

Soft Persuasion Through IEA Energy Policy Reviews: Transitions Towards Sustainable Energy?

Markku Lehtonen
Ph.D., Research Fellow
Sussex Energy Group
SPRU – Science and Technology Policy Research
The Freeman Centre, University of Sussex
Brighton BN1 9QE
United Kingdom
Tel: +44 (0)1273 872787
Mobile: +44 (0)774 7316190
fax: +44 (0)1273 685865
e-mail: M.Lehtonen@sussex.ac.uk

Abstract

The peer reviews carried out by the International Energy Agency (IEA) on its member countries' energy policies at four-year intervals, represent one of the most established examples of 'soft' methods of governance, which have recently gained popularity especially in the EU. Lacking regulatory authority, the IEA's policy impact relies on 'peer pressure', reinforced by the attention from the media, civil society, and other member country governments to the reviewed country's policy performance. Previous studies on OECD reviews have shown that peer pressure operates mainly by causing indirect, subtle and long-term 'conceptual' impact, shaping frameworks of thought and setting agendas. Few comparative studies exist of 'soft' mechanisms across different international organisations and policy contexts (e.g. federal or semi-federal systems), in multilevel governance systems. Based on experience from other similar experiences, notably the OECD environmental peer reviews, this paper suggests a framework for analysing the capacity of 'soft' persuasion by IEA reviews to stimulate the needed fundamental transitions in the member countries towards more sustainable energy policies.

Table of contents

1	Soft persuasion through multilateral surveillance: benchmarking, mainstreaming, OMC and peer review	3
2	Does soft governance have an impact?	6
3	IEA and OECD peer reviews	7
4	Framework of analysis	9
4.1	Types of influence.....	9
4.1.1	Direct, conceptual and political use	10
4.1.2	Three levels of influence: individual, interpersonal and collective	11
4.1.3	Pathways of influence: learning vs. accountability.....	14
4.1.4	Influence of evaluation on democracy	14
4.2	Factors affecting evaluation influence	16
4.2.1	Role of the evaluator	16
4.2.2	Policy actors and actor networks	19
4.2.3	Country-specific institutional context.....	23
4.2.4	Review design	24
4.3	IEA message: what type of framing advocated by IEA?	25
5	Discussion	28
6	References	29

1 Soft persuasion through multilateral surveillance: benchmarking, mainstreaming, OMC and peer review

‘Soft’ methods of governance have recently gained popularity in international politics, through increasing use of mechanisms such as benchmarking, mainstreaming, and the EU’s ‘Open Method of Coordination’ (OMC). Soft governance promises to alleviate the *legitimacy crisis* of top-down supranational policymaking in the EU (Meyer 2004, 815), provide *flexibility*, enabling states to minimise sovereignty losses and ‘shift the blame’ for unpopular decisions to EU, and e better able to help solve disagreements (Schäfer 2006, 83-84). It is slightly paradoxical that the soft mechanisms should be in the vogue within the EU at the moment when the ‘father’ of international peer review, the OECD is undergoing a persistent identity crisis. In energy policy, the rise of soft measures is related to the nature of energy as an area of ‘high politics’, which has effectively prevented greater international cooperation and diminished governments’ willingness to relinquish their sovereignty to international organisations (e.g. Keohane 1978).

Benchmarking has its origins in the business world, where it has been associated with decentralisation and devolution of management authority to lower levels, as well as greater reliance on market-based steering mechanisms (Arrowsmith and others 2004). Today, benchmarking is a major instrument of soft governance, not only within the OMC, but also in environmental policy, e.g. in the form of the European Environment Agency (EEA) ‘Country Scorecards’ (EEA 2005, 408). According to the EEA the scorecard allows the reader to make policy relevant and informative comparisons between countries and amongst issues against performance benchmarks.

‘*Mainstreaming*’ usually refers to the promotion of gender equality, but has in the environmental field been applied to denote the integration of environmental concerns into ‘harder’, economic fields of policymaking. Edquist (2006, 512), in turn, understands mainstreaming as the “conversion of policy solutions developed by NGO, alternative, or ‘parallel programmes’ into the generally and officially accepted policy framework”. In energy policy, mainstreaming can be seen as a major tool for promoting a ‘sustainability transition’, by integrating environmental concerns in the heart of the sector’s decision-making.

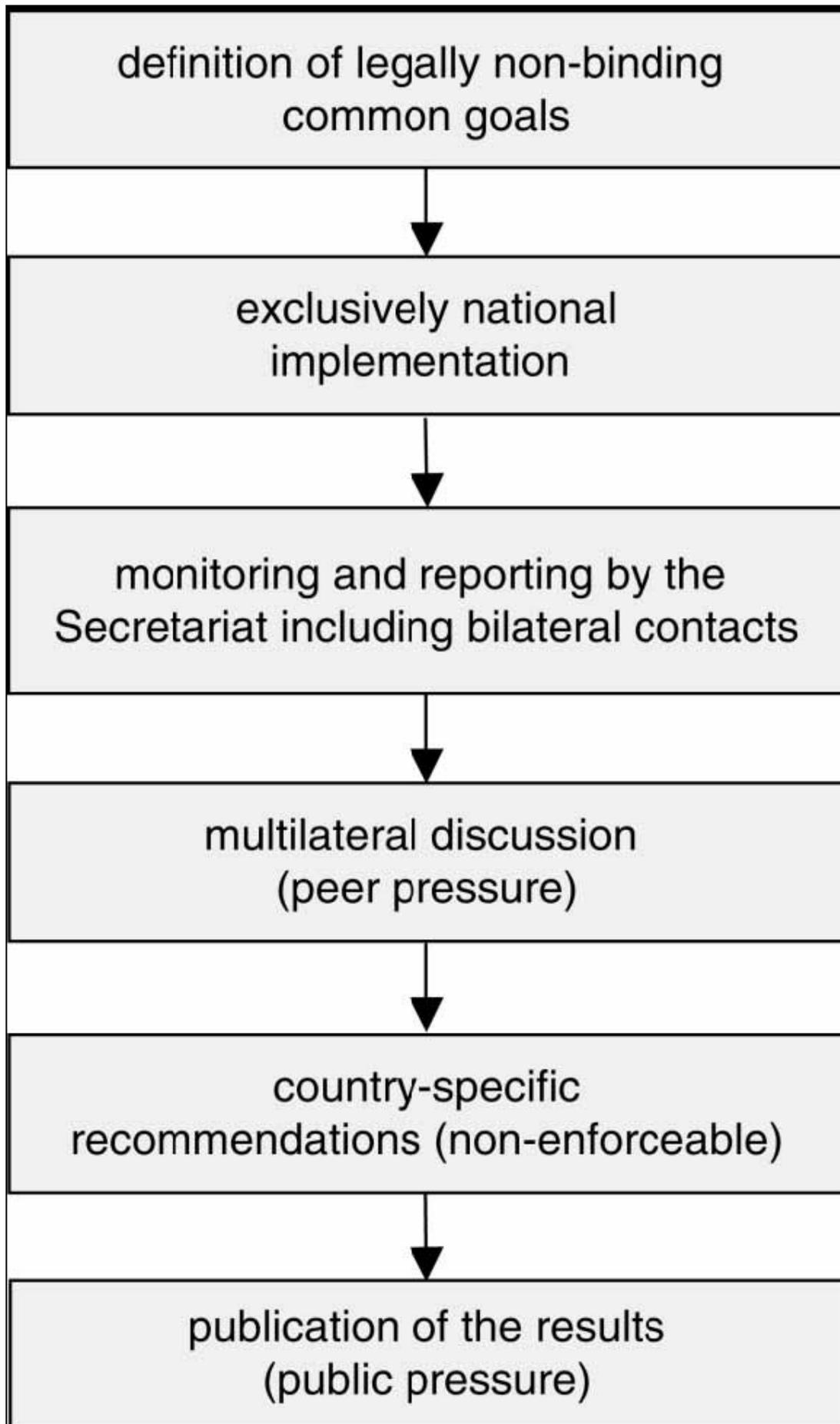
Finally, the Open Method of Coordination is designed to be a mutual learning process, involving the setting up, at the EU level, policy guidelines and goals, indicators for benchmarking national performance against the best in the world, and periodic monitoring, evaluation and peer review of member states (e. g. Borrás and Greve 2004; de la Rosa 2005; EU 2000, 12; Room 2005). While OMC falls within the general ‘soft law’ tradition in the EU, such as collective recommendations, review and monitoring, and benchmarking, it also constitutes a break from the older tradition, notably in entailing monitoring at a higher political level, being based on clear procedures and systematic linking across policy areas, and in explicitly seeking to enhance deliberation, participation, and learning from best practice (Borrás and Jacobsson 2004, 188-189; Eberlein and Kerwer 2002; Pochet and others 2001, 293).

All these mechanisms are to a greater or lesser extent inspired by the OECD peer review tradition, in place since the creation of the organisation in 1961. The OECD

(2003, 9) has defined international peer review 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*”. The review process is characterised by *dialogue* and interactive investigation, should be *non-adversarial* and rely on mutual *trust* between the States involved, as well as on shared *confidence* in the process. With these elements in place, peer reviews would create a system of mutual *accountability*. Media involvement, public scrutiny, dialogue and comparisons with peer countries exert pressure on the domestic public opinion, national administrations and policymakers. (ibid.) A country seldom wants to be seen in an unfavourable light among its peers, and *peer pressure* may therefore be effective in promoting compliance, notably in the ‘laggard’ states (Beyeler 2002; Marcussen 2003; Strang and Chang 1993). Peer reviews involve close involvement of the reviewed country government in the process, aimed at guaranteeing sufficient ‘ownership’, but try at the same time to maintain sufficient independence, which the credibility of the process relies upon.

Taken together, benchmarking, mainstreaming, OMC and peer review can be seen as different aspects of ‘*multilateral surveillance*’, which aims to bring states to comply with a code of conduct or specific goals, to develop common standards and acquire best practices through international comparison (Schäfer 2006, 80). Figure 1 presents the characteristics and phases of multilateral surveillance.

Figure 1. Characteristics and phases of multilateral surveillance (Schäfer 2006, 82)



2 Does soft governance have an impact?

Academic studies concerning the influence of soft mechanisms of governance have arrived at partly contrasting conclusions, largely because of the diverse theoretical understandings and analytical frameworks (Borrás and Greve 2004; López-Santana 2006, 486). For some, they provide a good example of advanced liberal government, individual freedom being conditioned by the techniques of government that shape the discourses. OMC's strength has been seen in its capacity to create precisely the political atmosphere and space that member states need: by freeing governments from the need to abide by details of community law-making, it would let them instead concentrate on vague and flexible generic output goals, and the ample room for interpretation and compliance can therefore make political action more practicable. (Borrás and Greve 2004)

Others have pointed out several defects. First, soft governance has been seen to lack effectiveness, because of the absence of sanctions and coercive mechanisms. Second, especially benchmarking has been criticised for focusing excessively on quantitative measurement, thereby hampering learning. Third, the context of multilevel governance presents particular challenges, because the soft methods are flexible at macro level, but impose strict accountability at local level. This also leads to unclear accountability structures and resistance from regional and local level against what is being perceived as a centralising method of governance (Arrowsmith and others 2004; Kaiser and Prange 2004). Fourth, the heterogeneity of country contexts has been seen as a real obstacle to soft methods having an impact. Fifth, the assumed deliberativeness of the soft methods has been seen largely as an unfulfilled promise (de la Porte and Nanz 2004; Kaiser and Prange 2004). Borrás and Greve (2004, 334) have summarised the critique in noting that the OMC runs the risk of ending up in a 'democratic legitimacy limbo', resulting from "the insufficient accomplishment of the deliberative and participatory democratic ideals, the virtual non-existence of democratic representation channels (by the European/national Parliament(s)), and the absence of output legitimacy owing to its indistinguishable results." These criticisms bear close resemblance with those levelled against the increasing use of performance measurement (Davies 1999; Greene 1999; Perrin 1998; Perrin 2002).

At the heart of soft mechanisms of governance lies the belief that tangible policy impacts can be produced not only through coercion and regulation, but also, and above all, through 'ideational' or 'framing' effects (Lopez-Santana 2006, 482): the direct impacts of OECD Economic Surveys on countries' social policies, for instance, tend to be modest (Armingeon 2003), whereas soft governance has the virtue of producing long term learning through socialisation (Marcussen 2003), creation of 'epistemic communities' (Armingeon 2003), by changing the mental models of the actors involved, affecting the ways in which problems are formulated and which issues come to national policy agendas and through dissemination of its 'organisational discourse' – in the case of the OECD dominated by a mainstream, liberal economics perspective (Dostal 2004). In this way, soft governance would be instrumental in promoting 'cognitive convergence' (Bruno and others 2006, 521). Seen from this perspective of ideational effects, soft governance is not simply about objective steering mechanisms, but has its normative dimension in 'instrumentation', whereby the different mechanisms of soft governance share the same 'management spirit', in which political issues are converted into specified target figures (ibid., 530).

This paper suggests a theoretical framework for analysing the capacity of the country reviews carried out by the International Energy Agency (IEA) to promote transitions towards more sustainable energy policies in its member countries. The framework is informed by experience from evaluation research in general, and in particular by previous studies on the influence of OECD peer reviews (e.g. Armingeon and Beyeler 2003; Dostal 2004; Lehtonen 2005a; Lehtonen 2005b; Lodge 2005). The following section describes the IEA review process and compares it with the OECD Environmental Review Process. Section 3 is devoted to the description of the analytical framework, which consists of three elements: identifying the types of influence from the reviews; analysing the factors that shape the type and extent of that influence; and examining the contents of the reviews, i.e. to what extent the IEA policy approach – as manifested in the reviews – corresponds to what is here understood as a sustainable energy policy perspective. The concluding section sums up the debate and highlights key issues for empirical research on IEA review influence.

3 IEA and OECD peer reviews

Since the theoretical framework suggested here is based on previous work on OECD peer reviews, notably the Environmental Performance Reviews, the IEA review process is contrasted with the OECD one. The general phases of the review process are the same in both reviews, consisting of the following stages:

Preparatory phase

- the IEA/OECD secretariat and the relevant ministry in the reviewed country agree on the timetable and organisation of the review
- government the secretariat gathers background material and prepares for the review partly on the basis of documents sent to it by the reviewed country
- a review team is set up, consisting of secretariat officials and experts from other member countries

Review mission

- Meetings between the review team and the officials and stakeholders in the reviewed country during the review visit of 6-10 days

Drafting phase

- Secretariat prepares a draft report, which is sent for comments to the reviewed country and various OECD directorates

Peer review meeting

- Draft report is discussed and conclusions and recommendations adopted
- Reviewed country delegation answers other member countries' questions

Release of the review report

- The report is published, often in a press conference in the reviewed country's capital

Follow-up of implementation of the recommendations

- Voluntary self-reporting (EPRs) or carried out by the secretariat (IEA)

The practical conduct of the reviews in the IEA secretariat is the responsibility of the Office of Long-Term Co-Operation and Policy Analysis (LTO). The Standing Group on Long-Term Co-operation (SLT), composed of member country energy officials, carries out the peer review and agrees on the review recommendations. The SLT meets four times a year in Paris, to share the results of the in-depth reviews, among other things. All of the IEA reviews are published annually, along with a synthesis report highlighting important cross-cutting policy issues.

The main differences between the two review processes are summarised in table 1.

Table 1. Main differences between EPRs and IEA reviews.

	EPRs	IEA
Design of reviews	Agenda & key themes defined in advance, partly in cooperation with the reviewed country government	Team's role is to define the agenda; the key themes to be addressed in review
Review process	Draft report only one month prior to peer meeting	10-15-page report with preliminary recommendations prepared at the end of the mission
Head of the review team	Head of the relevant OECD secretariat unit	Official from a reviewing country
Length of a full review cycle	7-8 years	4 years
Peer review meeting	Ceremonial & consensual debates in peer review Reviewed country delegations broad and cross-sectoral	Debates in peer review meetings are closed, and between energy policy officials and IEA secretariat
Independence of secretariat	Secretariat relatively independent in relation to reviewed country officials	Reviewed country officials have significant influence on review process and recommendations
Criteria of performance	Country's own objectives, international commitments and cost-effectiveness of policies (& shared OECD principles)	IEA Shared Goals approved by energy ministers in 1993
Follow-up	Voluntary follow-up report by reviewed country government 2-3 years after the review	In intervening years, brief annual reviews updating the main energy policy developments and report on progress of implementation

4 Framework of analysis

This section draws on experience from evaluation research in general, and studies on the influence of OECD peer reviews in particular, to suggest a general framework applicable for analysing the influence of IEA country reviews. The framework consists of three elements: first, the identification of the types of influence; second, the analysis of the factors conditioning the influence; and third, the type of policy message passed on through the framework of analysis applied by the IEA in its reviews.

4.1 Types of influence

The first step in the analysis of IEA reviews is to identify the different types of influence that the reviews engender. Direct use, whereby policymakers change their policies directly in response to evaluation findings, is only one – and relatively rare – type of impact. Often, evaluations are instead used conceptually or politically, partly in function of their varying purposes, and the impacts may take place at individual as well as collective levels. These different types of use and effects interact to create pathways of influence. Finally, evaluations' largely unintended impacts on democracy need to be taken into account.

4.1.1 Direct, conceptual and political use

The ways in which evaluations and assessments are used or otherwise influence policymaking are commonly divided into a few categories, such as direct/instrumental, conceptual, and political use or influence. The assumption that evaluations would be used *directly* by policymakers relies on a 'technocratic-rationalistic policymaking model, whereby evaluations would serve policy by identifying results of government interventions, thereby (1) helping the governments to decide whether or not to continue with particular policies; (2) expanding and institutionalising successful programs and policies, while cutting back unsuccessful ones; and (3) finding out which programmes or policies to modify and how. Rational policymakers, possessing clearly defined preferences and aiming at clearly established policy goals, would adjust policies according to the most recent available scientific information.

In practice, however, such direct use is often more of an exception than a rule. Many evaluations end up gathering dust on the shelves of policymakers and civil servants who seem to ignore the results of evaluation and carry on with policies chosen for other reasons. Since the rather naïve optimism embodied in the rational-technocratic policymaking model prevalent in the 1960s, research on the use of scientific knowledge in policymaking has repeatedly demonstrated the importance of the various indirect influences. In particular, Weiss (1980; 1999) coined the notion of 'enlightenment', or *conceptual use* of evaluations, entailing the idea that research affects decision-makers' problem definitions, 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. This view is put forward in particular by the various constructivist approaches, which see the evaluator as a facilitator in a negotiation process, bringing forward the various points of view involved and promoting consensus.

While important, direct and conceptual use overlook the fact that evaluations and scientific knowledge are frequently used '*politically*', in order to enhance the legitimacy of policies and actors, or to empower specific actor groups. Political use also encompasses tactical use, when an evaluation is commissioned e.g. to postpone a decision (Vedung 2001, 141), or symbolic use, aimed at conveying an image or a message. These political types of use could, however, also eventually lead to 'enlightenment', as policymakers gradually become sensitised to the perspectives of social science (Weiss 1980).

These types of use are intimately linked to the intended *purpose of evaluation*. An evaluation may be used to promote *control and accountability*, providing an overall

judgement on whether or not a policy has achieved its objectives – the evaluation is expected to be used directly to support decision-making. Often, evaluations seek to promote *learning* through the development and improvement of the evaluated policy or programme, and try to explain the causal relationships between a policy and outcomes. An evaluation may also have *'enlightenment'* as its explicit yet rarely as the primary objective: the evaluation would help the policymakers, programme personnel or citizens to better understand policy processes and factors affecting the course of events. Finally, an evaluation may function as a *'Potemkin village'*, when policymakers, government officials and programme managers seek to give the appearance that everything goes well, according to a rational policy plan, or to hide negative aspects and failures. (Vedung 1991, 74-81)

4.1.2 Three levels of influence: individual, interpersonal and collective

Classifying different types of evaluation use can be informative as such, yet the above-mentioned categories fail to account for a number of factors. First, the *'process use'* of evaluations points to the various types of learning that take place throughout the evaluation process – through interaction between individuals – rather than only at the end as a reaction to the publication of the evaluation results (Forss and others 2002; Patton 1998). Second, talking about the *'use'* or *'utilisation'* misses the point that evaluations may have indirect and unintended impacts – e.g. by framing the discussions – even without anybody expressly *'using'* them. In addition, focusing on use may create pressure to demonstrate that at least some *'use'* of evaluation has actually taken place, and limiting the attention to only one or a few forms of influence at a time, thereby overlooking broader impacts (Henry and Mark 2003). Third, distinguishing different types of use or influence does not in itself explain the ways in which different mechanisms interact to produce an impact. In particular, dialogue, argumentation and communication, and the above-mentioned *'process use'* promote the kinds of learning that constitute a major impact of evaluation.

In view of these shortcomings, a three-level framework of analysis is suggested here for examining the influence of evaluation. First, insights from evaluation findings or the evaluation process can act at the *individual level*, resulting in the confirmation of existing beliefs, rejection of evaluation findings, or new transformed beliefs or capacities. These new beliefs may mean that participants change their attitudes, attach greater or lesser salience to an issue, elaborate on aspects related to the topics dealt with in the evaluation, or start thinking in new ways about the issues at stake. The evaluation process may also help people acquire new skills such as evaluation methods and cooperation skills, ultimately leading to behaviour change.

At the second level of influence, the individual interpretations enter into collective processes of *deliberation, negotiation, argumentation and dialogue*. These processes involve different actors (governmental bodies, industry, academics, media, NGOs, etc.) using evaluations in order to (1) persuade others to adopt a given line of action; (2) legitimise their own actions; (3) criticise actions of others or the evaluation findings; and (4) defend themselves against criticism based on the evaluation findings (Valovirta 2002).

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*. 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 her own views on other actors.

Changes in decisions and actions are not, however, the only possible types of influence at collective level. Even in the absence of such concrete impacts, evaluations may create *new shared understandings* through ‘policy-oriented learning’. This may involve socialisation – i.e. modification of social norms – or rendering explicit different values, priorities and objectives of stakeholders, which should make arbitration between them easier. In general terms, new shared understandings can take the form of three alternative outcomes: production of a synthesis, a winner, or a deadlock. It is notable that even when decision-makers reject a policy recommendation, evaluation can nevertheless have had an impact, by providing stakeholders information about the each others’ reasoning and enabling them to sharpen their own arguments. These impacts, in turn, may influence policies in other areas, under different circumstances, etc. (Valovirta 2002.)

The third category of collective level effects consists of *increased or decreased legitimacy* of past and future actions, actors or organisations. Legitimacy can be internal or external: ‘process use’ may enhance *internal legitimacy* by ‘boosting moral’, as participation in the evaluation process may strengthen the actors’ commitment to the project, and motivate them to pursue their activities. This has been shown to have happened even in cases where evaluation has pointed out shortcomings, mistakes and abuses (Forss et al. 2002). The project or the policy may be strengthened as a result of changes in attitudes, salience and elaboration at the individual level. 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 (Forss et al. 2002). 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, while inversely, negative evaluation findings may decrease the legitimacy of a policy and incite the evaluated actor to imitate others.

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

Table 2 summarises the different types of influence from evaluations.

Table 2. Summary of the potential types and pathways of influence from an evaluation.

<u>Processes at the individual level</u>				
<u>BELIEFS</u>			<u>CAPACITIES</u>	
Confirmed	Transformed	Rejected		
	Attitude change		Skills	
	Saliency		Evaluation methods	
	Elaboration		Cooperation	
	Priming			
<u>Processes at interpersonal level: dialogue, deliberation and argumentation</u>				
PERSUASION	LEGITIMISATION	CRITICISM	DEFENCE	
	(justification)			
<u>Impacts at the collective level</u>				
DECISIONS AND ACTIONS	NEW SHARED UNDERSTANDINGS	LEGITIMACY	AGENDA-SETTING	NETWORKING
	(~ Policy-oriented learning)	Increased or decreased		

Finally, evaluations may also have negative impacts such as generating ‘evaluation fatigue’. The increasing frequency of different types of evaluations by supranational organisations, such as the EU, the IMF, the World Bank, coupled with the multiplication of reporting obligations under the various international conventions, has led many administrations and individual civil servants to perceive any new evaluation 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. Baron 1999, 84; Perrin 1998; Perrin 2002; Taut and Brauns 2003; Vedung 2001) The fact that the IEA and OECD reviews are not directly commissioned by any actor, but are in practice imposed upon the country’s government hardly makes the situation easier.¹

Previous studies on OECD reviews have shown evidence of all of these types of impact, with the processes of dialogue, deliberation and argumentation playing a central role. The reviews are frequently used by civil servants to legitimise their policies – e.g. greater use of economic instruments in environmental policy, or macroeconomic reforms by finance ministries. Sometimes also NGOs use reviews to criticise government policy.

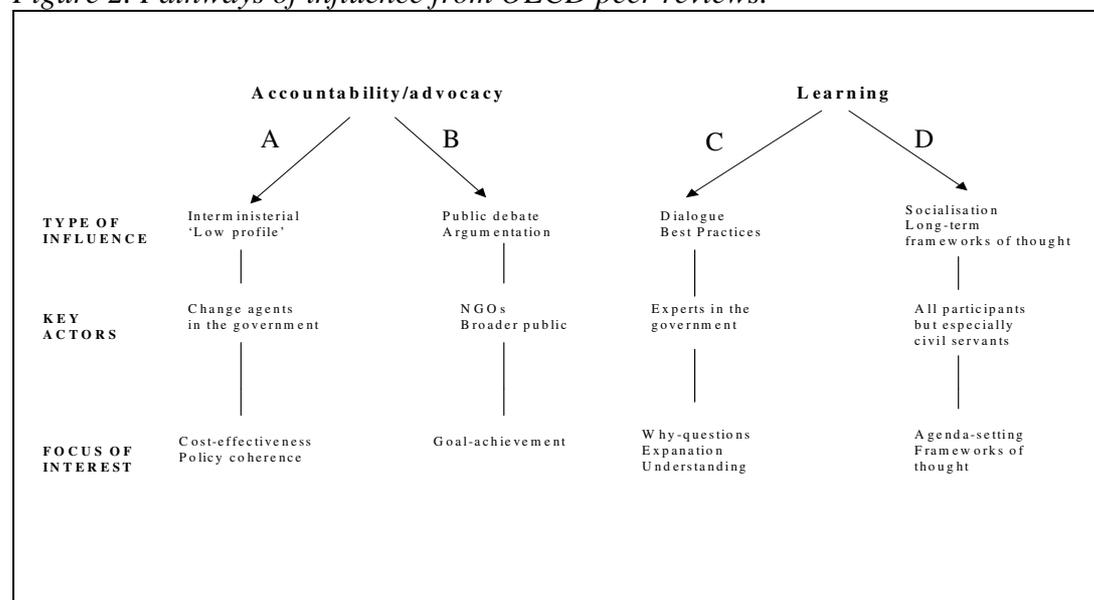
For understandable reasons the main focus of previous studies has been on the processes at the interpersonal and collective levels – after all, the ‘raison d’être’ of the OECD reviews is their presumed policy impact. More in-depth studies would be needed to establish the interaction between the individual and collective levels, i.e. how individual learning results in collective impacts, and vice versa.

¹ Of course, evaluations are not officially imposed upon countries, given that for instance carrying out a 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.

4.1.3 Pathways of influence: learning vs. accountability

Summing up the experience from other OECD reviews (notably environmental performance reviews), the pathways of influence can be categorised under four ‘ideal types’ (Figure 2). These pathways operate, on the one hand, through *pressure* inside the government (A) or through broader public pressure (B), and, on the other hand, through *learning* concerning the reviewed policies, which can involve relatively short-term and direct types impacts such as transfer of best practices, capacity building and building of professional networks (C), or more long-term influence through ‘framing effects’ (López-Santana 2006), gradual socialisation and formation of epistemic communities (e.g. Armingeon 2003), ‘enlightenment’ (e.g. Weiss 1980; 1999), and modification of mental models or frameworks of thought (D). OECD reviews therefore typically attempt to combine the learning and accountability functions of evaluation (Lehtonen 2005b).

Figure 2. Pathways of influence from OECD peer reviews.



The analysis of IEA reviews should identify the types of influence according to the typology presented above in table 2. Yet, given the existing experience from OECD EPRs, a major task would be to test the extent to which the four pathways in Figure 2 can be identified also in IEA reviews. In particular, research could try to establish how well the reviews succeed in combining the objectives of learning and accountability within a single evaluation structure, and identify the main (intended and unintended) beneficiaries of the reviews.

4.1.4 Influence of evaluation on democracy

Evaluations and IEA reviews have non only impacts on policies, but also have inevitable, mostly unintended, impacts on democracy. They can enhance, alternatively, elitist/representative, participatory or deliberative democracy. Different models of democracy and types of evaluation likewise entail different views on the role of civil society. Elitist democracy often goes together with a libertarian view, which sees the society as composed of two opposing ‘cells’ – the private sector and

the state – confronting each other in a zero-sum game. It starts from the assumption that individuals behave as ‘homo economicus’, maximising their individual economic utility. Participatory democracy can be inspired by a communitarian view, which perceives the civil society as a combination of more or less separate and autonomous groups united by a common identity. Deliberative democracy, in turn, holds a view of a ‘mediating’ civil society, defined as “civic communities that qualify as membership communities but are sufficiently open and egalitarian to permit civic participation on a voluntary basis” (Hanberger 2001). A feminist notion of civil society is similar, in seeing the distinction between the private and the civil sphere as blurred, with particular focus on benefiting all types of household and reducing gender inequalities.

An *expert-oriented evaluation*, adopting a positivistic and technocratic approach, sees evaluation essentially as scientific work that aims to keep facts and values apart and seek the objective truth. This type of evaluation tends to promote a libertarian civil society and an elitist/representative democracy. It may, however, reinforce the common misperception of science and evaluation as value-free activities.

An *advocacy approach* to evaluation would promote participatory democracy and a libertarian, communitarian, or feminist notion of civil society, depending on who is being empowered. The evaluator acts as an advocate serving her client’s needs, or a facilitator, educating and assisting the client to become a skilful self-evaluator. A downside of the advocacy approach is that it might promote a fragmented society, giving precedence to the interests of a specific deprived group, instead of serving the broader multicultural society.

A *mediating approach* to evaluation, finally, implies that some type of practical deliberation between stakeholders needs to take place, the evaluator acting as a counsellor and/or mediator inquiring, learning, and mediating together with various stakeholders, and seeking practical solutions to collective problems in the context of difference and conflict.

Especially in view of the legitimacy crisis that lies behind much of the interest in soft governance, Hanberger’s (2001) considerations about the ways of ‘taking evaluation to the people’ are relevant. He argues for the ‘mediating approach’ as the most promising one, since 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.

The IEA and OECD reviews typically combine elements of elitist-technocratic evaluations – by virtue of their strive to independence and objectivity – with the more or less explicit advocacy for the policy views of their client groups; the environmental and energy authorities. This does not mean that the reviews could not have a deliberative element. 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 represent an incentive to manipulate, and produce pathologies rather than better decision-making (Przeworski 1998; Stokes

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, an evaluation conducted in the spirit of a technocratic-positivist approach, can serve democracy by improving citizens’ possibilities to control the government, and has been shown to stimulate different types of learning (Owens and others 2004). Moreover, 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. Hence, 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).

In analysing the possibility of the IEA reviews to promote ‘sustainability transitions’, the key question is to what extent they serve to consolidate the existing, elitist perception of democracy, as opposed to helping to promote more deliberative democracy, which can be deemed necessary prerequisites for such transitions.

4.2 *Factors affecting evaluation influence*

For the examination of the role of IEA reviews, it is not sufficient to establish whether or not impacts have taken place or not – one must also try to understand the reasons for the impacts or the absence thereof. In this section, the factors possibly conditioning the degree of influence that an evaluation has on policy are divided into four categories: the role of an international organisation as evaluator; the involved actors with their expectations, interpretative frames and policy networks; the general country-specific policy context; and the way in which the evaluation has been designed.

4.2.1 Role of the evaluator

External or internal evaluation? Advocacy or independence?

IEA and OECD reviews at being external evaluations, carried out by an evaluator outside the evaluated entity. A strength of this type of evaluations is their greater credibility and legitimacy in the eyes of external stakeholders as compared to evaluations conducted internally by the evaluated entity itself. However, international peer review carries an element of internal evaluation, since the working party responsible for the reviews is made up of energy policy insiders from other countries, i.e. members of the same ‘epistemic community’. Indeed, persons involved in OECD peer reviews have mentioned the possibilities of decreased credibility and independence, pointing out that countries tend to refrain from all too harsh criticism against their peers, because they know that unduly harsh criticism could backfire once their own country is reviewed.

This double role of the IEA (and OECD) as supposedly independent outsider on the one hand, and an insider promoting the agenda of a specific ‘epistemic community’ on the other, creates dilemmas in particular in those areas of policy, which are not at the core of the organisation. Examples include areas such as social and environmental policy, in which the OECD reviews need to constantly balance between the advocacy

of sector policy views on the one hand, and coherence with OECD mainstream of liberal economics, on the other. Mabry (2002, 149-150) has formulated this dilemma more generally 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. 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).

One may assume such problems to be less acute in IEA reviews, whose main purpose is, in fact, to ensure that the member countries remain within the mainstream doctrine of the organisation, i.e. the reviews are the bread and butter of the organisation, rather than being at the margins as the EPRs in the OECD. As mentioned above, the differences in the review design and organisation suggest that the reviewed country government has more influence on the review findings than in the EPRs.

IEA as a creator of international norms and identities

Research placing emphasis on ‘ideational’ factors in international politics has demonstrated that international organisations act as creators of international norms and identities through processes such as socialisation, imitation and coercion through peer pressure. 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). In this role, the OECD peer reviews have been shown to promote particularly ‘*norm cascade*’ – a process of international socialisation intended to induce norm breakers to become norm followers (Finnemore and Sikkink 1998). This mechanism is particularly powerful with regard to ‘laggard’ states, whereas the frontrunners may be far more difficult to influence through the reviews. Experience shows that the OECD hardly plays the role of ‘*norm entrepreneur*’ through the environmental performance reviews, notably because the reviews seldom come up with genuinely new ideas, but instead rehearse the well-known OECD principles or arguments used in national policy discussion.² Finally, peer reviews enhance ‘*norm internalisation*’, by lending norms a taken-for-granted character, the polluter pays principle being among the most typical examples in the area of environmental policy.

It is fair to assume that the IEA reviews have a function similar to the OECD country reviews in mainly promoting norm cascade and internalisation. Whether the IEA reviews are better than the EPRs in enabling the IEA to act as a norm entrepreneur, i.e. introducing new policy approaches and ideas, is a question for further research. On the one hand, this would not seem likely as the reviews simply analyse progress in

² 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 (Anon. 2003).

relation to the IEA 'shared goals', but on the other hand, being at the heart of IEA activity, the reviews could have better chances than the EPRs to bring to the table new norms and ideas.

IEA and its member states

Four other issues are of relevance with regard to the role of the IEA. First, given its nature as an intergovernmental organisation, the power relations between its member countries are crucial. Four of the 30 OECD countries are not members of the IEA: Iceland, Mexico, Poland and the Slovak Republic.³ The US is the most influential country, not only by virtue of its superior economic and political resources, but also as the initiator of the IEA in early 1970s. Experience from other OECD reviews suggests that while these power relations cannot be ignored, their impact on the contents of the reviews is less than the asymmetries of power would suggest.

Second, OECD reviews clearly display a principal-agent dilemma. In principle, the member country governments decide upon the shape of the review programme, but in practice the OECD secretariat has a very strong influence. Member country officials lack the time and knowledge needed to engage in detailed discussions concerning the review programme, and country delegations have often been unable to agree on how the reviews should be developed.

Third, neither the OECD, nor the IEA secretariat is a monolith, with harmonious views on the policy issues at stake. One would expect divergences of views to be less pronounced within the IEA, which is far smaller and has a narrower policy remit than the OECD. Hence, it can be expected that the IEA reviews provide a more coherent analysis and stick closer to the organisation's core views than do the other than economic OECD reviews. The educational and professional background of the IEA staff is also likely to be more homogeneous than that of OECD officials.

Fourth, an organisation such as the IEA must take into account its role among the international governance structures. Since the IEA has no regulatory and coercive powers, it relies heavily on processes such as peer review in attempting to influence member states' policies. While the OECD has been going through an identity crisis since the 1990s, largely because of the end of the cold war and increased competition from other international organisations (e.g. WTO, EU), the IEA probably has a more clearly defined role as the most respected expert organisation providing reliable information on energy policies.

Kay questions concerning the role of the IEA can be summarised as follows:

- How do the IEA reviews balance between advocacy and independence?
- At which stage of the 'norm life cycle' do IEA reviews exert most influence?
- To what extent do power relations within the IEA affect the degree of influence of different member countries on the entire review programme and the contents of individual reviews?
- How far is the IEA secretariat able to influence the shape of the review programme?

³ The IEA membership negotiations of Poland and the Slovak Republic are at an advanced stage.

- What divergences of views on energy policy and the role of the reviews exist within the IEA secretariat on the one hand and among member countries on the other?
- What is the role of the IEA in international energy policy, and is its importance increasing or decreasing?

4.2.2 Policy actors and actor networks

The second element shaping the influence of IEA reviews is the role of policy actors and networks. The influence of these actors can take place through ‘epistemic communities’, but is conditioned by the actors’ expectations, mental framings (‘repertoires’), and the institutional context.

Expectations and epistemic communities: salience, credibility and legitimacy of reviews

IEA and OECD reviews typically represent mechanisms through which different epistemic communities⁴ advocate their views and policies and facilitate norm creation by international organisations. In broad terms, the OECD environmental performance reviews are a vehicle for the ‘environmentalist’ epistemic community (especially its ‘economically oriented’ part), congregated around environment ministries, to gain power and legitimacy. The Economic Surveys, in turn, are a tool in the hands of the powerful economic epistemic community, led by finance ministries and supported by the community of mainstream academic economists. It is fair to assume that the IEA reviews play a similar role in the hands of energy professionals and policymakers. One would therefore assume that the IEA reviews act to consolidate the mainstream energy policy approach, with particular emphasis on economic efficiency, competition and market liberalisation in line with the IEA/OECD policy doctrine.

In practice, the question of which epistemic community, if any, is dominant in using the reviews and defining their content is far more complex. In particular, the energy policy community is not homogeneous, but instead contains a number of different ‘sub-communities’, with different views regarding the role of the markets, the relative importance of environmental and energy security objectives, or the role of centralised as opposed to decentralised energy options, for instance. Relevant actors with varying interest in and influence on energy policies include not only the intended beneficiaries of the reviews – in OECD peer reviews typically the civil servants in the relevant policy sector – but also private sector actors (energy companies), politicians, NGOs (notably environmental ones), the civil society at large, and the media.

Furthermore, the influence of epistemic communities depends on their varying expectations concerning the reviews. Research suggests that 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

⁴ 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.” 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) (ibid., 18).

Dickson 1999; Eckley 2001).⁵ 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. Relationships between the actors also influence the criteria of credibility and legitimacy. 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) and the OECD and IEA influence crucially depends on their perceived independence and hence, credibility.

What different actors consider credible, salient and legitimate depends on their interests and expectations. Actors' expectations may be of *cognitive* nature (information and knowledge concerning the way a policy is being implemented, its outputs, results, outcomes, impacts, etc), but in reality, most actors' expectations are either *strategic* or *political* (Barbier 1999, 377), relating to power and legitimacy. 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).

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)".⁶ 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. Ambiguity stems not only from outside of one's own organisation or group, but also from within, because individuals are generally involved in multiple, partly related

⁵ *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.

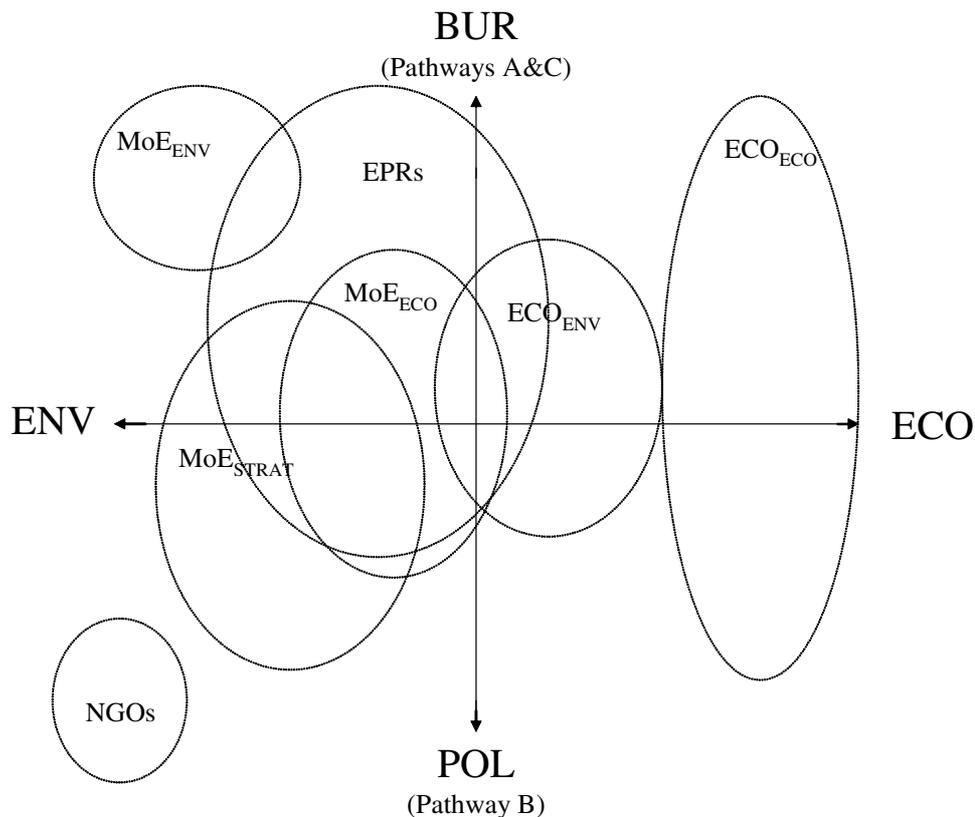
⁶ The concept is close to 'frame' (Giddens 1984, 87; see also Valve 1999) or 'ideological orientation' (Söderbaum 2001).

social contexts. A minister acts 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. 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 sections in the organisation. (van der Meer 1999, 390-392, 402-404.)

The example of the OECD environmental performance reviews illustrates this interaction between different actors and repertoires and gives clues as to how similar interaction may shape the outcomes and effects of IEA reviews. In the EPRs, two major dividing lines can be distinguished in order to describe the different repertoires. First, repertoires can be distinguished with regard to their emphasis on environmental and economic aspects, respectively. While environmental NGOs and sections of environmental administration working within the ‘traditional’ environmental policy approach can be placed in the environmentalist end of the continuum, the ‘economist’ pole is represented by the core of finance and economy ministries, as well as the economically oriented fractions of environment ministries on the other. The second dividing line is between a view that sees the reviews primarily as a government-led affair led by the bureaucracies (bureaucratic repertoire), and the opposing view of reviews only having impact if given sufficient public and media attention (political repertoire). The environmental NGOs tend to be located in the latter end of the continuum, whereas different parts of public administration – and the review programme itself – are situated at varying locations closer towards the government-led end of the spectrum. Figure 3 depicts the different actors as well as the review approach as well (EPRs) in the space defined by these two axes.

Figure 3. Repertoires of different stakeholders in the EPRs and the preferred pathways of EPR influence.



These elements for distinguishing repertoires remain valid in the case of the IEA, yet other distinctions may be equally important, e.g. the extent to which repertoires are dominated by the belief in centralised as opposed to decentralised energy solutions. The importance of analysing the repertoires becomes clear when seen in light of the question as to which group the reviews ‘empower’. At present, given the relatively government-driven style of the environmental performance reviews, the primary beneficiaries from the reviews are (1) the environmental authorities who can use them in order to legitimise their own work and their own position in battles against more powerful sectors of administration; (2) the economically-oriented sections in environment ministries, helping them to drive for more market-based approaches to environmental policy; and (3) the civil servants in charge of environmental issues in sector ministries, who see the reviews as an opportunity to raise the status of their work within their ministries. A question for the IEA reviews is the extent to which they allow space for actors and views beyond the mainstream energy policy approach to be heard. The analysis of repertoires also allows addressing convergences and divergences between and within the ‘epistemic community’ of international energy policy experts and officials on the one hand, and national energy policy communities on the other. For example, French interviewees pointed at the critical, if not hostile, attitudes towards the OECD that prevailed in the past.⁷ The role of the UK as a

⁷ In an interview concerning the impact of EPRs, a French civil servant stated, perhaps slightly optimistically, that “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.”

frontrunner in energy sector liberalisation, in turn, might make IEA recommendations more palatable to the public and the civil service.

4.2.3 Country-specific institutional context

The OECD faces formidable challenges in dealing with the heterogeneity of its member states in terms of their economic, political, geographic, ecological, cultural conditions, as well as their varying expectations concerning the reviews. It is enough to look at the differences in energy supply choices between countries such as Denmark and France, or energy resource endowments between, say, the US and Luxembourg to appreciate the significance of such differences in energy policy reviews as well.

The relevant country-specific factors can be divided into two groups: fundamental long-term framework conditions and the more ephemeral, short-term background factors. The aspects that most stood out in affecting the impact of the reviews had to do with the national policy style and policymaking structures; socio-economic conditions; changes in the politico-administrative structures; and the position of the potential change agents in the national policymaking structures.

Differences in national policy style and policymaking structures concern countries' attitudes towards outside opinions, the degree of cooperation and networking in their policymaking institutions, their policy style (adversarial vs. consensual), the degree of advancement of their environmental policy structures, and their political discussion cultures (for differences between 'Latin' and 'Anglo-Saxon' discussion cultures, see Gambetta 1998). Furthermore, the degree of policy integration and institutionalisation, as well as the existence or not of a tradition of policy evaluation in the country also affect the way in which reviews influence policies. It has also 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).

Changes in socio-economic conditions, e.g. economic conjunctures and the status of environmental issues on the national policy agenda may affect review impact: on the one hand, recommendations to increase funding for R&D or environmental policies, for instance, might be more difficult to argue for in an economic downturn. On the other hand, the emphasis that the reviews place on cost-effectiveness might actually make the recommendations more useful in times of economic difficulties. Obviously, changes in public awareness concerning matters such as climate change, energy security, or attitudes towards market liberalisation are likely to affect the way in which review messages are being received.

Changes in the governing coalitions and administrative structures during the review may adversely affect their ability to influence policies, by contributing to the loss of 'institutional memory' and continuity if persons responsible for the reviews in the reviewed country are transferred to other tasks. Often, a new government and environment minister is more eager than a well-established one to use the review as a

benchmark against which new policies could be assessed. On the other hand, a new minister may need to be ‘trained’ to understand the value of external critique.

Table 3 presents preliminary findings concerning the impact of country-specific, contextual factors on the potential of the reviews to bring about effects through the three first pathways mentioned above (figure 2). The impact of different contextual factors on pathway D, ‘long-term conceptual learning’, is more difficult to estimate, and has therefore been omitted.

Table 3. The impact of contextual factors on the likelihood of the reviews to bring about effects through the three first 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

- : reduces potential impact

Research on IEA reviews could test the significance of these factors in shaping the impact of the reviews. The variables would need to be modified, in order to take account of the particularities of energy policy. For example, the degree of environmental policy development could be replaced by the degree of advancement in energy efficiency and integration of environmental concerns into energy policies.

4.2.4 Review design

The final element conditioning the impact of peer reviews is the way in which the reviews are designed. In discussions concerning the reviews within the OECD working groups and committees, this is understandably the factor that attracts most

attention. However, in view of the above discussion on the role of the evaluator and the various contextual factors, one should not overestimate the possibilities of enhancing review influence simply by improving their design.

The questions concerning the review design in the EPRs often come down to dichotomous choices, which reflect differing views held by actors on the role of the reviews. The most frequently debated topics can be summarised under the following headings:

Recommendations: to what extent should the reviews aim to give critical and specific recommendations in order to create sufficient pressure, as opposed to ‘softer’, more general recommendations, which leave more room for national implementation and may therefore be more palatable especially on more controversial questions?

Approach to cross-country comparability: how much should the reviews use comparison between countries, e.g. through indicators and statistics showing the ranking of the reviewed country in relation to its peers? Do differences in country-specific aspects proscribe such comparison, and therefore be misleading and politically counterproductive?

Thematic focus: should the reviews focus on a selected number of ‘hot’ and difficult issues or seek to treat all issues in a balanced manner?

Breadth of the review: should the reviews try to focus on a limited selection of key topics, or attempt to cover the entire range of environmental policies?

Focus of analysis: to what extent should the reviews try to understand and explain the causal processes behind policy outcomes, as opposed to simply examining the degree to which the country has achieved its own policy objectives and international commitments?

Participation of stakeholders: should the participation of stakeholders in the review process be broadened and/or deepened in order to increase their legitimacy and ability to create broader debate in society, or should the reviews remain an essentially government-led affair?

There are no clear-cut ‘right’ answers to these questions. They are intimately linked with the differences in the perceived role of the reviews and the preferred pathways of influence, and therefore there is a need to understand actors’ repertoires and national policy contexts. The way in which these questions are answered in the review programme essentially defines where the reviews are located in the ‘repertoire space’ presented in figure 3 above.

4.3 IEA message: what type of framing advocated by IEA?

Since this research project is interested not only in the role of the IEA reviews and the factors conditioning the influence, but in their role in enhancing sustainable energy transitions specifically, it is necessary to look at the policy message – IEA doctrine – transmitted through the reviews. In other words, are the policy approaches and ‘mental models’ that the reviews advocate in line with what could be defined as

sustainable energy transition, or do they instead consolidate the prevailing unsustainable practices, policies and ways of thought?

Answering these questions requires choosing a point of departure, a perspective from which to judge 'which way is forward'. Given the nature of the OECD and IEA as organisations committed to enhancing free trade and economic prosperity, analysing the IEA approach from the perspective of economics is justified. More specifically, institutional economics is here adopted as an approach able better than the mainstream economic theory to deal with the complexity of sustainable development. The analysis of the OECD EPRs concluded that while the EPRs had made some modest approaches towards a more 'institutionalist' perspective, in particular by paying some attention to the procedural aspects of sustainability, and by rejecting the idea of one single efficiency criterion for sustainability, they nevertheless remained solidly anchored in the 'modern mainstream economics'⁸, which arguably lies at the basis of OECD work in general (Dostal 2004, 450-451).

The analysis of the IEA reviews could try to establish to what extent the approach they embody differs from that in the EPRs and examine the strength of the dominant liberal economics doctrine. Indeed, one could expect the IEA reviews to be closer to the OECD mainstream, given the nature of energy policy as a 'high politics' issue close to the core of economic policymaking. There is likely to be a tension between IEA's unreserved adherence to 'free and open markets' on the one hand, and the need for policy intervention to promote sustainable energy policies on the other.

Table 4 presents the list of criteria used to assess the compatibility of the OECD EPR approach with institutional economics. This list could serve as a basis for the analysis of IEA reviews as well, but it would probably need to be modified to account for specificities of energy policy.

⁸ 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 (Boyer and Saillard 1995, 9; Hodgson 1999, 29, 131; Norgaard 1994, 65-66; Pearce and Turner 1990, 10; Williamson 1985). 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. Bartelmus 2000; Mebratu 1998)

Table 4. The basic principles of sustainable development seen from the perspective of institutional economics.

1. Critical reflexivity: reviews should not take premises such as progress, modernity, economic growth and technological development as inherently desirable from the point of view of sustainable development and human well-being, but evaluating their value in the context.
2. Limits to growth. Take into account the limits to the scale of human activity posed by the biophysical life-support mechanisms, and identify ostensibly unsustainable situations for instance through concepts such as critical natural capital, ecosystem functions, resilience, as well as irreversible, synergistic and cumulative effects.
3. Procedural side of sustainable development. 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, in particular by empowering the weaker groups.
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. Measurement of sustainable development. 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.

In moving beyond this general list, insights from the literature on socio-technical literature could be useful in identifying ‘ideational’ factors that affect the capacity of energy reviews to operate in favour of desirable transitions, unlock undesirable lock-ins and empower change agents. Moreover, the analysis should endeavour to identify the meanings given by the IEA to its central concepts such as flexibility, diversity, efficiency, and energy security. This work would need to move between the analysis of the IEA Shared Goals, country review reports, and other IEA documents.

5 Discussion

It is clear that ‘soft’ mechanisms of governance alone are insufficient to bring about a change towards sustainable energy policies, yet experience from OECD peer reviews suggests that reviews can have impacts, in particular by shaping the mental models and frameworks of thought. Depending on the country-specific context, expectations and ‘repertoires’ of key actors, and the more short-term changes in political conditions, these conceptual changes may also have more lasting policy impacts. This paper has suggested a framework for analysing the role of IEA energy policy reviews in promoting (or preventing) sustainable energy transitions. The strength of this framework is that it allows comparisons with other OECD reviews, notably the Environmental Performance Reviews, and gathering much needed empirical evidence as to whether soft mechanisms actually deliver on their promises.

While the IEA and OECD, being part of the same OECD ‘family’, share a common policy approach, the position of the IEA reviews in the organisation’s work greatly differs from that of the EPRs in OECD work. The IEA reviews are clearly in the mainstream of the organisation’s activities, whereas the EPRs are marginal to OECD work, and must balance their message between environmental ‘advocacy’ and faithfulness to the OECD core doctrine to maintain their credibility. From the perspective of sustainability transitions, an interesting question in comparing IEA reviews with the EPRs is the capacity of the respective processes to ‘mainstream’ environmental concerns into ‘harder’ policy sectors. While EPRs constantly suffer from their marginal position in OECD work, the IEA reviews may by contrast have the opposite problem of being too close to the core of the dominant energy policy regime, and therefore incapable of triggering change. Struggles around IEA reviews are likely to be between those advocating centralised as opposed to decentralised energy systems, and take place within the energy expert community, rather than between different sectors.

A number of practical questions need to be solved before the suggested framework is ready to be applied into empirical research. Obtaining reliable information concerning the impacts of ‘soft’ governance is by no means simple, not least because the impacts often take place after a considerable time lag, which makes establishing causal links between a review and policies highly challenging. Yee (1996) has suggested three strategies for analysing such impacts. Co-variational analysis would try to identify correlation between the development of ideas and the development of policies, congruence procedure would seek congruence between ideas and policies to establish causality, while process tracing would involve studying a specific policy process in depth. The research on IEA reviews would seek to combine these strategies, in order to minimise their respective shortcomings. This would involve in-depth case studies of a limited number of countries, to gain better understanding of the pathways of influence, combined with more general analysis of the review programme and IEA work to allow cross-country comparisons.

6 References

- Anon. 2003 14 May. Englannin kieli ei tee ajatuksista tuoreita, Editorial, 14 May. Helsingin Sanomat.
- Armingeon K. 2003. OECD and national welfare state development. In: Armingeon K, Beyeler M, editors. *The OECD and European Welfare States*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar. p 226-241.
- Armingeon K, Beyeler M, editors. 2003. *The OECD and European Welfare States*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar.
- Arrowsmith J, Sisson K, Marginson P. 2004. What can 'benchmarking' offer the open method of co-ordination? *Journal of European Public Policy* 11(2):311-328.
- Barbier J-C. 1999. Inter-governmental evaluation: Balancing Stakeholders' Expectations with Enlightenment Objectives? *Evaluation* 5(4):373-386.
- Baron G. 1999. Evaluation, participation, apprentissage: Une conception de l'action public avec rationalité limitée. Analyse et étude de cas d'une évaluation d'un programme européen [Doctoral dissertation]. Rennes: Université de Rennes I. 440 p.
- Bartelmus P. 2000. Sustainable development: paradigm or paranoia? *International Journal of Sustainable Development* 3(4):358-369.
- Beyeler M. 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.
- Borrás S, Greve B. 2004. Concluding remarks: New method or just cheap talk? *Journal of European Public Policy* 11(2):329-336.
- Borrás S, Jacobsson K. 2004. The open method of co-ordination and new governance patterns in the EU. *Journal of European Public Policy* 11(2):185-208.
- Boyer R, Saillard Y. 1995. *Théorie de la regulation: état des savoirs*. Paris: Edition la découverte. 568 p.
- Bruno I, Jacquot S, Mandin L. 2006. Europeanization through its instrumentation: benchmarking, mainstreaming and the open method of co-ordination ... toolbox or Pandora's box? *Journal of European Public Policy* 13(4):519-536.
- Clark WC, Dickson N. 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*:6-7.
- Davies IC. 1999. Evaluation and Performance Management in Government. *Evaluation* 5(2):150-159.
- de la Porte C, Nanz P. 2004. The OMC - a deliberative-democratic mode of governance? The cases of employment and pensions. *Journal of European Public Policy* 11(2):267-288.
- de la Rosa S. 2005. The Open Method of Coordination in the New Member States - the Perspectives for its Use as a Tool of Soft Law. *European Law Journal* 11(5):618-640.
- Dostal JM. 2004. Campaigning on expertise: how the OECD framed EU welfare and labour market policies - and why success could trigger failure. *Journal of European Public Policy* 11(3):440-460.
- Eberlein B, Kerwer D. 2002. Theorising the new modes of European Union governance, *European Integration online Papers (EIoP)* 6.

- Eckley N. 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. Copenhagen. Report nr 26.
- Edquist K. 2006. EU social-policy governance: advocating activism or servicing states? *Journal of European Public Policy* 13(4):500-518.
- EEA. 2005. *The European Environment - State and Outlook*. Copenhagen: European Environment Agency.
- EU. 2000. Council of the EU. Note on the ongoing experience of the Open Method of Coordination, Portuguese Presidency Note, 9088/00, 14/06/00.
- Finnemore M, Sikkink K. 1998. International Norm Dynamics and Political Change. *International Organization* 52(4):887-917.
- Forss K, Reubien CC, Carlsson J. 2002. Process Use of Evaluations: Types of Use that Precede Lessons Learned and Feedback. *Evaluation* 8(1):29-45.
- Foulkes G. 1998. Evaluation and International Development: A British Perspective. *Evaluation* 4(3):359-362.
- Gambetta D. 1998. "Claro!" An Essay on Discursive Machismo. In: Elster J, editor. *Deliberative Democracy*. Cambridge: Cambridge University Press. p 19-43.
- Giddens A. 1984. *The Consequences of Modernity*. Cambridge: Polity Press. 186 p.
- Greene JC. 1999. The Inequality of Performance Measurements. *Evaluation* 5(2):160-172.
- Haas PM. 1992. Introduction: epistemic communities and international policy coordination. *International Organization* 46(1):1-35.
- Hanberger A. 2001. Policy and Program Evaluation, Civil Society, and Democracy. *American Journal of Evaluation* 22(2):211-228.
- Henry GT, Mark MM. 2003. Beyond Use: Understanding Evaluation's Influence on Attitudes and Actions. *American Journal of Evaluation* 24(3):293-314.
- Hodgson GM. 1999. *Evolution and Institutions: On evolutionary economics and the evolution of economics*. Cheltenham, UK and Northampton, MA: Edward Elgar. 345 p.
- Kaiser R, Prange H. 2004. Managing diversity in a system of multi-level governance: the open method of co-ordination in innovation policy. *Journal of European Public Policy* 11(2):249-266.
- Keohane RO. 1978. The International Energy Agency: State Influence and Transgovernmental Politics. *International Organization* 32(4):929-951.
- Lehtonen M. 2005a. *Environmental Policy Evaluation in the Service of Sustainable Development: Influence of the OECD Environmental Performance Reviews from the Perspective of Institutional Economics* [Unpublished doctoral dissertation]. Guyancourt: Université de Versailles Saint-Quentin-en-Yvelines. 335 p.
- Lehtonen M. 2005b. OECD Environmental Performance Review Programme: Accountability (f)or Learning? *Evaluation* 11(2):169-188.
- Lodge M. 2005. The importance of being modern: international benchmarking and national regulatory innovation. *Journal of European Public Policy* 12(4):649-667.
- López-Santana M. 2006. The domestic implications of European soft law: framing and transmitting change in employment policy. *Journal of European Public Policy* 13(4):481-499.

- Mabry L. 2002. Postmodern Evaluation - or Not? *American Journal of Evaluation* 23(2):141-157.
- March JG, Olsen JP. 1998. The Institutional Dynamics of International Political Orders. *International Organization* 52(4):943-969.
- Marcussen M. 2003. Multilateral surveillance and the OECD: playing the idea game. In: Armingeon K, Beyeler M, editors. *The OECD and European Welfare States*. Cheltenham, UK and Northampton, MA, USA: Edward Elgar. p 13-31.
- Mebratu D. 1998. Sustainability and Sustainable Development: Historical and Conceptual Review. *Environmental Impact Assessment Review* 8(6):493-520.
- Meyer CO. 2004. The hard side of soft policy co-ordination in EMU: the impact of peer pressure on publicized opinion in the cases of Germany and Ireland. *Journal of European Public Policy* 11(5):814-831.
- Norgaard RB. 1994. *Development Betrayed: the end of progress and a coevolutionary revisioning of the future*. London and New York: Routledge. 280 p.
- OECD. 2003. *Peer Review: An OECD Tool for Co-operation and Change*. Paris: OECD. 54 p.
- Owens S, Rayner T, Bina O. 2004. New agendas for appraisal: reflections on theory, practice, and research. *Environment and Planning A* 36(11):1943 - 1959.
- Patton MQ. 1998. Discovering Process Use. *Evaluation* 4(2):225-233.
- Pearce DW, Turner RK. 1990. *Economics of natural resources and the environment*. New York, London, Toronto, Sydney, Tokyo and Singapore: Harvester Wheatsheaf. 378 p.
- Perrin B. 1998. Effective Use and Misuse of Performance Measurement. *American Journal of Evaluation*. *American Journal of Evaluation* 19(3):367-379.
- Perrin B. 2002. How to - and How Not to - Evaluate Innovation. *Evaluation* 8(1):13-28.
- Pochet P, de la Porte C, Room G. 2001. Social benchmarking, policy making and new governance in the EU. *Journal of European Social Policy* 11(4):291-307.
- Przeworski A. 1998. Deliberation and Ideological Domination. In: Elster J, editor. *Deliberative Democracy*. Cambridge: Cambridge University Press. p 140-160.
- Radaelli C. 2000. Discourse and Institutional Change: the case of Italy in the Euro-Zone. *Queen's Paper on Europeanisation* No 5/2000.
- Room G. 2005. Policy Benchmarking in the European Union: Indicators and Ambiguities. *Policy Studies* 26(2):117-132.
- Sabatier PA. 1987. Knowledge, Policy-Oriented Learning and Policy Change: An Advocacy Coalition Framework. *Knowledge: Creation, Diffusion, Utilisation* 8(4):649-692.
- Schäfer A. 2006. A new form of governance? comparing the open method of co-ordination to multilateral surveillance by the IMF and the OECD. *Journal of European Public Policy* 13(1):70-88.
- Segerholm C. 2003. Researching Evaluation in National (State) Politics and Administration: A Critical Approach. *American Journal of Evaluation* 24(3):353-372.
- Söderbaum P. 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.
- Stokes S. 1998. Pathologies of Deliberation. In: Elster J, editor. *Deliberative Democracy*. Cambridge: Cambridge University Press. p 123-139.

- Strang D, Chang PMY. 1993. The International Labour Organisation and the Welfare State: Institutional Effects on National Welfare Spending, 1960-1980. *International Organization* 47(2):235-262.
- Taut S, Brauns D. 2003. Resistance to Evaluation: A Psychological Perspective. *Evaluation* 9(3):247-264.
- Valovirta V. 2002. Evaluation Utilization as Argumentation. *Evaluation* 8(1):60-80.
- Valve H. 1999. Frame Conflicts and the Formulation of Alternatives: Environmental Assessment of an Infrastructure Plan. *Environmental Impact Assessment Review* 19(2):125-142.
- Van der Meer F-B. 1999. Evaluation and the Social Construction of Impacts. *Evaluation* 5(4):387-406.
- Vedung E. 1991. *Utvärdering i politik och förvaltning (Evaluation in politics and administration)*. Lund: Studentlitteratur. 239 p.
- Vedung E. 2001. Utvecklingsdrag i utvärdering (Development Trends in Evaluation). *Hallinnon tutkimus (Administrative Studies)* 20(3):139-143.
- Weiss CH. 1980. Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilisation* 1(3):381-404.
- Weiss CH. 1998. Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation* 19(1):21-33.
- Weiss CH. 1999. The Interface Between Evaluation and Public Policy. *Evaluation* 5(4):468-486.
- Williamson OE. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York and London: The Free Press. 450 p.
- Wimbush E, Watson J. 2000. An Evaluation Framework for Health Promotion: Theory, Quality and Effectiveness. *Evaluation* 6(3):301-321.