communicative rationality, quantitative and qualitative elements. An evaluation should be based on a clear understanding of the context, endorse multiple rationalities and methodological approaches, and strive to improve conditions for deliberative democracy by 'empowering' weaker actors in society. Finally, the measurement and evaluation of environmental performance and sustainable development should be based on the idea of incommensurability of values in an attempt to enhance deliberation rather than coming up with one single, incontestable 'correct' solution or judgement.

These demands pose formidable challenges to environmental policy evaluation, and it is unlikely that all of them could be achieved simultaneously or that the ideal of an 'institutionalist' evaluation could ever be fully attained within a single evaluation exercise. In the end, taking the context seriously calls for certain modesty and attention to the broader picture: in a specific situation, an evaluation may better promote sustainable development by not trying to dogmatically comply with all the criteria, if the overall context requires a different approach.

9.1 Are the ERPs influential and if so, why?

Before asking the fundamental question – do the EPRs enhance sustainable development? – it is useful to first determine whether the EPRs produce any policy effect whatsoever. The two first research questions, related to the pathways of influence and the factors conditioning the influence, were analysed through frameworks based on experiences from evaluation research. The concept of 'evaluation influence' (Henry and Mark 2003), complemented by Valovirta's (2002) analysis of evaluation as an argumentative process, served to as the basis for examining the pathways of influence at individual, interpersonal and collective levels. Hanberger's (2001b) ideas concerning the impacts of evaluation on democracy provided an additional angle on evaluation influence. Finally, literature on the use and influence of evaluation was employed to identify the main categories of factors potentially conditioning the influence of the reviews: the role of an international organisation as evaluator; the involved actors with their expectations, interpretative frames and policy networks; the general policy context; and the evaluation design itself.

The specificity of the OECD peer reviews is that they attempt to combine the two seemingly contradictory functions of learning and accountability within a single evaluation exercise. The former focuses more on aspects such as cross-country comparability, performance measurement, and issues of high political interest, whereas the latter adopts a 'softer' approach, emphasising dialogue, mutual trust, explanation of (causal) relationships between policies and outcomes, success stories and policy failures, etc. These two types of influence were further divided in two, so as to constitute the four parallel pathways of influence described in section 7.1. The attempt to combine learning and accountability is challenging, to say the least, and often, the influence from the EPRs has been disappointing to the parties involved. It is not uncommon that reviews actually do not exert significant influence through any of the three first pathways. Pathway A ('advocacy through intragovernmental debate') may fail simply because of the lack of motivated change agents in key positions in the government; pathway B ('advocacy/accountability through public debate') frequently suffers from the lack of comparability and from the highly diplomatic tone of the reports, hardly appealing to the more conflict-oriented 'repertoires' dominant among the media and the environmental NGOs. Often, the EPR process is not sufficiently dialogical to engender a genuine learning process through pathway C ('learning from best practice'). Moreover, the lack of an explanatory element in the reviews does not facilitate learning either. Yet, the interviews showed that the EPRs are only rarely considered as completely useless, since they are seen to help

legitimise environmental policies and authorities on the one hand, and to 'infiltrate' the OECD-type results-based, market-oriented policy style within environmental administrations on the other.

Whichever, if any, of the three first pathways of influence dominates, and no matter how 'well' the EPRs are designed, their influence depends to a major degree on the country-specific institutional context. While this study did not allow far-reaching conclusions to be drawn on the country-specific factors, a number of tentative hypotheses can be put forward for further testing. In summary, the EPRs tend to be most influential in small countries that are sensitive to outside opinions, have recently joined the OECD, have relatively weakly developed environmental policy institutions and a weak tradition of environmental policy evaluation. Other factors identified in this study mostly affect the choice between the different pathways of influence and their final impact depends on their interaction with other aspects of the country-specific setting, relations of power between stakeholders, and the review design. Such aspects include the degree of networking within the administration, the country's 'discussion culture', whether 'hot' issues are clearly highlighted during the review process and at the moment when the report is being released, the country's discussion culture and policy style.

9.2 EPR message in the light of the 'institutionalist' conception of sustainable development

This thesis has suggested that 'enlightenment' – the indirect and subtle processes of long-term adoption and adaptation of OECD ideas into national policies and the conceptual frameworks applied by policy actors – may often be the most important and lasting type of influence from the reviews, a finding largely in line with experience from evaluation research. This leads us to consider what can be seen as the ultimate research question of this thesis: do the EPRs – taking into account the entire review process – promote sustainable development? At least 'insiders' in the EPR programme may have the tendency to simply think that the more influential the reviews, the better for sustainable development. However, the influence from a review may be substantial, for instance through the empowerment of environmental authorities, yet the broader consequences for sustainable development may be detrimental. The main concern in this regard has to do with the type of conception of sustainable development that the reviews transmit. Are they in line with the institutionalist premises outlined in chapters 2-4, or do they rather perpetuate the 'outdated ideas' (Boven 2003) that stand in the way of sustainable development?

As we saw in chapter 8, the EPRs are to a large extent based on a modern mainstream economics perspective on sustainable development, but they do integrate a number of 'institutionalist' elements. The reviews attach great importance to non-market institutions in environmental policies, integration of environmental concerns into sectoral policies, policy coherence across sectors, as well as the implementation and enforcement of policies. The willingness to look at the environmental-social interface, including environmental education, citizen participation and access to environmental information and the justice system demonstrates openness towards a perspective seeing environmental policy in the broader context of sustainable development. The analysis is based on a disaggregated approach, to the extent that no attempt is made to arrive at a single measure of environmental performance or sustainable development. However, the message of the EPRs concerning the substance of environmental policy and sustainable development is firmly based on the basic tenets of modernity, progress, and the ideal of objective, value-free science. The faith in rational, results-based, market-oriented environmental policy, and 'greening of growth' in the spirit of ecological modernisation are not called into question. The analysis of institutions remains rather 'atheoretical', without due attention given to informal institutions, behavioural aspects, and alternative rationalities beyond the economic one. The reviews hardly try to explain the reasons behind policy results, which compromises their capacity to promote policy learning. The three-pillar approach to sustainable development, with an emphasis on 'win-win' solutions, does not allow the political character of the issues at stake to be explicitly brought forward, or help the reader see the trade-offs between the different 'pillars'. The environmental-social interface is examined almost exclusively as a distributional question, without due attention given to behavioural and identity-related aspects. Both the global and local aspects of sustainable development are sometimes obfuscated by the focus on the nation-state as the key unit of analysis: the decoupling indicators only insufficiently take into account the effects of production and consumption patterns beyond the country's borders, and the articulation between local, national and global levels is not addressed to a greater extent.

9.3 Should the EPRs have influence?

We can conclude, then, that the EPRs do, indeed, influence policies through the three first pathways identified in chapter seven, representing accountability, advocacy and relatively short-term learning, but this impact is usually perceived as modest, and greater influence depends on favourable political conditions, and country-specific aspects. The long-term impacts on mental models and frameworks of thought may, in the end, be the most important ones, but the view of sustainable development transmitted through the EPRs is rather far from the institutionalist ideal. Why then, do I nevertheless argue that it is worthwhile an attempt to try and maximise the impact of the EPRs? My arguments are based on the three principal impacts of the reviews, each of which has its downsides.

Firstly, the EPRs have the capacity to improve the conditions for deliberative democracy, by providing a more equal 'playing field' and 'empowering' the weaker groups – mainly the environmental authorities – in relation to their more powerful economic counterparts. This was an aspect brought up in the majority of the discussions and was seen by many interviewees as the main contribution of the reviews.

Secondly, the reviews help integrate environmental concerns into other sectors of administration, with the reputation of the OECD as a respected and reliable economic organisation as an asset – a guarantee of quality, credibility and legitimacy – among the actors in the non-environmental sectors. Much of the value of the OECD work in the field of the environment stems from its capacity to build links between the economically and environmentally oriented repertoires within the state bureaucracies.

Thirdly, the EPRs act as a tool for 'infiltrating' economic rationality, performance-orientation, and evaluative thinking – the 'OECD doctrine' – to the environmental administrations. This is justified not only because of the benefits it brings in terms of more rational use of resources and concern for accountability concerning policy results, but can also help win support for environmental policies in governments and among the broader public. Moreover, even though the basic 'philosophy' behind the EPRs is to give credit for progress made and thereby foster continuous improvement of policies instead of applying an absolute standard of sustainability as a 'measuring rod', to a certain extent they also enhance understanding of the 'substantive' side of sustainable development, by reminding of the limits posed by ecological systems.

Each of these three advantages has its reverse side, of course. The empowerment of environmental administrations and environmental policies may not be desirable for broader sustainability concerns if it allows opportunistic governments, ministries or politicians to legitimise inefficient, non-egalitarian or otherwise 'poor' policies. It may likewise be counterproductive to the extent that the EPRs, rather than enhancing transparency and greater openness, empower environmental bureaucracies not only in relation to other sectors of administration, but also in relation to the civil society. And, in the final analysis, one has to ask whom the EPR in fact legitimise – the environmental administration or the government as a whole?

Nevertheless, there are good reasons to believe that legitimising environmental policies and authorities is, indeed, desirable from the point of view of sustainable development, already because the environmental bureaucracies tend to be the main advocates of sustainable development within the governments. However, such advocacy should not be based on political expediency, favouritism or random selection of issues to be advocated, but instead on clear analysis of the prevailing relations of power, and transparent methods and criteria of evaluation, so as to minimise the possibilities of opportunistic 'greenwashing' on the part of the governments. To achieve this, the EPRs would probably benefit from an approach that would allow better comparability across countries and more transparent criteria for assessing in particular the more 'institutional' issues of policy implementation, enforcement and environmental policy integration. The idea would not be to transform the EPRs into mere compilations of quantitative indicators – such an idea would go against the principle of 'weak comparability' and would risk provoking more resentment than interest in genuine learning. Rather, in order to bring the non-quantifiable, institutional aspects of environmental policies on par with the quantifiable ones, to facilitate comparisons across countries, and to exploit the role of the OECD as an 'ideational agency' (Marcussen 2001), the OECD should seek to operationalise concepts such as environmental policy integration by developing criteria flexible enough to accommodate variations in country-specific contexts, yet sufficiently rigorous to enable comparisons and make explicit the criteria underpinning the judgements.

The problems associated with my second argument for the EPRs – enhancement of environmental policy integration through the prestige of a respected international economic organisation – come back to the downsides of policy integration evoked by Hukkinen (1999). Do the EPRs enhance sustainability, if the problems is not so much the lack of cooperation and coordination across policy sectors, but rather the systematic preference given to short-term interests over the long-term

concerns for sustainability, and, more concretely, the unequal distribution of power in favour of sectoral and economic authorities in relation to their environmental counterparts? Do not the 'soft' advocacy for policy integration and 'win-win' rhetoric that underpin the EPRs win the support of the more powerful sectors only at the cost of diluting the message of sustainable development to an extent of making it simply an issue of efficient environmental policies – implying that environmental (or social) concerns should be taken into account, but only as long as they do not harm economic development?

One possibility of avoiding such risks of consensus around the smallest common denominator – largely dictated by the more powerful 'economist community' within the OECD – would be to abandon the requirement that, for instance, all public documents produced by the different units of the organisation should be in strict coherence with each other. In concrete terms, instead of attempting to harmonise the views across the different OECD peer reviews, to arrive at an 'OECD view on sustainable development', it might be more fruitful to allow each one to defend its own perspective, even if this would lead to contradictory conclusions across the reviews. Of course, this would require that any peer review would clearly make clearly explicit its underlying basic premises. Such an approach would abandon the idea of 'monism' or 'unity of science', in favour of a view recognising the irreducible differences between the descriptions of the reality from different disciplinary and methodological perspectives (see chapters 2 and 3).

My third argument evoked above in favour of strengthening the influence of the EPRs – the 'infiltration' of 'OECD thinking' into the environmental administrations – is perhaps the most problematic one. While the OECD doctrine certainly has its qualities in terms of efficiency and more results-based thinking, we are here faced with fundamental questions concerning what Norgaard (1994) has called the metaphysical and epistemological premises underpinning modernity. These issues still continue to engender heated debates also among scholars working on policy evaluation, between the advocates of the theory-based and 'realistic' evaluation, as opposed to the more constructivist approaches. It is certainly not possible, on the basis of this thesis, to unambiguously establish whether or not the EPRs serve to perpetuate the false beliefs in the capacity of modern societies to achieve sustainability through more control, more rational and sophisticated methods of monitoring and evaluation – more modernity, 'more of the same'. Neither can we assert that the EPRs would be an island of potential change and dissent within the ocean of OECD work that is essentially founded in the modernist tenets. Nevertheless, from a very down-to-earth standpoint, it is safe to assume that the OECD would not disappear even if the EPRs were

abolished, and that the organisation would continue to influence thinking within the state bureaucracies both in the OECD member countries and beyond. The EPRs certainly cannot completely overturn the dominance of economist thinking in the OECD doctrine, yet they do have the capacity to infiltrate some environmentalist arguments in the dominant OECD repertoire. To maximise such capacity, the EPRs should be based on a constant effort to ensure reflexivity – preparedness to call into question and test the validity of the OECD 'axioms' – instead of simply taking them as self-evident starting points for the EPRs. In practice, this would involve deeper analysis of the reasons behind policy outcomes: by studying the implementation and the consequences of policies following (or not) the OECD principles, the results from the EPRs could feed back to the OECD work more generally, and if needed, lead to adjustments in the organisation's policy doctrine. This would also lay the basis for a better use of the considerable expertise engendered by the EPRs within the OECD Environment Directorate, through synthesis reports drawing together and bringing to a more general, theoretical level the lessons learned from the EPRs.

9.4 OECD as an evaluator: independent source of Truth or a respectable policy advocate?

A great deal of the policy impact of the OECD in general, and through the EPRs in particular, stems from its symbolic power, based on its prestige, credibility, and perceived independence. Therefore, both the difficulties and the potential of the EPRs to make a difference relate to the role of the organisation, and its capacity to link with the dominant 'repertoires'. The EPRs face formidable challenges in trying to straddle between the contrasting demands from the different repertoires within the environmental administrations, other sectors of government, NGOs, the media, and the other parts of the OECD – the Economics Department and the associated 'economist community' in particular. Credibility, salience and legitimacy stem from different factors depending on the repertoire. Any changes proposed to the EPR framework need to be seen against this background. For example, sharper and more specific recommendations, as well as a wider use of quantitative indicators and comparison between countries would likely help enhance the salience and credibility of the reviews, in particular in the eyes of the NGOs and the media, but might on the other hand weaken their legitimacy in the eyes of their main clients – the environmental administration of the reviewed country – and bring about resistance on the part of sectoral ministries that would feel unjustly 'attacked'. Strengthening the explanatory element of the reviews, by contrast, would

enhance their salience and credibility especially to environmental administrations by helping understand the causal relationships between policies and outcomes, and thus strengthening the arguments behind the recommendations. The reverse side is that, given the limited resources available for the review programme, this would render it more difficult to simultaneously improve comparability through the use of quantitative indicators. Moreover, the civil servants whose work would thus come under closer scrutiny might perceive such an approach as threatening – and hence less legitimate.

Another dilemma for the OECD is how to balance between the conflicting roles of an advocate and a politically neutral, independent evaluator. The credibility of the organisation has traditionally rested on the latter, whereas the failure to acknowledge the former would be politically naïve, and would greatly compromise the potential of the reviews to influence policies. Finally, the OECD Environment Directorate – more specifically, its Environmental Performance and Information Division – must take into account its own credibility and legitimacy within the organisation, which limits the degree to which it can diverge from the general OECD policy doctrine. This limitation somewhat diminishes the credibility of the EPRs among the environmentalist community, in so far as the EPRs are seen as too accommodating in relation to the economic discourse dominant in the OECD. However, such attitudes seem relatively rare, the most common view being that the OECD's role as an economic organisation in fact increases the value and weight of its message in environmental matters.

A number of participants in the EPRs have argued for a more participatory approach, and greater involvement of stakeholders in the review process. This would at least at the first sight seem to be in line with the principles of deliberative democracy, and susceptible to dispel the criticisms arguing that the OECD reinforces an authoritarian, objective, top-down policymaking. But increased participation would come at a cost, and might not always improve the potential of the EPRs to work for sustainability. At least three arguments can be evoked against a more participatory approach in the EPRs. First, the resources available for the reviews would not allow genuine participation, and might simply frustrate the participants, who would feel participation has been organised only to legitimise the EPR approach. Second, the role of the EPRs as a forum for 'cool' discussions on 'hot' topics would be compromised if a broader range of participants were involved, aggravating the risk of political power play and demagogy. This could compromise the potential of the reviews to promote learning through relatively frank and open discussions among delegates. On the other hand, of course, if the stakes were higher in the EPRs, this could motivate those participants that at

present perceive the EPRs as a useless 'theatre play'. Third, the independence of the review team, which the credibility of the reviews rests upon, would come under greater pressure, as the reviews would become more 'politicised'.

It can be argued that the EPRs' capacity to equalise power relations, and thereby enhance deliberative democracy, in fact depends on certain non-transparency, which facilitates more genuine deliberation among the governmental stakeholders. Of course, truly deliberative, inclusive and dialogical processes are essential if one wishes to enhance deliberative democracy and sustainable development, but I would suggest that the comparative advantage of the OECD in carrying out evaluations lies elsewhere, notably in providing a reasonably independent expert view of a country's policies, thereby stimulating debate at the national level. The suggestions presented in the previous section should enhance the capacity of the EPRs to stimulate such debate.

The role of the OECD also helps understand some of the reasons for failure of many of the reform proposals put forward by the WPEP delegates and others involved in the EPRs. The widely contrasting views within the WPEP, which has made it difficult to achieve consensus within the group. Together with the need of the Environment Directorate to ensure the support from its main 'clients' – the environmental authorities – and remain close enough to the mainstream OECD doctrine largely influenced by the Economics Department, the absence of consensus has made reforms difficult. Moreover, the OECD Environment Directorate seems to have skilfully exploited the lack of consensus and determination among the member countries in order to avoid any drastic changes that might have threatened the status quo. While many of the suggested reforms to the review programme have reflected lack of realism and deeper understanding of the OECD working methods, the fact remains that the slow pace of change has frustrated a number of participants, and thus further reduced the enthusiasm and motivation of many actors needed to make the EPRs better known in the capitals and function as 'change agents'.

In part because of their diplomatic approach, the EPRs do not seem to fully satisfy anybody, but only very seldom do they engender strong resentment either. Likewise, the EPRs do not fully exploit the possibilities of the OECD in any of the four roles of ideational artist, arbitrator, agency or agent (sections 5.1.2 and 5.1.3). An optimist would affirm that the EPRs have succeeded in striking a good balance between the conflicting objectives of different stakeholders, whereas the response from a sceptic would involve noting that EPRs do not generate strong feelings since they are completely toothless and hence irrelevant. What seems clear is that the EPRs are bound to lose

importance unless they are capable of innovation – carrying out a third review cycle very similar to the two first ones would be unlikely to receive sufficient support from the member countries. The programme needs to break the self-reinforcing spiral: lack of innovation in the EPRs -> declining utility of the reviews -> diminished political support for the continuation of the programme -> cuts in budget and/or personnel -> reduced capacity in the secretariat to innovate -> increasing gap between countries' expectations and what the EPRs can deliver in practice -> declining utility of the reviews…

The highly diplomatic tone of the EPRs has been defended on the grounds that it enables sufficient ownership over the reviews by the key stakeholders, avoids resistance and resentment against the reviews, thereby creating an atmosphere of mutual trust and confidence, necessary for learning. It has been argued that a stronger, more coercive approach, in the spirit of accountability evaluation, would be against the nature of the OECD as an international organisation. The experience from the PISA educational assessments has clearly demonstrated that there is nothing in the role of the OECD that would prevent such an approach. The PISA assessments have generated lively debate both in countries arriving at the top of the ranking (e.g. Finland) and in those surprised by the mediocrity of their own performance (e.g. Germany). Whether applying in the EPRs ranking of the kind practiced in PISA would work in favour of sustainable development, is another issue; sceptical voices and outright rejection of the results have been expressed not only in the poorly performing countries, but also in those at the top of the ranking.[322] Nevertheless, one could expect the combined effect from the abandonment of the ideas of a single 'Truth' and the unity of science on the one hand, and a more courageous, comparative approach – even ranking of countries – on the other, to lead to a gradual learning process, in which the OECD reviews would enter into public debate, allowing critical assessment of not only the findings of the reviews, but also of the methods of assessment applied. This could be a way for the OECD to regain some of its lost status in the arena of global governance, yet without the burden created by the requirement to represent infallible objectivity. However, such reform might require the organisation to abandon its consensus rule at least in some areas of OECD decision-making – a reform far from easy to drive through.

---

[322] A comment by a Finnish professor of pedagogy, Michael Uljens, illustrates the doubts raised by PISA studies:*"Is this what the OECD wants? To bring about cross-country competition? The PISA phenomenon testifies to a neoliberal educational policy message, which we uncritically swallow. This is a ranking culture by God's grace: only one can win. In Finland, we have of course all the good reasons to be happy and content, but we also have to ask ourselves if this is the type of school system whose creation we want to contribute to"* (Skolranking blev…2005). (My translation).

9.5 Suitability of the research method to the task

The interdisciplinary research approach adopted in this study allowed an analysis of the OECD EPR programme from a broad perspective that would not have been possible by limiting the analysis to a single discipline. Combining the basic premises of institutional economics with the experience and concepts from evaluation research turned out useful in defining a set of criteria for institutional evaluation of environmental policies in the pursuit of sustainable development. Insights from evaluation research proved particularly helpful in identifying the indirect sources of influence from the EPRs, through learning, dialogue, argumentation and 'enlightenment' throughout the review process. The case study helped understand on the one hand the ways in which evaluations carried out by an intergovernmental organisation without direct regulatory power can influence policies in the evaluated countries, and on the other hand, the working methods of the OECD and its potential role in international governance.

The decision to look at the EPR programme as a whole, rather than concentrating on a more in-depth study on the impact of a review in one country, prevented me from studying in detail the mechanisms and pathways of influence in specific situations over a longer period of time. Such an approach would have provided a more nuanced picture of the ways in which the reviews actually influence patterns of thought in the long term, but would, on the other hand, have compromised the ability to obtain a broader picture of the review programme. Moreover, it would have not allowed seeing the variations in review influence across countries and in function of different policy situations. The present study could constitute the basis for further studies on the specific mechanisms through which the EPRs operate in the reviewed countries, and on the obstacles that stand in the way of review influence.

The informants used for analysing the impacts of the EPR programme consisted to a great extent of government officials, with scarce participation by representatives from the NGOs and other non-governmental stakeholders. This can be considered as a weakness in the research setting, as the results fail to fully take into account the views from the civil society. While such a bias towards government views was partly dictated by the time constraints and the fact that, in the Netherlands and in Portugal, the countries WPEP delegates chose the interviewed persons, it was also a conscious choice based on the observation that the NGOs did not seem to use the review to any substantial extent. This came out in several interviews, in informal discussions with WPEP

delegates, and through my own observations, particularly following the Finnish review. However, to the extent that an attempt will be made in the future to render the EPRs more attractive and useful for a broader range of stakeholders, for instance by strengthening the role of 'pathway B' described in this study, views from the civil society would be desirable.

Even though my position as an 'insider' in the OECD EPR programme may have introduced a certain bias in favour of the EPRs and the OECD work in general, on balance, the experience from OECD work was an indispensable asset, which gave me access to such tacit information that would would have been unavailable for a researcher coming from the outside'.

References

Abma, Tineke A. 1998. Text in an Evaluative Context: Writing for Dialogue. *Evaluation* 4(4): 434-454.

Abma, Tineke A., Jennifer C. Greene, Ove Karlsson, Katherine Ryan, Thomas A. Schwandt & Guy A.M. Widdershoven. 2001. Dialogue on Dialogue. *Evaluation* 7(2): 164-180.

Adger, Neil W. 2001. *Social Capital and Climate Change*. Tyndall Centre for Climate Change Research Working Paper No. 8. October 2001. School of Environmental Sciences, University of East Anglia.

Adler, Emanuel & Haas, Peter M. 1992. Conclusion: epistemic communities, world order, and the creation of a reflective research program. *International Organization* 46(1): 367-390.

Ahonen, Pertti. 2001. Evaluaatio-oppia edistyville (Evaluation for the advancing learner). *Hallinnon Tutkimus (Administrative Studies)* 20(3): 102-117.

Alasuutari, Pertti. 1999. *Laadullinen tutkimus (Qualitative Research).* 3. uudistettu pianos. Vastapaino. Tampere. 317 p.

Albæk, Erik. 1995. Policy Evaluation: Design and Utilization. In: Rist, R. (ed.), *Policy Evaluation: Linking Theory to Practice*. London. Edward Elgar. Pp. 5-18. (Reprint of: Albæk, E. 1989/1990. Policy Evaluation: Design and Utilization. *Knowledge in Society* 2(2): 6-19.)

Alchian, Armen A. & Demsetz, Harold. 1972. Production, Information Costs and Economic Organization. *American Economic Review* 62(5): 777-795.

Ayres, Robert U. 1998. *Turning Point: An End to the Growth Paradigm*. London. Earthscan. 192 p.

Ball, Alan & Peters, Guy. 2000. *Modern Politics and Government*. London. Macmillan Press.

Ballet, Jérôme, Jean-Luc Dubois & François-Régis Mahieu. 2003. *Le développement socialement durable : un moyen d'intégrer capacités et durabilité*. Paper presented at the third Conference on the Capability Approach, University of Pavia, 6-9 September 2003.

Balthasar, Andreas & Rieder, Stefan. 2000. Learning from Evaluations: Effects of the Evaluation of the Swiss Energy 2000 Programme. *Evaluation* 6(3): 245-260.

Barbier, Jean-Claude. 1999. Inter-governmental evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives? *Evaluation* 5(4): 373-386.

Barnes, Marian, Elizabeth Matka & Helen Sullivan. 2003. Evidence, Understanding and Complexity. *Evaluation* 9(3): 265-284.

Barnett, Michael & Finnemore, Martha. 1999. The politics, power, and pathologies of international organizations. *International Organization* 53(4): 699-732.

Baron, Gaëlle. 1999. *Evaluation, participation, apprentissage: Une conception de l'action public avec rationalité limité. Analyse et étude de cas d'une évaluation d'un programme européen.* Doctoral dissertation, Université de Rennes I. 440 p.

Baron, Gaëlle & Monnier, Eric. 2003. Une approche pluraliste et participative: Coproduire évaluation avec la société civile. *Revue « Informations sociales »,* no 10, septembre 2003.

Bartelmus, Peter. 2000. Sustainable development: paradigm or paranoia? *International Journal of Sustainable Development* 3(4): 358-369.

Bartelmus, Peter. 2003. Dematerialization and capital maintenance: two sides of the sustainability coin. *Ecological Economics* 46(1): 61-81.

Bazzoli, Laure. 1999. *L'économie politique de John R. Commons : Essai sur l'institutionnalisme en sciences sociales.* Paris et Montréal. L'Harmattan. Collection Etudes d'Economie Politique. 234 p.

Beck, Ulrich. 1992. *Risk Society: Towards a New Modernity*. London. Sage. 260 p.

Beck, Ulrich, Anthony Giddens & Scott Lash. 1994. *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Cambridge. Polity Press. 225 p.

Benton, Ted. 2002. Social Theory and Ecological Politics: Reflexive Modernization or Gree Socialism? In: Dunlap, R.E., F.H. Buttel, P. Dickens & A. Gijswijt (eds.), *Sociological Theory and the Environment: Classical Foundations, Contemporary Insights*. Lanham, Boulder, New York and Oxford. Rowman and Littlefield Publishers. Pp. 252-273.

Bernstein, David J. 1999. Comments on Perrin's "Effective Use and Misuse of Performance Measurement". *American Journal of Evaluation* 20(1): 85-93.

Beyeler, Michelle. 2002. *OECD Surveillance and Welfare States in Western Europe – Concepts and Methods of a Comparative Research Project*. Paper presented at COST A15 Second Conference 'Welfare Reforms for the 21st Century', April 5-6, Oslo.

BIAC. 2004. Consultation with the OECD Ad Hoc Group on Sustainable Development, 17 March 2004. Response from the Business and Industry Advisory Committee to the OECD, Room document 2. Paris.

Billig, Michael. 1987. *Arguing and Thinking: A Rhetorical Approach to Social Psychology*. Cambridge. Cambridge University Press.

Biott, Colin & Cook, Tina. 2000. Local Evaluation in a National Early Years Excellence Centres Pilot Programme: Integrating Performance Management and Participatory Evaluation. *Evaluation* 6(4): 399-413.

Birckmeyer, Johanna D. & Weiss, Carol H. 2000. Theory-based Evaluation in Practice: What Do We Learn? *Evaluation Review* 24(4): 407-431.

Bishop, R.C. 1978. Endangered species and uncertainty. The economics of a safe minimum standard. *American Journal of Agricultural Economics* 60(1): 10-18.

Blalock, Ann Bonar. 1999. Evaluation Research and the Performance Management Movement: From Estrangement to Useful Integration? *Evaluation* 5(2): 117-149.

Boaz, Annette & Hayden, Carol. 2002. Pro-active Evaluators: Enabling Research to Be Useful, Usable and Used. *Evaluation* 8(4): 440-453.

Bohman, James & Rehg, William (eds.). 1997. *Deliberative Democracy: Essays on Reason and Politics*. Cambridge, Massachusetts. MIT Press. 446 p.

Bourdieu, Pierre. 1986. Forms of Capital. In: Richardson, J.G. (ed.), *Handbook of Theory and Research for the Sociology of Education*. New York. Greenwood Press. Pp. 241-258.

Boven, Rick. 2003. *Choosing Your Future: The Role of Ideas in Managing Trade-Offs Between Economic and Environmental Objectives*. Ph.D. thesis. Department of Management Science and Information Systems, Faculty of Commerce, University of Auckland. 356 p.

Boyer, Robert & Saillard, Yves. 1995. *Théorie de la regulation: état des savoirs*. Paris. Edition la découverte. 568 p.

Bromley, Daniel. 1991. *Environment and Economy: Property Rights and Public Policy*. Oxford, UK and Cambridge, USA. Blackwell. 247 p.

Buchanan, James M. & Tullock, Gordon. 1962. *The Calculus of Consent: The Logical Foundations of Constitutional Democracy.* University of Michigan Press. http://www.econlib.org/library/Buchanan/buchCv3Contents.html (consulted 17 March 2005).

Bukkems, Gerard & de Groot, Hans. 2002. *Learning by auditing: mission impossible?* Paper presented at the Conference of the European Evaluation Society "Three Movements in Contemporary Evaluation: Learning, Theory and Evidence", Seville, 10-12 October 2002.

Burkett, Paul. 1999. *Marx and Nature: A Red and Green Perspective*. St. Martin's Press. New York. 312 p.

Buttel, 2000. Ecological Modernization as Social Theory. Geoforum 31(1): 57-65.

Bürgenmeier, Beat. 1999. Environment Protection: Towards a Socio-Economic Policy Design. In: O'Connor, M. & Spash, C. (eds.), *Valuation and the Environment: theory, method and practice*. Cheltenham, UK and Northampton, MA, USA. Edward Elgar. Pp. 73-86.

Bürgenmeier, Beat. 2005. *Développement durable et territoire: la dimension sociale comme fourre tout ?* Presentation at the European Summer School « Développement économique, environnement, espace », Université de Corse, 4-9 juillet 2005, Corté.

Caccia, Charles. 2003. The OECD and the World Economy. Doc. 9851 prov., 7 July. Provisional Report, to the Committee on Economic Affairs and Development of the Extended Parliamentary Asssembly of the Council of Europe. Rapporteur: Mr Caccia, Canada, Liberal Party.

Caldwell, Bruce J. 1982. *Beyond Positivism: Economic Methodology in the Twentieth Century*. London, Boston and Sydney. George Allen & Unwin. 277 p.

Campbell, Donald T. 1969. Reforms as experiments. *American psychologist* 24(4): 409-429.

Carroll, Michael C. & Stanfield, James R. 2003. Social Capital, Karl Polanyi, and American Social and Institutional Economics. *Journal of Economic Issues* 37(2): 397-404.

Chelimsky, Eleanor. 1997a. The coming transformations in evaluation. In: Chelimsky, E. & Shadish, W. R. (eds.), *Evaluation for the 21st century.* Thousand Oaks, California. Sage.

Chelimsky, E.leanor. 1997b. Thoughts for a new evaluation society. *Evaluation* 3(1): 97-109.

Chelimsky, Eleanor. 1998. The role of experience in formulating theories of evaluation practice. *American Journal of Evaluation* 19(1): 35-55.

Chen, Huey-tsyh. 1990. *Theory-driven evaluations.* Newbury Park, California. Sage. 326 p.

Chiesura, Anna & de Groot, Rudolf. 2003. Critical Natural Capital: a socio-cultural perspective. *Ecological Economics* 44(2-3): 219-231.

Ciriacy-Wantrup, S.V. 1952. *Resource Conservation: Economics and Policies.* Berkeley. University of California Press. 395 p.

Clark, William C. & Dickson, Nancy. 1999. The Global Environmental Assessment Project: Learning from efforts to link science and policy in an interdependent world. *Acclimations: Newsletter of the US National Assessment of the Potential Consequences of Climate Variability and Change*, no 8, pp. 6-7.

Coenen, Frans H.J.M., Doris A. Fuchs & Rob A. van de Peppel. 2000. *The environment and social well-being: An exploration of facts and figures and of possible relationships.* Centre for Clean Technology and Environmental Policy & Ministry of Housing, Spatial Planning and the Environment. Publicatiereeks milieustrategie (Environmental Strategy Publication Series) 2000/6.

Coglianese, Cary. 1999. The Limits of Consensus. The Environmental Protection System in Transition: Toward a More Desirable Future. *Environment* 41(3): 28-33.

Coleman, James S. 1988. Social Capital in the Creation of Human Capital. *American Journal of Sociology* 94 (suppl): S95-S120.

Commission of the European Communities. 2001. *A Sustainable Europe for a better world: a European Union Strategy for Sustainable Development.* Communication from the Commission (Commission's proposal to the Gothenburg European Council). COM(2001)264 final. http://europa.eu.int/eur-lex/en/com/cnc/2001/com2001_0264en01.pdf (consulted 30 August 2005)

Cook, Thomas D., Judith Levinson-Rose & William Pollard. 1980. The Misutilization of Evaluation Research: Some Pitfalls of Definition. *Knowledge: Creation, Diffusion, Utilisation* 1(4): 477-498.

Cousins, J. Bradley & Leithwood, Kenneth A. 1986. Current empirical research on evaluation utilization. *Review of Educational Research* 56(3): 331-364.

Cronbach, L.J., Ambron, S.R., Dornbusch, S.M., Hess, R.D., Hornik, R.C., Phillips, D.C., Walker, D.F. & Weiner, S.S. 1980. *Towards Reform of Program Evaluation: aims, methods and institutional arrangements.* San Francisco. Jossey-Bass Publishers. 432 p.

Dahler-Larsen, Peter. 2001. From Programme Theory to Constructivism: On Tragic, Magic and Competing Programmes. *Evaluation* 7(3): 331-349.

Daly, Herman E. 1992. *Steady-state Economics.* 2nd edition with new essays. London. Earthscan. 302 p.

Daly, Herman. 1996. *Beyond Growth: the Economics of Sustainable Development.* Boston. Beacon Press. 253 p.

Davies, Huw T.O. & Nutley, Sandra M. 2001. *Evidence-based policy and practice: moving from rhetoric to reality.* Paper presented at the Third International, Inter-disciplinary Evidence-Based Policies and Indicator Systems Conference, London, July 2001. http://cem.dur.ac.uk

Davies, Ian C. 1999. Evaluation and Performance Management in Government. Evaluation 5(2): 150-159.

Deblonde, Marian. 2000. Environmental Economics: The Meaning of an 'Objective' Policy Science. *Environmental Values* 9 (2000): 235-248.

DeGregori, Thomas R. & Shepherd, Deborah A. 1994. Technology, Theory of. In: Hodgson, G.M., W.J. Samuels & M.R. Tool (eds.), *The Elgar Companion to Institutional and Evolutionary Economics.* Volume II. Aldershot, England and Brookfield, Vermont. Edward Elgar. Pp. 320-325.

Denzau, Arthur T. & North, Douglass C. 1994. Shared Mental Models: Ideologies and Institutions. *Kyklos* 47(1): 3-31.

Dequech, David. 2002. The Demarcation between the "Old" and the "New" Institutional Economics: Recent Complications. *Journal of Economic Issues* XXXVI(2): 565-572.

Driessen, Peter P. J. & Glasbergen, Pieter (eds.). 2002. *Greening Society: The Paradigm Shift in Dutch Environmental Politics.* Dordrecht, Boston and London. Kluwer Academic Publishers.

Dryzek, John. 1990. Discursive Democracy: Politics, Policy and Political Science. Cambridge. Cambridge University Press. 244 p.

Dryzek, John S. 2000. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations.* Oxford. Oxford University Press. 195 p.

Dugger, William M. 1995. Beyond Technology to Democracy: The Tool Legacy in Instrumentalism. In: C.M.A. Clark (ed.), *Institutional Economics and the Theory of Social Value: Essays in Honor of Marc R. Tool.* Norwell, MA. Kluwer Academic Publishers. Pp. 195-207.

Eckersley, Robyn. 1992. *Environmentalism and Political Theory: Toward an Ecocentric Approach.* Albany. State University of New York Press. 274 p.

Eckley, Noelle. 2001. *Designing effective assessments : The role of participation, science and governance, and focus.* Report of a workshop co-organised by the European Environment Agency and the Global Environmental Assessment Project, Copenhagen, 1-3 March 2001. European Environment Agency Environmental issue report 26.

EEA. 2003. Europe's environment: the third assessment. Environmental assessment report no 10. European Environment Agency. Copenhagen. http://reports.eea.eu.int/environmental_assessment_report_2003_10/en/tab_content_RLR (consulted 30 August 2005)

Eggers, Helmut & Chelimsky, Eleanor. 1999. Purposes and Use: What Can We Expect? *Evaluation* 5(1): 92-96.

Eggertsson, Thráinn. 1991. *Economic Behavior and Institutions*. Cambridge. Cambridge University Press. 385 p.

Ekins, Paul. 1994. Growth, Limits to. In: Hodgson, G.M., W.J. Samuels & M.R. Tool (eds.), *The Elgar Companion to Institutional and Evolutionary Economics*. Volume I. Aldershot, England and Brookfield, Vermont. Edward Elgar. Pp. 289-297.

Ekins, Paul & Jacobs, Michael. 1995. Environmental Sustainability and the Growth of GDP: Conditions for Compatibility. In: Bhaskar, V. & Glyn, A. (eds). *The North, the South and the Environment: Ecological Constraints and the Global Economy.* London. Earthscan.

Ekins, Paul & Simon, Sandrine. 1999. The sustainability gap: a practical indicator of sustainability in the framework of national accounts. *International Journal of Sustainable Development* 2(1): 24-58.

Elster, Jon. (ed.). 1998a. *Deliberative Democracy*. Cambridge. Cambridge University Press. 282 p.

Elster, Jon. 1998b. Introduction. In: Elster, J. (ed.), *Deliberative Democracy*. Cambridge. Cambridge University Press. Pp. 1-18.

Englannin kieli ei tee ajatuksista tuoreita. Editorial, Helsingin Sanomat, 14 May 2003.

European Commission. 1997. *Evaluating EU Expenditure Programmes: A Guide to Ex Post and Intermediate Evaluation.* Luxembourg. European Commission. January.

Evans, Peter. 1995. *Embedded Autonomy: States and Industrial Transformation*. Princeton, NJ. Princeton University Press. 323 p.

Farrel, Alex & Hart, Maureen. 1998. What Does Sustainabiltiy Really Mean? The Search for Useful Indicators. *Environment*. 40(9): 4-9, 26-31.

Faucheux, Sylvie. 1997. Technological Change, Ecological Sustainability and Industrial Competitiveness. In: Dragun, A.K. & Jacobsson K.M. (eds.), *Sustainability and Global Environmental Policy: New Perspectives*. Cheltenham, UK and Lyme, US. Edward Elgar. Pp. 131-148.

Faucheux, Sylvie. 2000a. *Summary principles for sustainable development*. Article prepared for publication in the Encyclopedia of Life Support Systems (EOLSS), Unesco.

Faucheux, Sylvie. 2000b. *Sustainable Development of Natural Resource Capital*. Article prepared for publication in the Encyclopedia of Life Support Systems (EOLSS), Unesco.

Faucheux, Sylvie & Noël, Jean-François. 1995. *Economie des ressources naturelles et de l'environnement.* Paris. Armand Colin. 370 p.

Faucheux, Sylvie & O'Connor, Martin (eds.). 1997. *Valuation for Sustainable Development: Methods and Policy Indicators*. Cheltenham, UK and Lyme, US. Edward Elgar.

Faucheux, Sylvie, Eliot Muir & Martin O'Connor. 1997. Neoclassical theory of natural capital and 'weak' indicators for sustainability. *Land Economics* 73(4): 528-552.

Faucheux, Sylvie, Christelle Hue, Isabelle Nicolaï & Martin O'Connor. 2002a. *Integration of the Social Dimension of Sustainable Development in Enterprise Strategies within the Aluminium Industry*. Full Final Report (Phase One). Research Report prepared by the C3ED for a study supported by the European Aluminium Association through the EEA 'Aluminium for Future Generations' Programme, France, June.

Faucheux, Sylvie, Christelle Hue & Martin O'Connor. 2002b. *ITSAFE Project : Integrating Technological and Social Aspects of Foresight in Europe.* State of the Art Report: The 'Environment' Dimension. Report prepared for the European Commission. Final Report. December. C3ED, Université de Versailles Saint-Quentin-en-Yvelines. 64 p.

Feinstein, Osvaldo N. 2002. Use of Evaluations and Evaluation of their Use. *Evaluation* 8(4): 433-439.

Feller, Irwin. 2002. Performance Measurement Redux. American Journal of Evaluation 23(4): 435-452.

Fetterman, David .M. 1994. Empowerment evaluation. *Evaluation Practice* 15(1): 1-15.

Fetterman, David M. 1996. Empowerment evaluation: An Introduction to theory and practice. In: Fetterman, D.M, S. J. Kaftarian & A. Wandersman (eds). *Empowerment evaluation: Knowledge and tools for self-assessment & accountability.* Thousand Oaks. Sage. Pp 3–46.

Finnemore, Martha. 1993. International organizations as teachers of norms: the United Nations Educational, Scientific, and Cultural Organization and science policy. *International Organization* 47(4): 565-597.

Finnemore, Martha & Sikkink, Kathryn. 1998. International Norm Dynamics and Political Change. *International Organization* 52(4): 887-917.

Fishkin, James S. and Laslett, Peter (eds.). 2003. *Debating Deliberative Democracy.* Malden, MA, USA; Oxford, UK; Melbourne and Berlin. Blackwell. 234 p.

Flick, Uwe. 2002. *An Introduction to Qualitative Research*. 2nd edition. London, Thousand Oaks and New Delhi. Sage. 310 p.

Floc'hlay, Beatrice & Plottu, Eric. 1998. Democratic Evaluation: From Empowerment Evaluation to Public Decision-Making. Evaluation 4(3): 261-277.

Flyvbjerg, Bent. 2001. *Making Social Science Matter: Why social inquiry fails and how it can succeed again.* Cambridge, New York, Port Melbourne, Madrid and Cape Town. Cambridge University Press. 204 p.

Forester, John F. 1989. *Planning in the Face of Power.* Berkeley. University of California Press. 264 p.

Forss, Kim, Claus C. Reubien & Jerker Carlsson. 2002. Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback. *Evaluation* 8(1): 29-45.

Foster, Gladys Parker. 1995. In: Clark, C.M.A. (ed.), Institutional Economics and the Theory of Social Value: Essays in Honor of Marc R. Tool. Kluwer Academic Publishers. Boston, Dordrecht and London. Pp. 43-57.

Foucault, Michel. 1982. The Subject and Power. In: Faubion, J.D. (ed.), *Michel Foucault: Power. Essential Works of Foucault 1954-1984.* Volume 3. New York. The New Press. Pp. 326-348.

Foulkes, George. 1998. Evaluation and International Development: A British Perspective. *Evaluation* 4(3): 359-362.

Frey, Bruno.S. 1984. The public choice view of international political economy. *International Organization* 38(1): 199-223.

Froger, Géraldine. 2001. Gouvernance et nouvelles techniques de gestion des affaires communes : de quelles analyses dispose-t-on ? In: Froger, G., S. van den Hove, J. Haake & A. le Dars. *Gouvernance I: Gouvernance et développement durable.* Bâle, Genève et Munich. Helbing & Licthenhahn. Pp. 27-52.

Froger, Géraldine, Sybille van den Hove, Julia Haake & Aude le Dars. 2001. *Gouvernance I: Gouvernance et développement durable.* Bâle, Genève et Munich. Helbing & Licthenhahn 154 p.

Froger, Géraldine & Oberti, Pascal. 2002. Gouvernance et développement durable : L'aide multicritère à la décision participative. *Sciences de la société*, no 57, pp. 57-74.

Fukuyama, Francis. 1995. *Trust: The Social Virtues and the Creation of Prosperity.* New York. Free Press. 457 p.

Funtowicz, Silvio O. & Ravetz, Jerry .R. 1990. *Uncertainty and quality in science for policy.* Dordrecht: Kluwer Academic Publishers. 229 p.

Funtowicz, Silvio & Ravetz, Jerry. 1994. The worth of a songbird: Ecological economics as a post-normal science. *Ecological Economics* 10(3): 197-207.

Funtowicz, Silvio & Ravetz, Jerry. 2003. *Post-Normal Science.* International Society for Ecological Economics Internet Encyclopaedia of Ecological Economics. February 2003.

Funtowicz, Silvio O., Joan Martinez-Alier, Giuseppe Munda & Jerry R. Ravetz. 1999. *Information tools for environmental policy under conditions of complexity*. European Environment Agency, Environmental Issues Series, No 9. Luxembourg. Office for Official Publications of the European Communities. 34 p.

Gambetta, Diego. 1998. "Claro!" An Essay on Discursive Machismo. In: Elster, J. (ed.), *Deliberative Democracy*. Cambridge. Cambridge University Press. Pp. 19-43.

Gass, Ron. 2003. The battle for world progress: A strategic role for the OECD. *OECD Observer* no. 236, March. Pp. 29-31.

Gauchotte, Pierre. 1992. *Le pragmatisme*. Que sais-je ? Paris. Presses universitaires de France. 127 p.

Geva-May, Iris & Pal, Leslie A. 1999. Good Fences Make Good Neighbours: Policy Evaluation and Policy Analysis – Exploring the Differences. *Evaluation* 5(3): 259-277.

Giddens, Anthony. 1984. *The Constitution of Society: outline of the theory of structuration*. Cambridge. Polity Press. 402 p.

Giddens, Anthony. 1990. *The Consequences of Modernity*. Cambridge. Polity Press. 186 p.

Glasbergen, Pieter. 1997. Learning to Manage the Environment. In: Lafferty, W. & Meadowcroft, J. (eds.), *Democracy and the Environment*. Cheltenham and Lyme. Edward Elgar .Pp. 175-193.

Goodwin, Neva R. 2003. *Equity*. International Society for Ecological Economics Internet Encyclopaedia of Ecological Economics. February 2003. http://www.ecoeco.org/publica/encyc_entries/Equity.doc (consulted 28 April 2005)

Gowdy, John. 2003. *Contemporary Welfare Economics and Ecological Economics Valuation and Policy*. International Society for Ecological Economics Internet Encyclopaedia of Ecological Economics. February 2003. http://www.ecoeco.org/publica/encyc_entries/WelfeconEcolEcon.doc (consulted 15 May 2005)

Gowdy, John & McDaniel, Carl. 1999. The Physical Destruction of Nauru: An Example of Weak Sustainability. *Land Economics* 75(2): 333-338.

Graedel, Thomas E. & Klee, Robert J. 2002. Getting Serious about Sustainability. *Environmental Science & Technology* 36(4): 523-529.

Gray, Andrew. 2001. *Evidence-Based Policies and Indicator Systems: From Profane Arithmetic to a Sacred Geometry.* Paper presented Paper presented at the Third International, Inter-disciplinary Evidence-Based Policies and Indicator Systems Conference, London, July 2001.

Greenberg, George D., Jeffrey A. Miller, Lawrence B. Mohr & Bruce C. Vladeck. 1977. Developing Public Policy Theory: Perspectives from Empirical Research. *American Political Science Review* 71(4): 1532-1543.

Greene, Jennifer C. 1999. The Inequality of Performance Measurements. *Evaluation* 5(2): 160-172.

Greene, Jennifer C. 2002. With a Splash of Soda, Please: Towards Active Engagement with Difference. *Evaluation* 8(2): 259-266.

Greene, Jennifer C., Lehn Benjamin & Leslie Goodyear. 2001. The Merits of Mixing Methods in Evaluation. *Evaluation* 7 (1): 25-44.

Gregory, Amanda. 2000. Problematizing Participation: A Critical Review of Approaches to Participation in Evaluation Theory. *Evaluation* 6(2): 179-199.

De Groot, R.S., 1992. *Functions of Nature: evaluation of nature in environmental planning, management and decision-making*. Groningen. Wolters Noordhoff. 345 p.

Grootaert, Christiaan. 1998. *Social capital: the missing link?* The World Bank Social Capital Initiative, Working Paper No. 3.

Guba, Yvonna S. & Lincoln, Egon G. 1989. *Fourth Generation Evaluation*. Newbury Park, CA, London and New Delhi. Sage. 294 p.

Gunvaldsen, Jens A. & Karlsen, Rita. 1999. The Auditor as an Evaluator: How to Remain an Influential Force in the Political Landscape. *Evaluation* 5(4): 458-467.

Gutmann, Amy and Thompson, Dennis. 2003. Deliberative Democracy Beyond Process. In: Fishkin, J.S. & Laslett, P. (eds.), *Debating Deliberative Democracy*. Malden, MA, USA; Oxford, UK; Melbourne and Berlin. Blackwell. Pp. 31-53.

Haas, Peter M. 1992. Introduction: epistemic communities and international policy co-ordination. *International Organization* 46(1): 1-35.

Habermas, Jürgen. 1984. *The Theory of Communicative Action, Volume 1, Reason and the Rationalization of Society*. Boston. Beacon Press.

Hahtola, Kauko. 1990. Pragmatic-hermeneutical human action model for environmental planning. *Hallinnon tutkimus (Administrative Studies)* 9(4): 272-288.

Hahtola, Kauko. 1997. Tietävä, keskusteleva vai neuvotteleva tiede? (Knowing, discursive or negotiating science?) *Tiedepolitiikka* 2/97, pp. 49-51.

Hajer, Maarten A. 1995. *The Politics of Environmental Discourse: Ecological Modernization and the Policy Process*. Oxford. Clarendon Press. 344 p.

Hanberger, Anders. 2001a. What is the Policy Problem? Methodological Challenges in Policy Evaluation. *Evaluation* 7(1): 45-62.

Hanberger, Anders. 2001b. Policy and Program Evaluation, Civil Society, and Democracy. *American Journal of Evaluation* 22(2): 211-228.

Hanley, Nick, Shogren, Jason F. and White, Ben. 2001. *Introduction to Environmental Economics*. Oxford, NY. Oxford University Press. 350 p.

Hansson, Finn. 2000. *Research evaluation: Controlling or developing the production of knowledge. The peer review system in the changing knowledge production system.* Paper presented at Evaluation 2000, Conference of the American Evaluation Association, 1-5 November, Waikiki, Hawaii.

Harré, Rom. 1972. *The Philosophies of Science: An Introductory Survey*. London, Oxford and New York. Oxford University Press. 191 p.

Harribey, Jean-Marie. 1997. *L'économie économe: Le développement soutenable par la réduction du temps de travail.* Paris et Montreal. L'Harmattan. 298 p.

Harribey, Jean-Marie. 1998. *Le développement soutenable. Paris.* Economica. 111 p.

Harris, Jonathan M. 2000. *Basic Principles of Sustainable Development.* Global Development and Environment Institute, Tufts University, Working Paper 00-04. http://ase.tufts.edu/gdae (consulted 26 June 2005)

Hartwick, John M. 1977. Intergenerational equity and investing of rents from exhaustible resources. *American Economic Review* 67 (5): 972-974.

Heap, Shaun P. Hargreaves. 1994. Rationality and Maximization. In: Hodgson, G.M., W.J. Samuels, & M.R. Tool (eds.), *The Elgar Companion to Institutional and Evolutionary Economics*. Volume II. Aldershot, England and Brookfield, Vermont. Egar Elgar. Pp. 215-219.

Henry, Gary T. 2000. *Using Evaluation Findings for Policy: A Realist Perspective*. Paper presented at the European Evaluation Society Conference, Lausanne, 12-14 October 2000.

Henry, Gary T. & Mark, Melvin M. 2003. Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation* 24(3): 293-314.

Hilding-Rydevik, Tuija. 2001. A planning theory perspective on the EIA. In: Hilding-Rydevik, T. (ed.), *EIA, large development projects and decision-making in the Nordic countries.* Stockholm. Nordregio Report 2001:6.

Hindess, Barry. 1977. *Philosophy and Methodology in the Social Sciences*. Hassocks. Harvester Press. 258 p.

Hinterberger, Friedrich, Fred Luks, Marcus Stewen & Jan var der Straaten. 2000. Environmental policy in a complex world. *International Journal of Sustainable Development* 3(3): 276-296.

Hirsch, Fred. 1976. *Social Limits to Growth*. Cambridge. Harvard University Press (for the Twentieth Century Fund). 224 p.

Hisschemöller, Matthijs, Richard S.J. Tol & Pier Vellinga. 2001. The relevance of participatory approaches in integrated environmental assessment. *Integrated Assessment* 2(2): 57-72.

Hodgson, Geoffrey M. 1988. *Economics and Institutions: A Manifesto for a Modern Institutional Economics*. Cambridge and Oxford. Polity Press and Blackwell. 365 p.

Hodgson, Geoffrey M. 1999. *Evolution and Institutions: On evolutionary economics and the evolution of economics.* Cheltenham, UK and Northampton, MA. Edwar Elgar. 345 p.

Hodgson, Geoffrey M. 2002. Darwinism in economics: from analogy to ontology. *Journal of Evolutionary Economics* 12(3): 259-281.

Holling, C.S. 1973. Resilience and stability in ecological systems. *Annual Review of Ecology and Systematics* 4: 1-23.

Honkanen, Pertti. 1999. Marxin arvoteoria ja luonto (Marx's Theory of Value, and Nature). In: Heiskanen, J. (ed.), *Marx ja ekologia.* Kustannusyhtiö TA-Tieto Oy ja Demokraattinen Sivistysliitto ry. Pp. 93-110.

House, Ernest R. 1983. How We Think About Evaluation. In: House, E. R. (ed.), *Philosophy of Evaluation*. New Directions for Program Evaluation, no. 19. San Francisco. Jossey-Bass. Pp. 5-25.

House, Ernest R. 1993. *Professional Evaluation: Social Impact and Political Consequences*. Newbury Park, London and New Delhi. Sage. 215 p.

House, Ernest R. & Howe, Kenneth R. 1999. *Values in evaluation and social research*. Thousand Oaks, California. Sage.

Huber, Joseph. 1982. *Die verlorene Unschuld der Oekologie?* Frankfurt am Main. S.Fischer. 184 p.

Hukkinen, Janne. 1999. *Institutions of Environmental Management: Constructing mental models and sustainability.* London and New York. Routledge/EUI Environmental Policy Series. 226 p.

Hukkinen, Janne. 2001. Eco-efficiency as abandonment of nature. *Ecological Economics* 38(3): 311-315.

Hukkinen, Janne. 2003. From groundless universalism to grounded generalism: improving ecological economic indicators of human-environmental interaction. *Ecological Economics* 44(1): 11-27.

Ikenberry, John G. and Kupchan, Charles. 1990. Socialization and Hegemonic Power. *International Organization* 44(3): 283-315.

IMF:n ja OECD:n Suomi-kuvat ovat lähellä toisiaan. Helsingin Sanomat, 9.3. 2000.

Ingen risk för överhettning. Hufvusdstadsbladet, 5.8. 1998.

IUCN. 1980. *World Conservation Strategy: Living Resource Conservation for Sustainable Development*. International Union for Conservation of Nature and Natural Resources. Gland, Switzerland. 73 p.

Iyengar, Shanto, Mark .D. Peters & Donald R. Kinder. 1982. Experimental Demonstrations of the 'Not-So-Minimal' Consequences of Television News Programs. *The American Political Science Review* 76(4): 848-857.

Joas, Marko. 2001. *Reflexive Modernisation of the Environmental Administration in Finland: Essays of Institutional and Policy Change within the Finnish National and Local Environmental Administration*. Åbo. Åbo Akademi University Press. 271 p.

Johnson, Lyndon B. 1964. Great Society Speech. Public Papers of the Presidents of the United States, Lyndon B. Johnson, Book I (1963-64), p. 704-707. http://coursesa.matrix.msu.edu/~hst306/documents/great.html (consulted 21 February 2005).

Jones, Darryl. 2004. *Evaluating Agri-environment Policy in the OECD*. Paper presented at the OECD Workshop on Evaluating Agri-environmental Policies. Paris, 6-8 December 2004.

Julin, Jorma. 2003. The OECD: Securing the Future. *OECD Observer* No. 240/241, December.

Julnes, George, Melvin Mark & Gary T. Henry. 1998. Promoting Realism in Evaluation: Realistic Evaluation and the Broader Context. *Evaluation* 4(4): 483-504.

Jänicke, Martin. 1988. Ökologische Modernisierung. Optionen und Restriktionen präventiver Umweltpolitik. In: Simonis, U. (ed.), *Präventive Umweltpolitik*. Frankfurt am Main. Campus. Pp. 13-26.

Jänicke, Martin. 2000. *Ecological Modernization: Innovation and Diffusion of Policy and Technology*. Discussion Paper FFU – 00-08 - 2000. Berlin. Freie Universität Berlin, Forschungsstelle fûr Umweltpolitik.

Kallis, G. 2001. *A Co-evolutionary Institutional Approach to Sustainable Development: Theory and Potential Applications*. Paper presented at 'Frontiers 1' European Society for Ecological Economics Conference, Cambridge, 4-7 July 2001.

Kapp, William K. 1950. *The Social Costs of Private Enterprise.* Cambridge. Harvard University Press. 287 p.

Kazi, Mansoor A.F. & Spurling, Lucy J. 2000. *Realist Evaluation for Evidence-Based Practice*. Paper presented at Fourth Conference of the European Evaluation Society, Lausanne, Switzerland, October 2000.

Keppler, Jan H. 1998. The Genesis of 'Positive Economics' and the Rejection of Monopolistic Competition Theory: A Methodological Debate. *Cambridge Journal of Economics* 22(3): 261-276.

Khakee, Abdul. 2003. The Emerging Gap between Evaluation Research and Practice. *Evaluation* 9(3): 340-352.

Klein, Philip A. 1987. Power and Economic Performance: The Institutionalist View. *Journal of Economic Issues* 21(3): 1341-1377.

Knott, Jack & Wildavsky, Aaron. 1980. If Dissemination is the Solution, What Is the Problem? *Knowledge: Creation, Diffusion, Utilisation* 1(4): 537-578.

Koestler, Arthur. 1967. *The ghost in the machine*. London. Hutchinson. 384 p.

Kozak, David C. & Keagle, James G. (eds.). 1988. *Bureaucratic Politics and National Security: Theory and Practice.* Boulder, CO, and London. Lynne Rienner Publishers. 504 p.

Krosnick, Jon A. & Brannon, Laura A. 1993. The Impact of Gulf War on the Ingredients of Presidential Evaluations: Multidimensional Effects of Political Involvement. *American Political Science Review* 87(4): 963-975.

Krosnick, Jon A. & Kinder, Donald R. 1989. Altering the Foundations of Support for the President Through Priming. *American Political Science Review* 84(2): 497-512.

Kushner, Saville. 2002a. *The Object of One's Passion: Engagement and Community in Democratic Evaluation.* Keynote Address presented at the 2002 Australasian Evaluation Society International Conference, October/November 2002, Wollongong, Australia. Http://www.aes.asn.au

Kushner, Saville. 2002b. I'll Take Mine Neat: Multiple Methods but a Single Methodology. *Evaluation* 8(2): 249-258.

Lafferty, William M. & Hovden, Eivind. 2002. *Environmental Policy Integration: Towards an Analytical Framework?* Program for Research and Documentation for a Sustainable Society (ProSus), Centre for Development and the Environment, University of Oslo. Report No. 7/02.

Lampinen, Osmo. 1992. *The utilization of social science research in public policy.* Doctoral Dissertation. Suomen Akatemian julkaisuja 4. Helsinki. VAPK-kustannus. 294 p.

Lang, Jochen. 2001. Improving Structural Policy Under Conditions of Hybrid Governance: Multi-Actor Constellations, Integrated Feedback Instruments and the Use of Evaluation Results. *Evaluation* 7(1): 7-23.

Langhelle, Oluf. 1998. *Ecological Modernisation and Sustainable Development - differences and similarities.* Paper prepared for the workshop "Ecological Modernisation" at the University of Helsinki, 10-12 September 1998. 25 p.

Lascoumes, Pierre. 2001. Les ambiguities des politiques de développement durable. In: Université de tous les savoirs, tome 5. Editions Odile Jacob.

Latouche, Serge. 1984. *Le procès de la science sociale : introduction à une théorie critique de la connaissance*. Paris. Editions Anthropos. 219 p.

Latouche, Serge. 1989. *L'occidentalisation du monde: Essai sur la signification, la portée et les limites de l'uniformisation planétaire*. Paris. Editions la Découverte. 143 p.

Laukkanen, Reijo. 1998. *Opestustoimen keskushallinnon evaluaatioajattelun kehitys Suomessa 1970-luvulta 1990-luvulle (The Development of the Evaluation Thinking in the National Administration for Education in Finland from the 1970s to the 1990s).* Doctoral dissertation. University of Jyväskylä, koulutuksen tutkimuslaitos (Department for Education Research), tutkimuksia 5. Jyväskylä. Jyväskylän yliopistopaino. 178 p.

Lavoux, Thierry. 2003. Comment mesurer le développement durable? *Découverte – Revue du Palais de la découverte*, no 310 (juillet-août-septembre), pp. 50-59.

Lawrence, David P. 2000. Planning theories and environmental impact assessment. *Environmental Impact Assessment Review* 20(6): 607-625.

Le Bot, Jean-Michel. 2002. Du développement durable au bien public: Essai anthropologique sur l'environnement et l'économie. Paris. L'Harmattan. 297 p.

Leca, Jean. 1997. *L'Evaluation comme intervention, sciences sociales et pratiques administratives et politiques*. Paper presented at the GAPP-CNRS-ENS Seminar, Cachan, France, 5-6 February 1997.

Leeuw, Frans L. 2002. Evaluation in Europe 2000: Challenges to a Growth Industry. *Evaluation* 8(1): 5-12.

Leeuw, Frans L., Toulemonde, Jacques & Brouwers, André. 1999. Evaluation Activities in Europe: A Quick Scan of the Market in 1998. *Evaluation* 5(4): 487-496.

Leknes, Einar. 2001. The roles of EIA in the decision-making process. *Environmental Impact Assessment Review* 21(4): 309-334.

Les institutions internationals s'ouvrent sous la pression de la société civile. 2001. Le Monde, 23 January 2001. Journalist: Martine Laronche.

Leskinen, Antti. 1994. *Environmental Planning as Learning: The Principles of Negotiation, the Disaggregative Decision-making Method and Parallel Organization in Developing the Road Administration.* University of Helsinki, Department of Economics and Management, Publications No. 5, Land Use Economics. Helsinki. 162 p.

Leviton, L. & Hughes, E. 1981. Research on the utilization of evaluations. *Evaluation Review* 4(5): 525-548. (p. 102)

Liebowitz, Stan J. & Margolis, Stephen E. 1998. Path dependence. In: *The New Palgrave's Dictionary of Economics and the Law*. MacMillan. http://www.utdallas.edu/~liebowit/

Lorentsen, Lorents. 2004. Presentation of the Director of the OECD Environment Directorate, Lorents Lorentsen at the OECD Conference on Material Flows and Related Indicators: "OECD Work on Environment & Sustainable Development", Helsinki 17-18 June 2004.

Louhimaa, Eila. 2002. *Luonnon sosiaalinen konstruointi, ympäristödiskurssit ja ympäristöön orientoiva kasvatus. Tutkimus institutionaalisen ympäristökasvatuksen institutionaalisista rakenne-ehdoista ja kulttuuristen mahdollisuuksien kentistä. (The social construction of nature, environmental discourses, and environmentally oriented education. A study of structural conditions and cultural possibilities of environmental education)*.Doctoral dissertation. Department of Educational Sciences and Teacher Education; the Unit of Sociology, Women's Studies, and Environmental Education, University of Oulu, Finland. 296 p. (abstract in English)

Lukes, Steven. 1982. Relativism in its Place. In: Lukes, S. & Hollis, M. (eds.), *Rationality and Relativism*. Oxford. Blackwell. Pp. 261-310.

Luks, Fred. 1998. The rhetorics of ecological economics. *Ecological Economics* 26(2): 139-149.

Lundqvist, Lennart J. 2000. *The International Spread of Ecological Modernisation Ideas*. Paper presented at the International Workshop on Diffusion of Environmental Policy Innovations, Berlin 8-9 December 2000.

Lyons, Michal, Carin Smuts & Andrea Stephens. 2001. Participation, Empowerment and Sustainability. *Urban Studies* 38(8): 1233-1251.

Mabry, Linda. 2002. Postmodern Evaluation – or Not? *American Journal of Evaluation* 23(2): 141-157.

March, James G. & Olsen, Johan P. 1998. The Institutional Dynamics of International Political Orders. *International Organisation* 52(4): 943-969.

Marcussen, Martin. 2001. *The OECD: Transnational Regulation Through the Idea-Game*. Paper prepared for the SCANCOR Workshop, "Transnational Regulation and the Transformation of States", Stanford University, 22-23 June 2001.

Maréchal, Jean-Paul. 2000. *Humaniser l'économie.* Paris. Desclée de Brouwer. 226 p.

Martin, Steve & Sanderson, Ian. 1999. Evaluating Public Policy Experiments: Measuring Outcomes, Monitoring Processes or Managing Pilots? *Evaluation* 5(3): 245-258.

Martinez-Alier, Joan, Giuseppe Munda & John O'Neill. 1998. Weak comparability of values as a foundation for ecological economics. *Ecological Economics* 26(3): 277-286.

Martinez-Alier, Joan, Giuseppe Munda & John O'Neill. 1999. Commensurability and Compensability in Ecological Economics. In: O'Connor, M. & Spash, C. (eds.), *Valuation and the Environment: Theory, Method and Practice.* Cheltenham, UK and Northampton, MA. Edward Elgar. Pp. 37-58.

McCloskey, Deirdre N. 1998. *The Rhetoric of Economics.* 2nd edition. Madison. Madison University Press. 223 p.

McCombs, Maxwell E. & Shaw, Donald L. 1972. The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly* 36(2): 176-185.

McKechnie, Rosemary B. & Welsh, Ian. 2002. When the Global Meets the Local: Critical Reflections on Reflexive Modernisation. In: Dunlap, R.E., F.H. Buttel, P. Dickens & A. Gijswijt (eds.), *Sociological Theory and the Environment: Classical Foundations, Contemporary Insights.* Lanham, Boulder, New York and Oxford. Rowman and Littlefield Publishers. Pp. 286-310.

McKie, Linda. 2003. Rhetorical Spaces: Participation and Pragmatism in the Evaluation of Community Health Work. *Evaluation* 9(3): 307-324.

Meadows, Donella H., Dennis L. Meadows, Jorgen Randers & WilliamW. Behrens III. 1972. *The Limits to Growth: a report for the Club of Rome's Project on the Predicament of Mankind.* London. Earth Island Ld. 205 p.

Mebratu, Desta. 1998. Sustainability and Sustainable Development: Historical and Conceptual Review. *Environmental Impact Assessment Review* 8(6): 493-520.

Mertens, Donna M. 1999. Inclusive Evaluation: Implications of Transformative Theory for Evaluation. *American Journal of Evaluation* 20(1): 1-14.

Mickwitz, Per. 2002. Effectiveness Evaluation of Environmental Policy: the Role of Intervention Theories. *Hallinnon tutkimus (Administrative Studies)* 21(4): 77-87.

Mickwitz, Per. 2003..A Framework for Evaluating Environmental Policy Instruments. *Evaluation* 9(4): 415-436.

Mickwitz, Per & Kivimaa, Paula. 2004. *Evaluating Policy Integration – The Case of Policies for Environmentally Friendlier Technological Innovations*. Paper presented at the Sixth Biennial Conference of the European Evaluation Society, "Governance, Democracy and Evaluation", Berlin, 30 September – 2 October 2004.

Mintrom, Michael. 1997. Policy Entrepreneurs and the Diffusion of Innovation. *American Journal of Political Science* 41(3): 738-770.

Mirowski, Philip. 1987. The Philosophical Basis of Institutionalist Economics. *Journal of Economic Issues* 21(3): 1001-1038.

Mishan, Edward J. 1967. *The Costs of Economic Growth*. London. Staples Press. 190 p.

Mishan, Edward J. 1977. *The Economic Growth Debate: an Assessment.* London. George Allen and Unwin. 277 p.

Mol, Arthur P.J. & Sonnenfeld, David A. 2000. Ecological Modernisation Around the World: An Introduction. *Environmental Politics* 9(1): 3-16. Available at: http://www.sls.wau.nl/es/ecomod_intro.htm (consulted 5 May 2005).

Mol, Arthur P.J. & Spaargaren, Gert. 2000. Ecological Modernisation Theory in Debate: A Review. *Environmental Politics* 9 (1): 17-49.

Montague, Steve & Allderings, Erwin. 2004. *Building Accountability Structures into Agri-environmental Policy Development.* Presentation at the OECD Workshop on Evaluating Agri-Environmental Policies, 6-8 December 2004, Paris.

Morabito, Stephen M. 2002. Evaluator Roles and Strategies for Expanding Evaluation Process Influence. *American Journal of Evaluation* 23(3): 321-330.

Muller, Pierre & Surel, Yves. 1998. L'Analyse des politiques publiques. Clefs/Politique. Montchrestien. Paris. 156 p.

Munda, Giuseppe. 1997. Environmental Economics, Ecological Economics, and the Concept of Sustainable Development. *Environmental Values* 6(2): 213-233

Munda, Giuseppe. 2004. 'Social multi-criteria evaluation (SMCE)': methodological foundations and operational consequences. European Journal of Operational Research 158(3): 662-677.

Myrdal, Gunnar. 1973. *Against the Stream: Critical Essays on Economics*. New York. Pantheon Books. 336 p.

Myrdal, Gunnar. 1978. Institutional Economics. *Journal of Economic Issues* 12(4): 771-783.

Nelson, Julie A. 2003. Confronting the science/value split: notes on feminist economics, institutionalism, pragmatism and process thought. *Cambridge Journal of Economics* 27(1): 49-64.

Norgaard, Richard B. 1989. The Case for Methodological Pluralism. *Ecological Economics* 1(1): 37-58.

Norgaard, Richard B. 1994. *Development Betrayed: the end of progress and a coevolutionary revisioning of the future.* London and New York. Routledge. 280 p.

North, Douglass C. 1990. Institutions, institutional change and economic performance. Cambridge. Cambridge University Press. 152 p.

North, Douglass C. 1996a. Institutions, Organizations, and Market Competition. Available in www.format: ftp://econwpa.wustl.edu/econ-wp/eh/papers/9612/9612005.html (Consulted 29 July 2005)

North, Douglass C. 1996b. Economic Performance Through Time: The Limits to Knowledge. Available in www format: ftp://econwpa.wustl.edu/econ-wp/eh/papers/9612/9612004.html (Consulted 29 July 2005)

Nussbaum, Martha C. 2000. *Women and Human Development: The Capabilities Approach.* Cambridge. Cambridge University Press. 312 p.

"L'OCDE est un piston de la mondialisation". Libération, le 7 août 2001.

O'Connor, Martin. 1998a. *Walking in the Garden(s) of Babylon: An overview of the VALSE Project.* C3ED Rapport de Recherche. July. 3E Press. Université de Versailles Saint-Quentin-en-Yvelines. 36 p.

O'Connor, Martin. 1998b. Ecological-Economic Sustainability. In: Faucheux, S. & O'Connor, M. (eds.), *Valuation for Sustainable Development: methods and policy indicators.* Cheltenham, UK and Northampton, MA. Edward Elgar. Pp.19-42

O'Connor, Martin. 1999. Dialogue and debate in a post-normal practice of science: a reflection. *Futures* 31(7): 671-687.

O'Connor, Martin. 2000a. The VALSE project – an introduction. *Ecological Economics* 34(2): 165-174.

O'Connor, Martin. 2000b. *Natural Capital*. Environmental Valuation in Europe Policy Research Brief no. 3. Cambridge Research for the Environment. 24 p.

O'Connor, Martin. 2000c. *"Our Common Problems": ICT, the Prisoner's Dilemma, and the Process of Working out Reasonable Solutions to Impossible Environmental Problems.* Cahiers du C3ED No 00-06, Juin 2000. 24 p.

O'Connor, Martin & Muir, Eliot. 1995. Endowment Effects in Competitive General Equilibrium: A Primer for Paretian Policy Analysts. *Journal of Income Distribution* 5(2): 145-175.

O'Connor, Martin & Spash, Clive. 1999. Introduction. In: O'Connor, M. & Spash, C. (eds.), *Valuation and the Environment: theory, method and practice.* Cheltenham, UK and Northampton, MA. Edward Elgar. Pp. 1-36.

O'Connor, Martin, Anton Steurer & Marialuisa Tamborra. 2001. *Greening National Accounts*. Environmental Valuation in Europe Policy Research Brief no. 9. Cambridge Research for the Environment. 24 p.

O'Connor, Martin, Sybille van Hove, John O'Neill & Christelle Hue. 2001/2002. *Le Modèle Politique de la Démocratie Délibérative.* C3ED, Département de Sciences Economiques et Management. Université de Versailles Saint-Quentin-en-Yvelines. EEE-SDP Module "Gouvernance". 50 p.

O'Connor, Martin & Spangenberg, Joachim. 2004. *A Methodology for CSR Reporting: Assuring a Representative Diversity of Indicators Across Stakeholders, Scales, Sites and Performance Issues.* Cahiers du C3ED, no. 04-02. Université de Versailles Saint-Quentin-en-Yvelines, France. 27 p.

OECD. 1996. *Environmental Performance in OECD Countries: Progress in the 1990s*. Paris.

OECD. 1997a. *OECD Environmental Performance Reviews: A Practical Introduction*. OECD Environment Monograph. OCDE/GD(97)35. Paris. 60 p.

OECD. 1997b. *Workshop on the Second Cycle of Reviews/ Lessons from the first cycle of reviews* (document prepared by the Finnish Delegation). OECD Environment Directorate/Enviornmental Policy Committee, Group on Environmental Performance. Paris, 17 November 1997. Room document No. 12.

OECD. 1997c. *Guiding the Transition to Sustainable Development: A Critical Role for the OECD*. OECD High-Level Advisory Group on the Environment. Paris.

OECD. 1997d. *Evaluating Economic Instruments for Environmental Policy*. Paris. 141 p.

OECD. 1998a. *OECD Environmental Performance Reviews: Second Cycle Work Plan*. EPOC Document ENV/EPOC (98)21. Paris. 16 p.

OECD. 1998b. *Water Management in OECD Countries*. Paris. 37 p.

OECD. 1999. *Improving Evaluation Practices: Best Practice Guidelines for Evaluation and Background Paper*. Public Management Service, Public Management Committee. PUMA/PAC(99)1. Paris, 22 January 1999.

OECD. 2000a. *Environmental performance reviews*. Presentation by Christian Avérous (OECD Environment Directorate) at the Special session of the Working Party on Environmental Performance, 7 November 2000.

OECD. 2000b. *Resolution of the Council Concerning the Mandate of the Public Management Committee.* C(99)175/FINAL. Paris, 14 January.

OCDE. 2001a. *Du bien-être des nations: le rôle du capital humain et social.* Paris. 136 p.

OECD. 2001b. *The Concept of Socially-Sustainable Development: Review of Literature and Preliminary Conclusions* (SG/SD(2001)13). Paris.

OECD. 2001c. *The Concept of Socially-Sustainable Development: A Survey.* (SG/SD(2001)13/ANN). Paris, 22 October.

OECD. 2001d. Environmental Performance Reviews: Achievements in OECD Countries. Paris. 126 p.

OECD. 2001e. Policies to Enhance Sustainable Development. Paris. 106 p.

OECD. 2001f. Sustainable Development: Critical Issues. Paris. 487 p.

OECD. 2001g. Towards a Sustainable Future. Press release from the OECD Council at Ministerial level. 17 May 2001. http://www.oecd.org/document/25/0,2340,fr_2649_37361_1914137_1_1_1_1,00.html (consulted 31 August 2005)

OECD 2001h. *Aggregated Environmental Indices: Review of Aggregation Methodologies in use.* ENV/EPOC/SE(2001)2/ FINAL. OECD Paris. 43 p.

OECD. 2001i. *Civil Society and the OECD*. OECD Policy Brief, October. Available at: http://www.oecd.org/dataoecd/36/9/2367378.pdf (Consulted 6 May 2005).

OECD. 2001j. *Monitoring Progress with the OECD Environmental Strategy: Environmental Performance Reviews and Environmental Indicators*. Working Party on Environmental Performance. ENV/EPOC/GEP(2001)8. Paris. 8 p. + 3 annexes.

OECD. 2002a. *Governance for Sustainable Development: Five OECD Case Studies*. OECD. Paris. 35 p.

OECD. 2002b. *OECD Work on Sustainable Development in Response to the 2001 Ministerial Mandate: A Report of Progress to the 2002 Ministerial Council Meeting.* The Ad Hoc Group on Sustainable Development. General Secretariat. SG/SD(2002)6. Paris. 9 p.

OECD. 2002c. *Environmental Performance Reviews – Second Cycle: Review of Progress*. 21 December. ENV/EPOC/GEP(2002)4. 17 p.

OECD. 2002d. *Overview of Sustainable Development Indicators used by National and International Agencies.* Ad Hoc Group on Sustainable Development. 9 April. SG/SD(2002)5. Paris. 84 p.

OECD. 2003a. *Maximising the Impact of the OECD: Reorganising the Committee Structure and Related Matters*. Report by Mr. Peter Nicholson. SG(2003)1. Paris. 111 p.

OECD. 2003b. *OECD Priorities in the Medium Term: Report on a Survey of Members' Views*. Prepared by Peter Nicholson, Special Advisor to the Secretary-General. July 2003. Paris.

OECD. 2003c. *Sustaining Whose Development? Analysing The International Effect of National Policies*. A Project Of The Round Table On Sustainable Development. September 2003. Paris.

OECD. 2003d. *Scoping Paper on Possible Future Work on Sustainable Development: Ad Hoc Group on Sustainable Development.* SG/SD(2003)10. General Secretariat. 1 October 2003. Paris. 21 p.

OECD. 2003e. *OECD Environmental Indicators: Development, Measurement and Use.* Reference Paper. OECD Environment Directorate. 30 p.

OECD. 2003f. *Water: Performance and Challenges in OECD Countries.*

OECD. 2003g. *OECD Environmental Performance Reviews: Austria.* Paris. 205 p.

OECD. 2003h. *Future Direction of the OECD: Report on OECD's Role in Global Architecture.* Report prepared by Ambassador Jorma Julin. HOD(2003)2. 24 March 2003. Paris.

OECD. 2003i. *Second Cycle of the OECD Environmental Performance Reviews: Mid-term Review.* ENV/EPOC(2003)30. 28 October 2003. Paris. 9 p.

OECD. 2003j. *Peer Review: An OECD Tool for Co-operation and Change*. Paris. 54 p. (Also contains a French version: *L'examen par les pairs: un instrument de l'OCDE pour la coopération et le changement*. 61 p.)

OECD. 2004a. *Draft Final Report to Ministers: Ad Hoc Group on Sustainable Development.* OECD General Secretariat. SG/SD(2004)4/REV1. 8 April. Paris. 57 p.

OECD. 2004b. *Getting to Grips with Globalisation: The OECD in a Changing World*. OECD. Paris. 39 p.

OECD. 2004c. *OECD Environmental Performance Reviews: Sweden.* Paris. 210 p.

OECD. 2004d. *OECD Environmental Performance Reviews: Spain*. Paris. 211 p.

OECD. 2004e. *Measuring Sustainable Development: Integrated Economic, Environmental and Social Frameworks.* Paris. 426 p.

OECD. 2004f. *OECD Environmental Performance Reviews: Canada*. Paris. 228 p.

OECD. 2005a. *OECD Environmental Performance Reviews – Beyond the Second Cycle*. Working Party on Environmental Performance. ENV/EPOC/GEP(2005)3. 4 May. Paris. 20 p.

OECD. 2005b. *OECD Environmental Performance Reviews: France*. Paris. 248 p.

OECD. 2005c. *OECD Environmental Performance Reviews: The United States*. Paris. 284 p.

OECD ärsyttää yhä enemmän poliittisiin kiistoihin. 1998. Helsingin Sanomat, 6 August.

Olson, Mancur. 1965. *The Logic of Collective Action*; Cambridge MA. Harvard University Press. 176 p.

O'Hara, Sabine. 1996. Discursive ethics in ecosystems valuation and environmental policy. *Ecological Economics* 16(2): 95-107.

O'Neill, John. 1996. Cost-Benefit Analysis, Rationality and the Plurality of Values. *The Ecologist* 26(3): 98-103.

O'Neill, John. 1997. Managing Without Prices: The Monetary Valuation of Biodiversity. *Ambio* 26(8): 546-550.

O'Neill, John. 2001. *Representing People, Representing Nature, Representing the World*. Paper presented at the European Society for Ecological Economics 'Frontiers 1' Conference, Cambridge, 4-7 July 2001.

Opschoor, Hans & van der Straaten, Jan. 1993. Sustainable development: and institutional approach. *Ecological Economics* 7(3): 203-222.

Owen, John M. 2002. *Organisational Learning as an Accountability Commitment: Lessons from Innovative Practice*. Abstract. Paper presented at the Conference of the European Evaluation Society "Three Movements in Contemporary Evaluation: Learning, Theory and Evidence", Seville, 10-12 October 2002.

Owens, Susan & Rayner, Tim. 1999. 'When Knowledge Matters': The Role and Influence of the Royal Commission on Environmental Pollution. *Journal of Environmental Policy and Planning*, 1(1): 7-24.

Palerm, Juan R. 2000. An Empirical-Theoretical Analysis Framework for Public Participation in Environmental Impact Assessment. *Journal of Environmental Planning and Management* 43(5): 581-600.

Parkinson, John. 2001. *Legitimacy and Deliberative Democracy*. Paper presented at the European Society for Ecological Economics 'Frontiers 1' Conference, Cambridge, 4-7 July 2001.

Parlett, Malcolm & Hamilton, David. 1977. Evaluation as illumination: a new approach to the study of innovatory programmes. In: Hamilton, D., D. Jenkins, C. King, B. MacDonald, M. Parlett (eds.), *Beyond the Numbers Game: a Reader in Educational Evaluation*. London. Macmillan. Pp. 6-22.

Parson, Edward Anthony. 1995. Integrated assessment and environmental policy making:: in Pursuit of Usefulness. *Energy Policy* 23(4-5): 463-475.

Passet, René. 1996. *L'Economique et le vivant*. 2. édition. Paris. Economica. 291 p.

Patton, Michael Q. 1997a. *Utilization-Focused Evaluation: The New Century Text*. 3rd edition. Thousand Oaks, London and New Delhi. Sage. 431 p.

Patton, Michael Q. 1997b. Toward Distinguishing Empowerment Evaluation and Placing It In A Larger Context. *Evaluation Practice* 18(2): 147-163.

Patton, Michael Q. 1998. Discovering Process Use. *Evaluation* 4(2): 225-233.

Pawson, Ray. 2002a. Evidence-based Policy: In Search of a Method. *Evaluation* 8(2): 157-181.

Pawson, Ray. 2002b. Evidence-based Policy: The Promise of 'Realist Synthesis'. *Evaluation* 8(3): 340-358.

Pawson, Ray & Tilley, Nick. 1997. *Realistic Evaluation*. London, Thousand Oaks and New Delhi. Sage. 235 p.

Pearce, David. 1998. *Economics and Environment: Essays on Ecological Economics and Sustainable Development*. Cheltenham, UK and Northampton, USA. Edward Elgar. 363 p.

Pearce, David W. & Turner, R. Kerry. 1990. *Economics of natural resources and the environment*. New York, London, Toronto, Sydney, Tokyo and Singapore. Harvester Wheatsheaf. 378 p.

Pearce, David W., Anil Markandya & Edouard Barbier. 1989. *Blueprint for a Green Economy*. London. Earthscan. 192 p.

Peled, Enat & Spiro, Shimon E. 1998. Goal Focused Evaluation: Lessons from a Study of a Shelter for Homeless Youth. *Evaluation* 4(4): 455-468.

Perrin, Burt. 1998. Effective Use and Misuse of Performance Measurement. *American Journal of Evaluation* 19(3): 367-379.

Perrin, Burt. 1999. Performance Measurement: Does the Reality Match the Rhetoric? A Rejoinder to Bernstein and Winston. *American Journal of Evaluation* 20(1): 101-111.

Perrin, Burt. 2002a. How to – and How Not to – Evaluate Innovation. *Evaluation* 8(1): 13-28.

Perrin, Burt. 2002b. *Towards a New View of Accountability*. Paper presented at the Conference of the European Evaluation Society "Three Movements in Contemporary Evaluation: Learning, Theory and Evidence", Seville, 10-12 October 2002.

Petit, Olivier. 2002. *De la coordination des actions individuelles aux formes de l'action collective : une exploration des modes de gouvernance des eaux souterraines*. Doctoral thesis. C3ED, UFR des Sciences Sociales et des Humanités, Université de Versailles Saint-Quentin-en-Yvelines, France. 410 p.

Pezzey, John. 1997. Sustainability constraints versus optimality versus intertemporal concern, and axioms versus data. *Land Economics* 73(4): 448-466.

Pitarelli, Emilio & Monnier, Eric. 2000. *Benchmarking: the mission link between evaluation and management?* Paper presented at the European Evaluation Society Fourth Conference 'Taking Evaluation to the People: Between Civil Society, Public Management and the Polity'. Lausanne, 12-14 October.

Pollit, Christopher. 1998. Evaluation in Europe: Boom or Bubble? *Evaluation* 4(2): 214-224.

Popper, Karl R. 1945. *The Open Society and its Enemies*. London, Henley. Routledge & Kegan Paul. 268 p.

Power, Michael. 1997. *The Audit Society: Rituals of Verification*. Oxford. Clarendon Press. 183 p.

Przeworski, Adam. 1998. `Deliberation and Ideological Domination. In: Elster, J. (ed.), *Deliberative Democracy*. Cambridge. Cambridge University Press. Pp. 140-160.

Putnam, Robert (with Robert Leonardi & Raffaella Nanetti). 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton. Princeton University Press. 258 p.

Putnam, Robert. 2000. *Bowling Alone : The Collapse and Revival of American Community*. New York, London, Toronto. Simon and Schuster. 541 p.

Rajavaara, Marketta. 1999. Arviointitutkimuksen hyödynnettävyys (Usability of evaluation research). In: Eräsaari, R., T. Lindqvist, M. Mäntysaari & M. Rajavaara (eds.), *Arviointi ja asiantuntijuus (Evaluation and expertise)*. Helsinki. Gaudeamus. Pp. 31-53.

Rammel, Christian & van den Bergh, Jeroen C.J.M. 2003. Evolutionary policies for sustainable development: adaptive flexibility and risk minimising. *Ecological Economics* 47(2-3): 121-133.

Reinalda, Bob. 2001. *Decision Making within International Organisations: An Overview of Approaches and Case Studies*. Paper presented at the European Consortium for Political Research (ECPR), 29th Joint Sessions of Workshops, Grenoble, France, 6-11 April 2001. Draft.

Rist, Gilbert. 2001. *Le Développement : Histoire d'une croyance occidentale*. Deuxième edition mise à jour et augmentée d'une postface. Paris, Presses de Sciences Po. 442 p.

Robeyns, Ingrid. 2000. *An unworkable idea or a promising alternative? Sen's capability approach re-examined*. Center for Economic Studies Discussion Paper DPS 00.30. Katholieke Universiteit Leuwen. http://www.econ.kuleuven.ac.be/eng/ew/discussionpapers/Dps00/DPS0030.pdf

Robeyns, Ingrid. 2005. The Capability Approach: a theoretical survey. *The Journal of Human Development* 6(1): 93-117.

Roman, Michael & Vedung, Evert. 2000. Intervention Theory as a Tool for Evaluation of Global Environmental Regimes. In: *Proceedings of the Final Conference within the EU-financed Concerted Action Programme on the Effectiveness of International Environmental Agreements and EU Legislation*. Barcelona, 9-11 November 2000. Fridtjof Nansen Institute. http://www.fni.no.

Rossi, Peter H. & Freeman, Howard. 1989. *Evaluation: A Systematic Approach*. Fourth edition. Newbury Park. California. Sage. 496 p.

Rotmans, Jan. 2001. Integrated Assessment: A bird's eye view. In: van Asselt, M.B.A., J. Rotmans & S.C.H. Greeuw (eds.), *Puzzle-solving for policy: A provisional handbook for Integrated Assessment*. Proceedings of the Advanced summer course, organised by the International Centre for Integrative Studies (ICIS), 30 August – 17 September 1999, Maastricht. Pp. 20-68.

Rotmans, Jan & van Asselt, Marjolein B.A. 2001. Uncertainty in integrated assessment modelling: A labyrinthic path. *Integrated Assessment* 2 (2): 43-55.

Rotschild, Kurt W. 1994. Power II. In: Hodgson, G.M., Samuels, W.J. & Tool, M.R. (eds.), *The Elgar Companion to Institutional and Evolutionary Economics*. Volume II. Aldershot, England and Brookfield, Vermont. Edward Elgar. Pp. 173-177.

Rutherford, Malcolm. 1996. *Institutions in Economics: The Old and the New Institutionalism.* Cambridge. Cambridge University Press. 225 p.

Rutherford, Malcolm. 2001. Institutional Economics: Then and Now. *Journal of Economic Perspectives* 15 (3): 173-194.

Ryan, Katherine. 2004. Serving Public Interests in Educational Accountability: Alternative Approaches to Democratic Evaluation. *American Journal of Evaluation* 25(4): 443-460.

Saarikoski, Heli. 2000. Environmental impact assessment (EIA) as collaborative learning process. *Environmental Impact Assessment Review* 20(6): 681-700.

Sabatier, Paul A. 1987. Knowledge, Policy-Oriented Learning and Policy Change: An Advocacy Coalition Framework. *Knowledge: Creation, Diffusion, Utilisation* 8(4): 649-692.

Sabatier, Paul A. & Jenkins-Smith, Hank C. 1993. *Policy Change and Learning: An Advocacy Coalition Approach*. Boulder, CO and San Francisco. Westview Press. 290 p.

Sabatier, Paul A. & Jenkins-Smith, Hank C. 1999. The Advocacy Coalition Framework: An Assessment. In: Sabatier, P.A. (ed.), *Theories of the Policy Process*. Boulder, CO. Westview Press. 117-166.

Sachs, Ignacy. 1980. *Stratégies de l'écodéveloppement.* Paris. Editions ouvrières. 140 p.

Sachs, Ignacy. 1993. *L'écodéveloppement : stratégies de transition vers le XXIe siècle.* Paris. Syros. 120 p.

Sachs, Ignacy. 1999. Social Sustainability and Whole Development: Exploring the Dimensions of Sustainable Development. In: Becker E. & Jahn, T. (eds.), *Sustainability and the Social Sciences*. MOST Project, UNESCO/ISOE. Zed Books.

Sachs, Wolfgang (ed.). 1992. *The Development Dictionary: A Guide to Knowledge as Power*. London; Atlantic Highlands, N.J. Zed Books. 306 p.

Sager, Tore. 2001. A planning theory perspective on the EIA. In: Hilding-Rydevik, T. (ed.), *EIA, large development projects and decision-making in the Nordic countries*. Nordregio Report 2001:6. Stockholm. Pp. 197-222.

Sairinen, Rauno. 2000. *Regulatory Reform of Finnish Environmental Policy*. Helsinki University of Technology, Centre for Urban and Regional Studies Publications A 27. Espoo. 284 p.

Samuels, Warren J. 1971. The Interrelations Between Legal and Economic Processes. *Journal of Law and Economics* 14(2): 435-450.

Samuels, Warren J. & Schmid, A. Allan. (eds.) 1981. *Law and Economics: An Institutional Pers*pective. Boston. Martinus Nijhoff. 268 p.

Sanders, James R. 1994. *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs.* The Joint Committee on Standards for Educational Evaluation. Thousand Oaks, New Delhi, and London. Sage. 222 p.

Sanderson, Ian. 2000. Evaluation in Complex Policy Systems. *Evaluation* 6(4): 433-454.

Sanderson, Ian. 2002. *Evaluating Complexity: Theory, Evidence and Practical Guidance*. Paper presented at the European Evaluation Society Fifth International Conference, Seville, 10-12 October 2002.

Sarbin, Theodore R. & Kitsuse, John I. (eds.). 1994. *Constructing the Social*. London and Thousand Oaks. Sage. 256 p.

Schmid, A. Allan. 1978. *Property, Power, and Public Choice: An Inquiry Into Law and Economics.* New York. Praeger. 316 p.

Schmid, A. Allan. 1994. Cost-Benefit Analysis. In: Hodgson, G.M., W.J. Samuels & M.R. Tool (eds.), *The Elgar Companion to Institutional and Evolutionary Economics*. Volume I. Aldershot, England and Brookfield, Vermont. Edward Elgar. Pp. 104-108.

Schnaiberg Alan, Adam Weinberg & David Pellow. 1999. *The Treadmill of Production and the Environmental State*. Paper prepared for mini-Conference on "The Environmental State under Pressure: The Issues and the Research Agenda", 6 August 1999, Chicago, Illinois. http://www.northwestern.edu/ipr/publications/papers/2004/schnaiberg/17_TreadmillEnvirState.pdf (consulted 1 September 2005)

Schubert, Uwe, André Martinuzzi, Elisabeth Huchler, Markus E. Langer & Bernhard Obermayr. 2000. Evaluation Nachhaltiger Entwicklung: Herausforderung für Methodenentwicklung und Evaluationspraxis. In: Härtel, M. & Stockmann, R. (eds.), *Berufliche Umweltbildung und Umweltberatung: Grundlagen, Konzepte und Wirkungsmessung.* Buchreihe des Bundesinstituts für Berufsbildung (BIBB). Bertelsmann-Verlag. Http://www.sustainability.at/easy.

Schwandt, Thomas A. 1997. Evaluation as Practical Hermeneutics. *Evaluation* 5(1): 69-83.

Schwandt, Thomas A. 2001a. A Postscript on Thinking about Dialogue. *Evaluation* 7(2): 264-276.

Schwandt, Thomas A. 2001b. Understanding Dialogue as Practice. *Evaluation* 7(2): 228-237.

Schwandt, Thomas A. 2003. 'Back to the Rough Ground!' Beyond Theory to Practice in Evaluation. *Evaluation* 9(3): 353-364.

Schwartz, Robert. 1998. The Politics of Evaluation Reconsidered: A Comparative Study of Israeli Programs. Evaluation 4(3): 294-309.

Scriven, Michael. 1991. Prose and Cons of Goal Free Evaluation. *Evaluation Practice* 12(1): 55-62.

Scriven, Michael. 1997. Truth and Objectivity in Evaluation. In: Chelimsky, E. & Shadish, W.R. (eds.), *Evaluation for the 21st century: A handbook.* Thousand Oaks, CA: Sage. Pp. 477-500.

Scriven, Michael. 1998. Minimalist Theory: The Least Theory That Practice Requires. *American Journal of Evaluation* 19(1): 57-70.

Scriven, Michael. 2001. Evaluation – Future Tense. *American Journal of Evaluation* 22(3): 301-307.

Segerholm, Christina. 2003. Researching Evaluation in National (State) Politics and Administration: A Critical Approach. *American Journal of Evaluation* 24(3): 353-372.

Seippel, Ornulf. 2002. Modernity, Politics, and the Environment: A Theoretical Perspective. In: Dunlap, R.E., F.H. Buttel, P. Dickens & A. Gijswijt (eds.), *Sociological Theory and the Environment: Classical Foundations, Contemporary Insights*. Lanham, Boulder, New York and Oxford. Rowman and Littlefield Publishers. Pp. 197-229.

Sen, Amartya K. 1985. Commodities and Capabilities. Amsterdam, New York and Oxford. North Holland. 130 p.

Sen, Amartya K. 1999. *Development as Freedom.* Oxford. Oxford University Press. 366 p.

Sen, Amartya K. 2000. *The ends and means of sustainability.* Keynote address at the international conference on "Transition to Sustainability", Tokyo. Http://www.iisd.org/pdf/sen_paper.pdf (consulted 1 September 2005)

Shaw, Ian & Crompton, Amanda. 2003. Theory, Like Mist on Spectacles, Obscures Vision. *Evaluation* 9(2): 192-204.

Siebenhühner, Bernd. 2001. *How Do Scientific Assessments Learn? A Comparative Study of LRTAP and IPCC.* Paper presented at 'Frontiers 1' European Society for Ecological Economics Conference, Cambridge, 4-7 July 2001.

Simon. 2003. Sustainability Indicators. The ISEE Internet Encyclopaedia of Ecological Economics. 10 p. http://www.ecoeco.org/publica/encyc_entries/SustIndicator.doc (consulted 1 September 2005).

Skolranking blev politisk verktyg. Huvsfudstadsbladet, 31 March 2005, p. 11.

Solow, Robert M. 1974. The Economics of Resources or the Resources of Economics. *American Economic Review* 64(2): 1-14.

Solow, Robert M. 1986. On the intergenerational allocation of natural resources. *Scandinavian Journal of Economics* 88(1): 141-149.

Solow, Robert M. 1994. An Almost Practical Step Toward Sustainability. In: *Assigning Economic Value to Natural Resources.* National Research Council: Commission on Geosciences, Environment, and Resources; Commission on Behavioral and Social Sciences and Education. National Academy Press. Washington D.C. Pp. 17-29. http://books.nap.edu/books/0309051436/html/19.html#pagetop (consulted 1 September 2005)

Soromenho-Marques, Viriato. 1997. Politica de ambiente em Portugal: Balanço e perspectivas. *Seara Nova* 57 (Julho/Agosto/Setembro), pp. 3-9.

Soromenho-Marques, Viriato. 1999. O Estado do Ambiente em Portugal: uma Perspectiva Critica sobre a Politica de Ambiente em Portugal. In: Coloquio « Ambiente, economia e sociedade », Conselho Economico e Social, 15 Março 1999, Lisboa. Pp. 83-95.

Soromenho-Marques, Viriato. 2001. Ambientalismo e politica em Portugal: Valores e Facetas de um Arco-iris Débil. In: Beckert, C. (ed.), *Natureza e Ambiente: Representações na cultura portuguesa*. Centro de Filosofia da Universidade de Lisboa. Pp. 115-128.

Spangenberg, Joachim H. 2001. Investing in sustainable development: the reproduction of manmade, human, natural and social capital. *International Journal of Sustainable Development* 4(2): 184-201.

Spash, Clive L. 1999. The Development of Environmental Thinking in Economics. *Environmental Values* 8(4): 413-435.

Spash, Clive L. 2000. Ehtical Motives and Charitable Contributions in Contingent Valuation: Empirical Evidence from Social Psychology and Economics. *Environmental Values* 9(4): 453-479.

Stake, Robert E. (ed.). 1975. *Evaluating the Arts in Education: a Responsive Approach.* Columbus, Ohio. Charles E. Merrill. 122 p.

Stake, Bob. 2004. How Far Dare an Evaluator Go Toward Saving the World? *American Journal of Evaluation* 25(1): 103-107.

Stecher, Brian M. & Davis, Alan W. 1987. *How to focus an evaluation.* Newbury Park, California. Sage. 94 p.

Stern, David I. 2003. The Environmental Kuznets Curve. The ISEE Internet Encyclopaedia of Ecological Economics. Http://www.ecoeco.org/publica/encyc_entries/Stern.pdf (consulted 5 May 2005).

Stern, Eric & Verbeek, Bertjan. 1998. Conclusions: Towards a Neopluralist Approach to Bureau-Governmental Politics. *Mershon International Studies Review* 42(2): 240-255.

Stevenson, Rodney. 2002. An Ethical Basis for Institutional Economics. *Journal of Economic Issues* 36(2): 263-277.

Stokes, Susan C. 1998. Pathologies of Deliberation. In: Elster, J. (ed.), *Deliberative Democracy.* Cambridge. Cambridge University Press. Pp. 123-139.

Strang, David & Chang, Patricia M.Y. 1993. The International Labour Organisation and the Welfare State: Institutional Effects on National Welfare Spending, 1960-1980. *International Organization* 47(2): 235-262.

Suomen pysyvä edustusto. 2003. OECD/Kestävä kehitys; tilannekatsaus. Taustamuistio OEC0010-48. Suomen pysyvä OECD-edustusto, Pariisi. Hintikka, Juva, Kyöstilä, Lehto, Suvanto. 17.7.2003.

(Finland's permanent delegation to the OECD, Paris. OECD/Sustainable Development; survey of the situation.)

Swaney, James A. 1985. Economics, Ecology, and Entropy. *Journal of Economic Issues* 19(4): 853-865.

Swaney, James A. 1987. Elements of Neoinstitutional Environmental Economics. *Journal of Economic Issues* 21(4): 1739-1779.

Söderbaum, Peter. 1999. Valuation as Part of a Microeconomics for Ecological Sustainability. In: O'Connor, M. & Spash, C. (eds.), *Valuation and the Environment: Theory, Method and Practice*. Cheltenham, UK and Northampton, MA, USA. Edward Elgar. Pp. 87-107.

Söderbaum, Peter. 2000. *Ecological Economics: A Political Economics Approach to Environment and Development*. London. Earthscan. 152 p.

Söderbaum, Peter. 2001. Political Economic Person, Ideological Orientation and Institutional Change: On competition between schemes of interpretation in economics. *The Journal of Interdisciplinary Economics*. 12(3): 179-197.

Taut, Sandy & Brauns, Dieter. 2003. Resistance to Evaluation: A Psychological Perspective. *Evaluation* 9(3): 247-264.

Taylor Fitz-Gibbon, Carol. 2002. Evaluation in an Age of Indicators: Challenges for Public Sector Management. *Evaluation* 8(1): 140-148.

Theys, Jacques. 2002a. Environmental Governance. From Innovation to Powerlessness. In: Grote, J. R. & Gbikpi, B. (eds.), *Participatory Governance: Political and Societal Implications*. Opladen. Leske & Budrich. Pp. 213-244.

Theys, Jacques. 2002b. L'approche territoriale du "développement durable", condition d'une prise en compte de sa dimension sociale. *Développement Durable & Territoires*. 23 septembre 2002. http://www.revue-ddt.org/dossier001/D001_A05.htm (consulted 1 September 2005).

Thoenig, Jean-Claude. 2000. Evaluation as Usable Knowledge for Public Management Reforms. *Evaluation* 6(2): 217-229.

Tietenberg, Tom. 1988. *Environmental and Natural Resource Economics* (2nd edition). Glenview, Illinois. Scott, Foresman & Co. 559 p.

Torres, Rosalie T., Hallie S. Preskill & Mary E. Piontek. 1997. Communicating and Reporting: Practices and Concerns of Internal and External Evaluators. *Evaluation Practice* 18(2): 105-125.

Tool, Marc R. 1979. *The Discretionary Economy: A Normative Theory of Political Economy*. Santa Monica, California. Goodyear Publishing Company. 350 p.

Tool, Marc R. (ed.) 1993. *Institutional Economics: Theory, Method, Policy*. Boston, Dordrecht and London. Kluwer Academic Publishers. 360 p.

Toulemonde, Jacques. 1997. Faut-il libérer l'évaluation de ses liens de causalité? Une réponse illustrée par les politiques structurelles européennes. *Gérer & comprendre*, mars, no. 47, pp. 76-88.

Touraine, Alain. 1999. *Comment sortir du liberalisme ?* Paris. Fayard. 164 p.

Truzzi, Marcello. 1971. Sociology: The Classic Statements. Oxford University Press. New York. Pp. 145-154. http://www2.pfeiffer.edu/~lridener/courses/GEMEIN.HTML (consulted 27 August 2005)

Tullock, Gordon, Arthur Seldon & Gordon L. Brady. 2000. *Government: Whose Obedient Servant? A Primer in Public Choice*. IEA Readings 51. The Institute of Economic Affairs. London. 184 p.

Turner, Kerry R. 1999. Environmental and ecological economics perspectives. In: J.C.J.M. van den Bergh (ed.), *Handbook of Environmental and Resource Economics*. Chapter 67, pp. 1001-1033.

Turner, Kerry R., David Pearce & Ian Bateman 1993. *Environmental Economics: An Elementary Introduction.* Baltimore. Johns Hopkins Univeristy Press. 328 p.

Ulkoasiainministeriö. 2003a. Kokousmuistio: OECD:n ympäristökomitean alaisen globaali- ja rakennepolitiikka –työryhmän (WPGSP) kokous 24.-25.4.2003. HEL0472-8. Ulkoasiainministeriö, Globaaliasioiden osasto; Kati Heikinheimo & Outi Berghäll. (Ministry of Foreign Affairs. Report of the meeting of the OECD Working Party on Global and Structural Policies – WPGSP, 24-25 April 2003. Ministry of Foreign Affairs, Finland, Department of Global Issues).

Ulkoaisiainministeriö. 2003b. Pöytäkirja. OECD-koordinaatioryhmän kokous 18.3.2003. Ulkoasiainministeriö, Kauppapoliittinen osasto, Juuso Moisander. HEL0419-4. (Report of the meeting of the interministerial OECD coordination group, 18 March 2003. Ministry of Foreign Affairs, Department of Trade Policy).

Ulkoasiainministeriö. 2003c. Kokousmuistio: Suppean OECD-koordinaatioryhmän kokous 3.9.2003. Ulkoasiainministeriö, Kauppapoliittinen osasto, Petra Paasilinna. HEL0419-58. (Report of the meeting of the small interministerial OECD coordination group, 3 September 2003. Ministry of Foreign Affairs, Department of Trade Policy).

Underdal, Arild. 1980. Integrated marine policy: What? Why? How? *Marine Policy* 4(3): 159-169.

United Nations. 1994. Agenda 21: Programme of Action from Rio. New York. United Nations Department of Public Information. 338 p. http://www.sidsnet.org/docshare/other/Agenda21_UNCED.pdf (consulted 1 September 2005).

Upton, Simon. 2002. Retour à l'essentiel. *L'Observateur OCDE*, no. 233, août 2002. Pp. 8-10.

Uusikylä, Petri. 1999. Politiikan ja hallinnon arviointi (Evaluation of policy and administration). In: Eräsaari, R., T. Lindqvist, M. Mäntysaari & M. Rajavaara (eds.), *Arviointi ja asiantuntijuus (Evaluation and expertise)*. Helsinki. Gaudeamus. Pp. 17-30.

Uusikylä, Petri & Virtanen, Petri. 2000. Meta-Evaluation as a Tool for Learning: A Case Study of the European Structural Fund Evaluations in Finland. *Evaluation* 6(1): 50-65.

Valovirta, Ville. 2002. Evaluation Utilization as Argumentation. *Evaluation* 8(1): 60-80.

Valve, Helena. 1999. Frame Conflicts and the Formulation of Alternatives: Environmental Assessment of an Infrastructure Plan. *Environmental Impact Assessment Review* 19(2): 125-142.

Valve, Helena. 2003. *Social Learning Potentials Provided by EU Rural Development Programmes: A Comparative Study on Three Institutionalisation Processes*. Doctoral dissertation. Acta Universitatis Tamperensis 918. Tampere University Press.

Van Asselt, Marjolein & Rotmans, Jan. 2001. Introduction. In: van Asselt, M.B.A., J. Rotmans & S.C.H. Greeuw (eds.), *Puzzle-solving for policy: A provisional handbook for Integrated Assessment*. Proceedings of the Advanced summer course, organised by the International Centre for Integrative Studies (ICIS), 30 August – 17 September 1999, Maastricht. Pp. 7-19.

Van den Bergh, Jeroen C.J.M. & Gowdy, John M. 2000. Evolutionary Theories in Environmental and Resource Economics: Approaches and Applications. *Environmental and Resource Economics* 17(1): 37-57.

Van der Knaap, Peter. 1995. Policy Evaluation and Learning: Feedback, Enlightenment or Argumentation? *Evaluation* 1(2): 189-216.

Van der Knaap, Peter. 2000. Performance Management and Policy Evaluation in the Netherlands: Towards an Integrated Approach. *Evaluation* 6(3): 335-350.

Van der Meer, Frans-Bauke. 1999. Evaluation and the Social Construction of Impacts. *Evaluation* 5(4): 387-406.

Van Waarden, Frans. 1995. Persistence of National Policy Styles: A Study of Their Institutional Foundations. In: Unger, B. & van Waarden, F. (eds.), *Convergence or Diversity: Internationalization and Economic Policy Response*. Aldershot and Brookfield. Avebury. Pp. 333-372.

Vartiainen, Pirkko. 2001. Intressiryhmät evaluointiprosessissa: Metodologista pohdintaa monitahoarvioinnin avulla (Interest Groups in the Evaluation Process: Methodological Reflections Through Multiple Constituency Evaluation). *Hallinnon tutkimus (Administrative Studies)* 20(1): 30-40.

Vatn, Arild. 2002. *Rationality, institutions and environmental policy*. Paper presented at the 7th Biennial Conference of the International Society for Ecological Economics, Sousse, Tunisia, March 2002.

Veblen, Thorstein. 1919. *The Place of Science in Modern Civilisation*. New York. Viking. Republished in 1990, New Brunswick. Transaction Publishers. 509 p.

Vedung, Evert. 1991. *Utvärdering i politik och förvaltning (Evaluation in politics and administration)*. Lund. Studentlitteratur. 239 p.

Vedung, Evert. 1997. *Public Policy and Program Evaluation*. New Brunswick. Transaction Publishers. 336 p.

Vedung, Evert. 2001. Utvecklingsdrag i utvärdering (Development Trends in Evaluation). *Hallinnon tutkimus (Administrative Studies)* 20(3): 139-143.

Verbeek, Bertjan. 1998. International organizations: The ugly duckling of international relations theory? In: Reinalda, B. & Verbeek, B. (eds.), *Autonomous policy making by international organizations.* London. Routledge. Pp. 11-26.

Verbeek, Bertjan. 2001. *Decision making within international organisations. Problems, opportunities, prospects*. Paper presented at the workshop on Decision Making in International Organizations at the Joint Sessions of the ECPR, Grenoble, April 5-11, 2001. Available at http://www.essex.ac.uk/ECPR//events/jointsessions/paperarchive/Grenoble/ws12/verbeek.pdf (consulted 15 May 2005)

Victor, Peter A. 1991. Indicators of sustainable development: Some lessons from capital theory. *Ecological Economics* 4(3): 191-213.

Victor, Peter, Susan Hanna & A. Kubursi. 1998. How strong is weak sustainability? In: Faucheux, S., M. O'Connor & J. van der Straaten (eds.), *Sustainable Development: Concepts, Rationalities and Strategies.* Dordrecht, Boston and London. Kluwer Academic Publishers. Pp. 195-210.

Vi måste satsa allt på jobben. Hufvudstadsbladet, 8.8. 1998.

Vivien, Franck-Dominique. 1998. Bioeconomic Conceptions and the Concept of Sustainable Development. In: Faucheux, S., M. O'Connor & J. van der Straaten (eds.), *Sustainable Development: Concepts, Rationalities and Strategies.* Dordrecht, Boston and London. Kluwer Academic Publishers. Pp. 57-68.

Wallerstein, Immanuel. 1974. *The Modern World System: Capitalist Agriculture and the Origins of the European World Economy in the Sixteenth Century.* New York. Academic Press.

WCED. 1987. *Our Common Future.* World Commission on the Environment and Development. Oxford and New York. Oxford University Press. 383 p.

Weale, Albert. 1992. *The New Politics of Pollution.* Manchester. Manchester University Press. 227 p.

Webler, Thomas. 1995. *'Right' discourse in citizen participation: an evaluative yardstick*. In: Renn, O., T. Webler & P. Wiedeman (eds.), *Fairness and Competence in Citizen Participation: Evaluating Models for Environmental Discourse.* Dordrecht and Boston. Kluwer Academic Publishers. Pp. 35-86.

Weiss, Carol H. 1980. Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilisation* 1(3): 381-404.

Weiss, Carol H. 1983. The stakeholder approach to evaluation: Origins and promise. In: Byrk, A.S. (ed.), *Stakeholder-based Evaluation.* New directions for program evaluation, 17. San Francisco. Jossey-Bass Publishers. Pp. 3-14.

Weiss, Carol H. 1987. The Circuitry of Enlightenment: Diffusion of Social Science Research to Policymakers. Knowledge: Creation, Diffusion, Utilisation 8(2): 274-281.

Weiss, Carol H. 1988. Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice* 9(1): 5-20.

Weiss, Carol H. 1998. Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation* 19(1): 21-33.

Weiss, Carol H. 1999. The Interface Between Evaluation and Public Policy. *Evaluation* 5 (4): 468-486.

Westergaard, John & Resler, Henrietta. 1976. *Class in a Capitalist Society: A Study of Contemporary Britain.* Harmondsworth. Penguin. 432 p.

Wilber, Charles K. & Harrison, Robert S. 1978. The Methodological Basis of Institutional Economics: Pattern Models, Storytelling, and Holism. *Journal of Economic Issues* 12(1): 66-89. (Reprinted in Samuels, Warren J. (1988), *Institutional Economics*, Vol II. Aldershot. Edgar Elgar. Pp. 95-123.)

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York and London. The Free Press. 450 p.

Williamson, Oliver E. 2000. The New Institutional Economics: Taking Stock, Looking Ahead. *Journal of Economic Literature* 38(3): 595-613.

Williamson, Oliver E. & Winter, Sidney G. (eds.). 1991. *The Nature of the Firm: Origins, Evolution, and Development.* New York and Oxford. Oxford University Press. 235 p.

Wimbush, Erica & Watson, Jonathan. 2000. An Evaluation Framework for Health Promotion: Theory, Quality and Effectiveness. *Evaluation* 6(3): 301-321.

Winston, Jerome A. 1999. Performance Indicators – Promises Unmet: A Response to Perrin. *American Journal of Evaluation* 20(1): 95-99.

Woolcock, Michael. 2001. The Place of Social Capital in Understanding Social and Economic Outcomes. In: *The contribution of Human and Social Capital to Sustained Economic Growth and Well-Being.* International Symposium Report. Human Resources Development Canada (HRDC) and OECD. Chapter 5, pp. 65-88.

Woolcock, Michael & Narayan, Deepra. 2000. Social capital: implications for development theory, research and policy. *World Bank Research Observer* 15(2): 225-249.

Yandle, Bruce, Madhusudan Bhattarai & Maya Vijayaraghavan. 2004. *Environmental Kuznets Curves: A Review of Findings, Methods, and Policy Implications.* PERC Research Study 02-1. Update, April 2004. Http://www.perc.org/pdf/rs02_1a.pdf (consulted 1 September 2005).

Yearley, Steven. 2002. The Social Construction of Environmental Problems: A Theoretical Review and Some Not-Very-Herculean Labors. In: Dunlap, R.E., F.H. Buttel, P. Dickens & A. Gijswijt (eds.), *Sociological Theory and the Environment: Classical Foundations, Contemporary Insights*. Lanham, Boulder, New York and Oxford. Rowman and Littlefield Publishers. Pp. 197-229.

Yee, Albert S. 1996. The causal effect of ideas on policies. *International Organization* 50(1): 69-108.

Zaccaï, Edwin. 1999. Développement durable: caractéristiques et interprétations. Cahiers du CEDD, no. 4. 13 p.

Zaccaï, Edwin. 2002. *Le développement durable: Dynamique et constitution d'un projet.* Bruxelles, Bern, Berlin, Frankfurt/M, New York, Oxford and Wien. P.I.E.-Peter Lang. 358 p.

Zhu, J. 1992. Issue competition and attention distraction: A zero-sum theory of agenda setting. *Journalism Quarterly* 69(4): 825-836.

Zoeteman, Kees. 2002. Globalisation and Sustainability: on Governance and the Power(lessness) of the Nation State. Inaugural speech as professor in Sustainable Development Policies and Globalization at the Tilburg University, the Netherlands. Globus – Institute for Globalisation and Sustainable Development. Tilburg University. 60 p.

ANNEX I.   Interviewed persons

The individuals are below classified under the country or the institution they were primarily considered to represent in the interview.

Austria

Manfred Schneider (WPEP delegate)
Director
International Environmental Affairs
Federal Ministry of Agriculture, Forestry,
Environment and Water Management
1 July and 10 November 2004

Canada

Guy Serge Côté (WPEP delegate)
Senior Policy Advisor
Environment Canada
15 July 2004

Denyse Rousseau
Deputy Director
Division of Environmental Relations and
Sustainable Development
Department of Foreign Affairs and International
Trade
15 July 2004

Finland

Jorma Julin
Ambassador
Finland's permanent delegation to the OECD
14 January 2005

Pirkko-Liisa Kyöstilä
Counsellor
Finland's permanent delegation to the OECD
11 January 2005

Arvi Suvanto
Counsellor
Finland's permanent delegation to the OECD
2 February 2005

France

Christophe Blanc
Chef du Bureau C4, Environnement et
Agriculture
Direction de la prévision et de l'analyse
économique
Ministère de l'économie, des finances et de
l'industrie
2 February 2005

Jean-Christophe Boccon-Gibod
Bureau C4, Environnement et Agriculture
Direction de la prévision et de l'analyse
économique
Ministère de l'économie, des finances et de
l'industrie
2 February 2005

Louis De Gimel
Sous-directeur
Direction des études économiques et de
l'évaluation environnementale (D4E)
Ministère de l'écologie et du développement
durable
10 February 2005

Roger Jumel
Chargé de mission pollution nitrates, qualité de
l'eau
Ministère de l'agriculture, de l'alimentation, de
la pêche et des affaires rurales
4 February 2005

Jean-Luc Laurent
Chef de l'inspection générale de
l'environnement
Ministère de l'écologie et du développement
durable
3 February 2005

Richard Lavergne
Secrétaire Général
Observatoire de l'Energie
Direction générale de l'énergie et des matières premières
Ministère de l'économie, des finances et de l'industrie
23 February 2005

Gérard Lebourdais
Sous-directeur du soutien aux territoires et aux acteurs ruraux
Ministère de l'agriculture, de l'alimentation, de la pêche et des affaires rurales
4 February 2005

Louis Meuric
Adjoint au Secrétaire Général de l'Observatoire de l'Energie
Direction générale de l'énergie et des matières premières
Ministère de l'économie, des finances et de l'industrie
23 February 2005

Hungary

Istvan Pomazi (WPEP delegate)
Chief Governmental Offical
Ministry of Environment and Water
Department for International Environmental Policy
1 July 2003

Japan

Tsuyoshi Kawakami
Administrator
OECD Environment Directorate
Environmental Performance and Indicators Division
(during the review of Japan, Deputy Director of the Division of Recycling and Waste Management Bureau at the Ministry of the Environment)
17 August 2004

Akinori Ogava (former WPEP delegate)
Deputy Director Planning Division
Air Quality Bureau
Environment Agency of Japan
3 July 2003

Mexico

Veronique Deli Meadows
Minister for Environmental Affairs
Permanent Delegation of Mexico to the OECD
23 November 2004

The Netherlands

Henk Jan Bakker
Counsellor
Permanent Delegation of the Netherlands to the OECD
11 February 2005

Yvo de Boer
Deputy Director General,
Ministry of Housing, Spatial Planning and the Environment
International Environmental Affairs Directorate
10 November, 2004

Richard Braakenburg
Coordinator of environmental affairs
Ministry of Transport, Public Works and Water Management
31 March 2004

Drs. Olav-Jan van Gerwen
Head SCA
Coordination of Policy Advice
Netherlands Environmental Assessment Agency
30 March 2004

Dr. Gerard Keijzers
Professor Sustainable Entrepreneurship
Center for Sustainability
Universiteit Nyenrode – the Netherlands Business School
29 March 2004

Dr. Paul Koutstaal
Inspectorate of the Budget
Ministry of Finance
30 March 2004

Klaas Jan Moning (WPEP delegate)
Head of the Department for External Integration
Directorate for Strategy and Policy Affairs
Directorate-General for Environmental
Protection
Ministry of Housing, Spatial Planning and the
Environment
29 March 2004

Prof. Lucas Reijnders
Stichting Natuur en Milieu
1 April 2004

Portugal

José Manuel Alho
Liga para a Protecção da Natureza (LPN)
11 April 2003

Teresa Avelar
Environmental Auditor
Ministry of Agriculture, Rural Development and
Fisheries
10 April 2003

Jaime Braga
Counsellor of Industrial Affairs
Confederation of Portugese Industry
11 April 2003

João Gonçalves
Director General
Environment Institute
Ministry for Environment and Land Use
Planning
10 April 2003

Antonio Gonçalves Henriques
Director General
Ministry for Environment and Land Use
Planning
10 April 2003

Maria Isabel F. P. Guerra
Environmental Auditor
Ministry of Public Works, Transport and
Housing
11 April 2003

Pedro Liberato (WPEP delegate)
Director
ProgramaPolis
Ministry for Environment and Land Use
Planning
(Portugal's WPEP delegate)
8 April 2003

Christiano Afonso de Oliveira Domingues
Counsellor
Permanent Delegation of Portugal to the OECD
13 January 2005

Lia Vasconcelos
Liga para a Protecção da Natureza (LPN)
11 April 2003

Viriato Soromenho-Marques
Professor of Political Philosophy
University of Lisbon
10 April 2003

Republic of Slovakia

Frantisek Kolocany (WPEP delegate)
Deputy Director
Department of Environmental Concepts and
Planning
Ministry of the Environment
19 May 2003

OECD

Christian Avérous
Head of Division
Environmental Performance and Information
Environment Directorate
17 May 2002

Sveinbjorn Blondal
Head of Structural Policy Analysis Division II
Economics Department
18 February 2005

Heino von Meyer
Head of the OECD Berlin centre
(former administrator at the OECD Environment
Directorate, Environmental Performance and
Information Division)
4 October 2004

ANNEX II.   Interview guide for the semi-structured interviews

Prior attitudes and expectations

- what benefits/impacts did you personally / the government expect from the review?

Background information concerning the political and administrative situation in the country

- status of the Ministry of the Environment in the government
- position of the WPEP delegate in the Ministry
- changes in government coalitions and administrative structures during the review period
- relationships between different actors in the field of environmental policies
- is there a tradition of evaluating public policies in the country?

Attitudes towards the OECD

- what is / should be the role of the OECD in international (environmental) politics?
- what is / should be the role of the Environmental Performance Reviews in OECD work and in international environmental policies?
- how is the OECD perceived among the various stakeholders and by the general public?
- did the review influence your attitudes towards the OECD?

The review process

- were the roles of the different participants made clear enough in the beginning of the review process?
- how did you / other participants perceive the meetings with the OECD team during the review mission?
- did the review process influence relationships between actors in the country?
- did the review process enhance networking between the participants?
- how were the processes of preparation for the review and commenting to the draft review report organised?
- how did you / other participants perceive the peer review meeting in Paris?
- did you learn new skills in the course of the review?

Dissemination

- what was done in order to disseminate the review report? (seminars, press conferences; who participated?)
- what was the breadth and tone of the media coverage following the release of the review?
- was there a person / group of persons in the administration committed to bringing forward the recommendations and advocating for their implementation?
- was a follow-up report produced, and if so, what was its impact?

Conclusions and recommendations of the review report

- what is your overall opinion concerning the recommendations (concrete/abstract, relevant/irrelevant, critical/'kind', etc.)
- did the C&R contain new, innovative ideas?

- to what extent were there recommendations adopted as such / rejected?
- did any of the actors use the recommendations for persuasion, legitimisation, critique, or defence?

Influence of the reviews

- in which ways, if any, did the review (process or report) influence policies and actors in the country?
- did the review change your opinions and beliefs concerning environmental policies?

Social-environmental interface

- did you consider the chapter on environmental-social interface as useful ? why?

Reasons for the review influence

- which were the main reasons for the review having / failing to have an influence?