TABLE OF CONTENTS

I. TOPIC REPORT 1: EVALUATION PRACTICES

EXECUTIVE SUMMARY

This report describes the result of the analysis of 16 national reports on the evaluation practices with regard to Energy RTD in the European Member States plus Norway, but with the exception of Luxemburg. This analysis has been carried out within the framework of the project SENSER (Synergies between European and National Strategies for Energy RTD), funded in part by the European Commission. The following conclusions have been drawn from this study.

GENERAL OVERVIEW OF ANALYTICAL RESULTS

- evaluation of NNE-RTD is **not accountancy but a strategic activity**. Although justification of public expenses and assessment of goal fulfilment are normally the main initiators of an (ex post) evaluation exercise, evaluation normally goes beyond the verification of initial objectives, as evaluation results serve to set new priorities and to redefine research programmes;
- most of the research in the Countries under scrutiny is organised into **research programmes**. Evaluation of NNE-RTD is usually performed around such research programmes;
- in most countries the tendency is moving from more ad hoc NNE-RTD evaluation practices **toward a certain standardisation of evaluation practices**, even if these differ from one country to another. Apart from the countries in which no NNE-RTD evaluation is carried out at all (Italy, Belgium), only in Denmark is formalisation of evaluation practices not explicitly sought;
- there is a **very unequal experience** with NNE-RTD evaluation across countries. UK, France and Finland have a tradition dating back to the beginning of the 1980s if not earlier, whereas countries such as Austria or Greece are just starting;
- **political levels,** on behalf of which evaluation is carried out, **are different** across countries; the national (government) level appears to be the most important, sometimes through an agency. The longer the tradition in evaluation a country has, the more varied the levels are at which it is carried out;
- evaluation revolves mainly around:
  - definition of the entire national non-nuclear energy Programme: mainly ex ante
  - assessment of research programmes: mainly ex post
  - projects: both ex ante (project selection) and ex post
  - but (to a far lesser extent) institutes, industrial sectors, research areas, procedures, technologies or partnerships;
- with very few exceptions, **evaluation of NNE-RTD relies on peer review** and expert opinion and is scarcely quantified;
- the issues addressed by NNE-RTD evaluation are (exhaustive list according to frequency of appearance in sum of national reports):
  1. scientific and technical quality
  2. programme operation
  3. market opportunities
  4. collaborations
  5. users
  6. energy issues
  7. relation with other sectoral policies

The first four issues are addressed by all countries. The last three issues are addressed to a marginal extent. The relationship with European policies or research programmes is not explicitly mentioned as being an issue addressed by NNE-RTD evaluation at national levels;

- ex post evaluation normally addresses the first four issues, and the last three to a lesser extent; ex ante evaluation addresses issues 1 and 3, and sometimes 6 & 7. With three exceptions, monitoring focuses on programme operation only (issue 2);
- the types of parties which carry out evaluation are as follows:

  - ex ante: by public authorities and government, managing agencies, sometimes by independent expert panels but rarely by consultancies only; often performed through specially established ad hoc working groups involving several of the aforementioned parties;
  - ex post: most often by expert panels, sometimes by consultancies or managing agency;
  - monitoring: mostly performed by managing agency or public authority involved;
- there are **no explicit links between Technology Characterisation and Foresight (cf. Topic 2) and evaluation**. TC&F results are among the elements mobilised, together with the results of ex post evaluation or market assessments, within *ex ante* evaluation.

RECOMMENDATIONS

Following the analysis of the country reports, it is recommended that the priority in the further development of evaluation concepts and methodology should not be put on scientific-technical assessment and on the assessment of programme operation. These seem rather well developed at both Member States and European levels and, without losing the current level of quality, can probably only be marginally improved. They simply 'have to be done.'

Instead, it is proposed that further improvement of evaluation practices beneficial at both the European and Member State levels can only come if it is made subject to the following six conditions. A new evaluation approach should simultaneously:

1. be able to assess socio-economic benefits of research. Socio-economic 'indicators' should be of equal importance to evaluation as indicators for scientific and technical quality and programme operation;
2. integrate foresight and evaluation, or more broadly, 'backward-looking' and 'forward-looking' activities, without problems of continuity;
3. incorporate a type of monitoring which looks beyond programme operation only;
4. make use of 'good practices' of Member States and those used in European programmes;
5. take into account both European and Member States' energy RTD and energy and non-energy policies;
6. be able to take into account, again without problem of continuity, the short, medium and long term.

1. INTRODUCTION

This report describes the result of the analysis of 16 national reports on the evaluation practices with regard to energy RTD in the European Member States plus Norway, but with the exception of Luxembourg. This analysis has been carried out within the framework of the project SENSER (Synergies between European and National Strategies for Energy RTD), financed in part by the European Commission.

The results given in this report come on top of the national reports produced in December 1996. Together these are to be discussed at the peer review workshop. After this workshop, the energy RTD evaluation practices will be related to the results of the other three Topic Areas contained in the SENSER project.[1] Also indicators – the anticipated final result of the SENSER project – will be obtained only after sharing of and discussion on the results of the different topics areas within the framework of this workshop.

2. METHOD

In April 1996 an 'instructions document' with examples was prepared by SENSER's Topic Leaders, including one chapter on evaluation of energy RTD.[2] This document was discussed at a two-day meeting in Brussels involving members of all national teams ('launch meeting', April 22-23, 1996). The structure of the evaluation part of the instructions report is given in Box 2.1. Also, an example of the 'datasheet' which national teams were asked to fill in on the level of individual evaluations is annexed to the present report.

As far as the Topic Area 'evaluation' is concerned, the instructions document appeared to give enough basis to make a satisfactory national report. Most countries were able to follow the instructions. Exceptions were Italy and Belgium-Flanders, since no evaluation is systematically conducted in these countries, and Belgium-Wallonia and Denmark, for having only a very weakly institutionalised evaluation practice.

The (drafts of) a subset of national reports were analysed in a preliminary fashion, the results of which were presented to all responsible for the national reports, in order to obtain feedback, agreement on interpretation, supplementary information and amendments where necessary.[3] This discussion made possible a revised analysis resulting in the present report.

UNLESS STATED OTHERWISE,
'EVALUATION' IN THIS REPORT REFERS TO EVALUATION OF ENERGY RTD

[1] Technology Characterisation and Foresight (Topic 2), Market developments (Topic 3), and the Organisation of and expenditures for national E-RTD (Topic 4).
[2] SENSER. Instructions for national teams, May 1996.
[3] Meeting of October 15, 1996, summarised in 'Emerging Trends' report (Olthof, de Laat, Smith, Virdis).

3. STATUS: EVALUATION PRACTICES DIFFER

A first important observation is that in some countries evaluation of E-RTD (energy RTD) has a long-standing tradition whereas in others it is simply absent. In all but one of the cases in which evaluation of E-RTD exists, the extent to which evaluation practices are formalised and institutionalised seems to relate directly to the tradition in E-RTD a country has. Finally, the level as well as the multitude of levels on which evaluations are carried out differ, from country to country.

Results are given in Table 1, and will be briefly discussed afterwards.

Table 1. Status of E-RTD evaluation in Member States

Member State[4]	***Experience since...***	***Institutionalisation & Formalisation***	***Political levels at which evaluation is carried out***
Austria	starting	weak (1)	government
Belgium - Flanders	no evaluation	absent	-
Belgium - Wallonia	no evaluation (2)	absent	-
Denmark	2nd half 1980s	weak	national
Finland	2nd half 1980s	strong	programme board; national
France	beginning 1980s	strong	agency; operator
Germany	recent	weak	national government & ministries; länder (3)
Greece	recent (>1990)	weak	national
Ireland	no evaluation	absent	-
Italy	no evaluation	absent	-
Netherlands	since 1980s	strong	agency; ministries; research councils; programme
Norway	since 1980s	strong	research council; programme
Portugal	beginning 1990s	weak	national
Spain	recent	weak	operator; national
Sweden	late 1980s	strong	government; agency; programme
United Kingdom	1970s	strong & regular basis	government; agencies, operators

(1) there are few specific E-RTD research programmes in Austria.
(2) E-RTD falls under general research programmes so no specific E-RTD evaluations exist. For general research however, the Research Directorate (DGTRE) evaluates projects' eligibility criteria; there exists a general Orientation Committee which monitors programmes.
(3) in Germany, formalised or officially required procedures for the evaluation of federal RTD programmes in the energy sector for which BMBF is responsible do not exist; the Bundesländer's evaluations were unfortunately not included in the national report (they mainly concern programmes aiming at market introduction, not pure RTD).
(4) in Ireland E-RTD is carried out on an extremely small scale so no evaluation could be identified, apart from project evaluations within the framework of selection procedures.
(5) According to the national report an informal ex ante evaluation-type of activity exists in Italy but, '*...one must bear in mind that choice of the strategic areas of energy R&D in Italy is rarely based on, or supported by, an in-depth and systematic work of technology assessment the way it is done in other countries.*' The national report has thus been interpreted as there being no evaluation practice in Italy as meant in the SENSER project.

The heading '*Experience since...*' refers to the tradition of E-RTD evaluation existing within a country, i.e. from which period onwards has evaluation of Energy RTD become regular practice in the Member State of concern. The heading '*institutionalisation and formalisation*' refers to *how* evaluation in a Member State is established as

[4] Norway is not a member state but nevertheless we will use the term in this report.

a policy instrument, i.e. is it a routine matter, do general guidelines, rules or standards exist for evaluation and for organising the evaluation process, are specific bodies responsible for performing evaluation etc. The heading 'level' finally relates to the levels at which evaluation is carried out. The latter issue will be dealt with more thorougly below (cf. the 'units of evaluation,' 'parties involved,' and so forth).
Note that although they can be related, 'experience' does not necessarily equal 'formalisation and institutionalisation': evaluation can be of long-standing tradition in a Member State but this does not necessarily mean that formal, standardised evaluation methods or guidelines are established along which evaluations are carried out. This is, for instance, the case in Denmark, where energy RTD has been evaluated for nearly 10 years now, but nevertheless a '*standard programme for evaluation of Energy RTD does not exist*' according to the national report.

3.1 MATURITY

Table 2. First structured E-RTD evaluations in Member States

1980	85	90	95	97+
(<=UK), F, N, NL	DK, SF, S	A, Bw, D, E, GR, P		(I, IR, Bf=>?)

Maturity of NNE-RTD in the different Member States also differs. In Norway, France, the Netherlands and the UK, evaluation of non-nuclear energy RTD has a fairly long tradition. They are closely followed by the Scandinavian countries Denmark, Sweden and Finland who started evaluation in the second half of the 1980s. Inversely, there are about the same number of Member States in which evaluation has only recently become a regular practice (Austria, Greece, Portugal, Spain) and/or in which it is only a weakly developed activity (e.g. in Germany at federal level). Finally, in some Member states, there is no evaluation at all (Belgium, Italy).
In the cases where a more long-standing tradition of E-RTD evaluation exists, evaluation appears to have emerged at about the same time as the establishment of the first structured research programmes and/or national agencies (e.g. France) in the energy field, and seems strongly connected to this evolution. In the countries which have been performing evaluation more recently, a direct link between establishment of agencies, programmes and evaluation could not be identified.

3.2 FORMALISATION & INSTITUTIONALISATION

Table 3. Institutionalisation of E-RTD evaluation

absent	weak	strong
Bf, I, IR	Bw, A, D, DK, E, GR, P	FR, N, NL, S, SF, UK

The degree of formalisation and institutionalisation of evaluation practices and procedures appears to differ enormously across the different countries. For instance, Finland has a long historical record of evaluation and well-established procedures and protocols. In 1990, this country established the *WGEERP* (Working Group for the Evaluation of Energy Research Programmes) the aim of which was to produce *guidelines* along which the evaluation of RTD, including energy RTD, should be carried out.[5] In contrast, in Ireland or Italy no energy RTD evaluation is said to be performed at all, and in Belgium it is weakly developed. In Norway and the Netherlands evaluation is institutionalised at all kinds of different levels, whereas in Greece evaluation is only recently becoming institutionalised, and for the moment it is initiated and performed almost exclusively on a national level by the GSRT (General Secretariat for Research and Technology). A general tendency is toward more formalisation, institutionalisation and standardisation of evaluation procedures in the countries covered in the present study.

[5] See the table in the Finnish National Report explaining these guidelines.

3.3 LEVELS ON WHICH EVALUATION IS CARRIED OUT

The way in which energy RTD evaluation is carried out appears to relate strongly to the following four issues:

- political organisation of a country;
- the organisation of the E-RTD;
- the maturity of E-RTD evaluation;
- the type of evaluation.

As far as **political organisation** is concerned, in Germany for instance the BMFT only carries out evaluations relating to the federal level, the *Länder* having their only evaluation practices for their own research. Consequently, the latter do not appear in the national report. In this respect, Belgium can also be mentioned as an example. The fact that this country is divided into different autonomous regions, ensures that these regions do not evolve equally in so far as NNE-RTD (non-nuclear energy RTD) and its evaluation are concerned. In the Wallonia region NNE-RTD evaluation is now visibly developing, whereas in Flanders this is not yet the case. Also because of a regional structure, but for the remainder completely different from Belgium, France is a good example to illustrate the relations between political organisation and organisation of evaluation of NNE-RTD. The national energy agency ADEME has a structure which consists of, apart from a central body, regional delegation, corresponding to France's 22 regions. The Eval-system used by the agency (cf. national report) allows it to conduct a bottom-up evaluation and monitoring of energy (and environment) research projects and programmes from the regional to the national level, *via* the agency.

Evaluation of energy RTD is also strongly dependent on the **organisation of energy RTD**. It notably appears to depend on the number of different bodies which co-ordinate energy RTD, since this multiplies the number of levels on which evaluation is carried out.[6] Although this phenomenon can be observed in all countries, the Netherlands might be the most striking example. Even though the co-ordination of energy research is performed by the Ministry of Economic Affairs in collaboration with Novem (the national energy agency), non-nuclear energy RTD is further distributed over a huge set of other bodies.[7] The evaluation of energy RTD is performed correspondingly, i.e. it is initiated by these different bodies. This situation, in which evaluation gets multiplied to such an extent and takes place in different (at least three) and relatively separated circuits has recently been identified as needing eventual reconsideration.

The way in which and the levels on which evaluation is carried out depends on the **tradition of evaluation** which exists and the experience acquired with it. The more long-standing the tradition in evaluation, the 'lower' the levels are on which it is carried out.[8] In countries in which evaluation is only a recent practice it appears mainly to be initiated from central government level (e.g. Portugal, Greece), whereas in countries with a longer experience, evaluation will be found 'deeper down' i.e. at the agency and programme level (e.g. Finland, France, Netherlands). Exceptions to this rule exist: in Denmark, for reasons given earlier; and in Germany, since although no general evaluation framework exist, projects are individually evaluated, though this evaluation does not seem to amount to a formal 'programme' or 'national' evaluation.

Finally, the level at which evaluation is carried out also strongly depends on the **type of evaluation** performed. Whereas ex post evaluation and monitoring (see below), are carried out at different levels, the same conclusion is not valid for ex ante evaluation in most of the cases: the UK for instance has the most long-standing tradition in E-RTD evaluation of all, but ex ante evaluation is always carried out on a national level and initiated by government. Also in France, ex ante evaluations in energy RTD are co-ordinated at a central level, i.e. between the national agency and the responsible ministries.

[6] This is not evident a priori since one could very well imagine situations with, for instance, decentralised energy research but a centralised evaluation of this research.

[7] Ministry of Environment, Ministry of Education and Research, several thematic research councils, the academic and technological research councils, as well as several research institutes which each have their own and sometimes several responsible ministries.

[8] 'Levels' here relate to the 'organisational hierarchy' within the national energy RTD system. A 'high' level refers to national government bodies, a 'low' level refers to research institutes or even projects.

3.4 ONE COMMON ELEMENT: THE RESEARCH PROGRAMME

There is one recurrent element in the description of evaluation processes: the existence of **national research programmes**. This is an interesting feature since national research programmes are a relatively recent (i.e. not older than 15, 20 years) way of co-ordinating national research. Apparently though, they have come to be established in *all* European countries as a major way of organising energy RTD. Not surprisingly then, they also turn out to be the main unit around which evaluation revolves.

3.5 CONCLUSIONS

- In most Member States the tendency is moving from more ad hoc NNE-RTD evaluation practices toward a certain standardisation of evaluation practices, even if these differ from one country to another. Apart from the countries in which no NNE-RTD evaluation is carried out at all (Italy, Belgium), the only country where no explicit formalisation of evaluation practices seems to be sought is Denmark.
- The development of NNE-RTD evaluation practice is very unequal across countries. France and Finland have a tradition which seems to date back to the beginning of the 1980s if not earlier, whereas countries such as Austria or Greece are only just starting.
- The political levels on behalf of which evaluation is carried out also differs from country to country, although the national (government) level appears to be the most important one, sometimes through an agency.
- Evaluation is often performed at the level of research programmes, which have become a popular way of organising energy RTD in Member States at both national and other levels.

4. WHY AND HOW EVALUATION IS USED

4.1 THREE TYPES OF EVALUATION: DEFINITIONS FROM THE 'INSTRUCTIONS' DOCUMENT

In the instructions document, a distinction between ex post evaluation, ex ante evaluation and monitoring was made.

Table 4. Different types of evaluation - definitions from the instructions report

Item	***Definition from instructions document***
evaluation ex post	...concerns the process of checking how a **completed** set of (research) actions (project, programme, organisation ...) has met a set of given objectives (or not).
Evaluation ex ante	... concerns the process of identifying benefits, defining rationale, setting objectives, examining options and weighing up the costs and benefits **before** a set of (research) actions (project, programme, organisation ...) starts.[9]
Monitoring	... concerns the regular checking of **ongoing** research actions.

With the exception of Sweden, this threefold distinction applies in all other countries, though with somewhat different meanings and, notably, different terms used across countries (these are summarised in Annex 1). Most of the countries emphasise that a difference be made between the different types of evaluation and that it is important to take their relationships into account. Then, from a general point of view, ex ante evaluation is often used to define **new phases in a programme**, if not a whole new programme. In order to do this, it often relies on data and information on previous actions coming from, *inter alia*, ex post evaluations (but also from TC&F, market assessments and more implicit judgements). Monitoring is mostly close to the project level and focuses on administrative, not strategic, matters.[10]

4.2 SIMILAR REASONS TO CONDUCT AN EVALUATION

An overview of the general reasons for evaluation is given, by Member State, in Table 5.

Table 5. Why evaluation is used

Member State	***Why evaluation is used***
Austria	reorganisation of existing projects; input into new programmes
Belgium - Flanders	-
Belgium - Wallonia	-
Denmark	quality; implementation; goal fulfilment; organisation and administration; future programmes

[9] Following the discussion at the launch meeting, the difference between evaluation ex ante here and technology foresight and characterisation of chapter 4 lies in the fact that evaluation ex ante leads in the end to the definition of a set of research actions, whereas technology foresight and characterisation do not necessarily. In some cases there might indeed be overlap as has been suggested at the launch meeting.

[10] Finland is an exception to this since they use a special type of monitoring, called 'interim evaluation' or 'evaluation during implementation.' It is conducted at regular intervals during the course of research programmes and focuses not only on administrative matters, as in most other countries, but integrates aspects of more in-depth ex post and ex ante evaluation, in order to make steering during programme implementation possible.

Member State	*Why evaluation is used*
Finland	to assess impact and significance of RTD; to compare S&T level with international standards; to assess social and energy impact, and impact on technological development and innovation; to evaluate a programme's progress, productivity, results and quality
France	justification of public expenses; assessment of research results; assessment of establishment, maintenance and evolution of national RTD-capacity and structures
Germany	-
Greece	develop understanding / provide information and data for decision-making and selection of policies, mainly by national agencies
Ireland	-
Italy	-
Netherlands	research programming; improve effectiveness of RTD efforts with regard to policy goals; organisational transparency
Norway	justification and control of government expenditures; learning and motivating by feedback into research environment; steering (goal fulfilment, benefits and risks, socio-economic return on investments)
Portugal	identify benefits; examine options; set precise objectives; integration of national policies in general EU policies
Spain	assessment of implementation of RTD-programmes
Sweden	learning for improvement of strategy, cost-effectiveness and efficiency; support for strategic decision-making; assessment of outcomes, impacts, achievement of objectives; reinforcing, future priority setting
United Kingdom	objectives attained, value for money, quality of work and delivery, effectiveness (ex ante); direction and priority of RTD actions, potential benefit to cost, rationale and objective (ex post); technical progress, continuing merits, quality and performance of management, attainment of targets/milestones (monitoring)

The general purpose of evaluating energy RTD is expressed very similarly across countries. First of all it concerns direct justification of public expenses and therefore focuses on goal achievement, implementation and the management of the unit being evaluated. However, evaluation is also a highly strategic activity since all countries mention that it helps to define new research actions. Hence it aims at learning from past experience and assessing funded RTD in relation to a country's broader RTD, and socio-economic factors.

The reasons for specifically initiating an evaluation which are most frequently mentioned in national reports, are the following:

I. A 'unit' (see below) comes to its end; in this case, the following three reasons are mentioned for conducting an evaluation:
   - A. assessment of goal fulfilment
   - B. justification and verification of public expenses
   - C. definition of new action

II. A legal regulation exists
   - A. a legal regulation already exists (N, SF, NL, FR, ...)
   - B. a legal regulation is in the making (E, GR, P, UK, ...)

An overview of reasons for initiating an evaluation is given in Table 6. This table is of course a highly aggregated result - empty cells do not mean by definition that this reason for initiating has never been applied in a country, only that they do not appear as such in the national report.

Table 6. Specific reasons for initiating an evaluation, explicitly mentioned

	A 'unit' comes to its end			***(Legal) regulation***		
Member State	***goal fulfilment***	***justification of funding***	***new action***	***absent***	***exists***	***in prep.***
Austria			x	x		
B - Flanders				x		
B - Wallonia					x	
Denmark	x		x	x		
Finland	x		x		x	
France	x	x	x		x	
Germany			x	x		
Greece						x
Ireland				x		
Italy				x		
Netherlands	x	x	x	x		
Norway	x	x			x	
Portugal			x			x
Spain	x					x
Sweden	x				x	
U Kingdom	x	x	x			x

Table 7, below, describes where the different legal prescriptions (in the countries where one exists), are applicable.

Table 7. Legal or other official prescriptions to evaluation of energy RTD[11]

Member State	***The (legal) prescription relating to NNE RTD evaluation***
B - Wallonia	a decree (1991) for evaluation of publicly funded R&D programmes in all fields including energy RTD
Finland	state council (1987) for research and technology suggested that especially sectoral activities be evaluated, followed up by the WGEERP's guidelines for energy RTD evaluation
France	in the managing agency's 'contrat d'objectifs' with the State (1995), declares that an evaluation of the geothermy area be conducted, as well as several other specific evaluations. It is also explicitly stated that the agency conduct a 'permanent policy of internal evaluations of its interventions'
Norway	all projects financed by the NFR are officially required to deliver cost statements

It is obvious that, if a country has a legal prescription to perform evaluation, a whole process has already been gone through which in the end led to the decision to convert evaluation of energy RTD (but in most cases of RTD in general) into legislation. This also means that in all cases the way in which evaluation has been translated into legal terms will be country-specific. For instance, in one country the legal prescription might only concern a general financial control by a National Accounting Chamber, whereas in another it might well incorporate the specifics of R&D. A detailed analysis of these issues lies beyond the scope of the SENSER project but would be worthwhile to perform, as well as an assessment of the pros and cons of legally specifying evaluation in general.

[11] Does not include financial evaluations performed by national accounting chambers, which are sometimes referred to in the national reports.

4.3 THE TYPES OF [UNIT] EVALUATED

In the instructions report, the 'unit of evaluation' was defined as follows:

« Evaluation normally applies to defined entities: a research programme, a technological institute, etc. We will call these entities the **units of evaluation** ».

Table 8 shows, per country, which type of units are evaluated.

Table 8. The types of units evaluated

Member State	***Types of Units evaluated***								
Unit:	Pg	pg	pj	ins	sec	area	prcd	tec	par
Austria		x		x	x			x	
B - Wallonia	x	x	x						
Denmark		x							
Finland	x	x	x	x		x			x
France	x	x	x	x	x	x	x	x	x
Germany		x	x						
Greece	x	x	x	x	x				
Ireland			x						
Netherlands	x	x	x	x		x		x	
Norway		x	x	x		x			
Portugal	x	x		x					
Spain	x	x	x	x			x		
Sweden	x	x	x	x		x		x	
United Kingdom	x	x	x			x			

Legend:
Pg = entire national non-nuclear energy programme
pg = a specific research programme
pj = research project or operation
ins = institute
sec = a given industrial sector
area = a given research field/area
prcd = procedure
tec = a given technology
par = partnership(s)
The table is exhaustive as concerns the different types of 'units' mentioned in the national reports.
Flanders and Italy have been left out of the table, since no systematic evaluation is conducted.

The research programme appears to be the main type of 'unit' evaluated. As mentioned before, research programmes and their evaluation have become central elements in European national energy RTD. Projects are also evaluated, though in most cases this refers to ex ante evaluation of projects within the framework of project selection procedures.

To a far lesser extent, institutes, research organisations or agencies are evaluated (explicitly mentioned in E, GR, FR only). Finally, some national reports mention the evaluation of project proposals. This was not explicitly requested in the instructions report but of course project evaluation and selection are important aspects of energy RTD, since they *de facto* determine which research will be carried out.

During the discussion of the intermediary results[12] the following additional data, not appearing in the national reports, came to the fore. Ex ante evaluation of programmes is mostly based on ex post evaluation of projects, either implicitly or explicitly. Moreover, even if institutes were evaluated in a Member State it could well be that

[12] See 'Emerging Trends' document.

results of such evaluations were not necessarily made public and thus could not be taken up in the national report. Finally, evaluation of institutes might depend on their ownership, i.e. evaluations of public institutes would be made public whereas evaluations of private institutes would not.

4.4 WHICH TYPE OF EVALUATION FOR WHICH TYPE OF 'UNIT'?

The national teams were asked to describe in which situations, and for which type of [unit] the three types of evaluations are applied. The results are given in Table 9.

Table 9. Types of evaluations and units evaluated

Country	***Ex ante***	***Ex post***	***Monitoring***
Austria	Pg; pg	ins; sec; tec	
Belgium - Wallonia	pj	pg	pj
Denmark (1)			
Finland	pg; pj	pg; pj; ins; area; par	pg; pj; ins
France	Pg; pg; pj; sec	pg; sec; ins;prcd; tec; par	Pg; pg; pj
Germany	pg; pj	pg; pj	pj
Greece	pg; pj; sec	pg; pj; ins; sec	pg; pj; ins
Ireland	pj		
Netherlands	Pg; pg; pj; ins; area; tec	pg; pj; ins; area; tec	pj
Norway	pg; pj; area	pg; area; ins	pg; pj
Portugal (1)			
Spain	pg; pj; prd; ins	Pg	Pg; pg; prd; ins
Sweden[13]	Pg; area; tec	pg(?)	Pg; pg; pj; ins
United Kingdom	Pg; pg; pj; area	pg; pj; area	pg; pj; area

(1) Denmark and Portugal not known; Flanders and Italy are left out of the table.

The previous table can be converted in the following table (Table 10) which summarises which type of evaluation is most often used for which type of evaluation.

Table 10. Which type of evaluation for which type of unit

Type of eval.	***Types of Units evaluated***								
Unit:	Pg	pg	pj	ins	sec	area	prcd	tec	par
Ex ante	++	-	++	-	+	+	-	+	-
Ex post	-	++	+	++	+	+	+	+	+
Monitoring	-	++	+	+	-	-	-	-	-

'-' = hardly or not (e.g. yielding: « institutes are hardly or not evaluated ex ante »)
'+' = sometimes
'++' = in most cases (for ex ante evaluation of projects this concerns project selection)

The table does not account for 'indirect' links, that is, although entire *national* research programmes are usually not evaluated in an ex post manner as such, they often are *through* the evaluation of other units, like for instance research programmes and institutes. None of the countries perform a simultaneous assessment of all (energy) research programmes.[14]

[13] The terms have a different meaning in Sweden: see below.

[14] As per the recently performed (European level) *Five-Year Assessment*, which has of course been performed with a view to the preparation of the Fifth Framework Programme. The energy research foresight exercise in the Netherlands however made an overview of the different energy research programmes and parties and can thus in some way be viewed as such a *meta*-evaluation.

The global picture arising from the foregoing findings is that ex ante evaluation of programmes and projects is based on ex post evaluations of programmes and projects, as well as on many other inputs (modelling results for instance - cf. Topic Area 2). Ex ante evaluation of energy RTD is mostly said to be a process in which many different parties/stakeholders intervene.
Ex ante evaluation is often used to define the contents of new research programmes and/or **new research priorities**. As confirmed by the discussions of the intermediary results,[15] ex ante evaluation is clearly distinguished from foresight and technology characterisation.

Monitoring exists at programme and project levels, and, sometimes at the level of institutes. However, in most cases monitoring serves to control the evolution of the 'unit' under concern from an operational/management point of view, and hardly ever in order to scrutinise technical or socio-economic progress of the 'unit' under concern.[16] Thus, in nearly all cases monitoring serves administrative matters such as time scheduling, cost statements etc., rather than focusing on content of projects or programmes.

Finally, it might be worthwhile to mention the slightly different conception of evaluation existing in Sweden, where terms and uses are different. In this country, no 'real' ex post evaluation is carried out since, according to the national report, '*what might be learned from such a study cannot be implemented, the actual moments when decisions were made are long since gone*' (national report, § 5.3). Ex post evaluation, then, refers to what is called evaluation 'at end' in Sweden, i.e. just before the moment a unit comes to its end. Monitoring as such does not exist either in Sweden (at least not at the level of research programmes), but there are so-called 'mid-term' evaluations, normally coming in year 2.5 of a 3+3 programme. The purpose of these is to see whether *strategic* redefinition of ongoing programmes at their half-life is necessary.[17]

[15] See 'Emerging Trends' document: Olthof, de Laat, Virdis and Smith, 1996.
[16] Finland and Sweden (cf. 'interim' and 'mid-term' evaluation (see below), respectively) and France ('suivi') excepted.
[17] Of course the European Commission has also implemented Mid-Term Reviews of both Specific Programmes and funded projects, with a similar scope.

5. PROCEDURES

5.1 TYPES OF INFORMATION USED

Three types of information are generally used for evaluating energy RTD.

1. The first type of information is internal to the unit evaluated, but readily available. One should think here of documents such as project descriptions, technical reports, annual reports, scientific articles etc.
2. A second type of information is internal to the unit evaluated but is specially produced for the purpose of the evaluation itself. Here one should think of interviews with researchers, project managers, etc. and of indicators produced from outputs of the unit (aggregated figures such as totals of articles produced, patents, etc.).
3. 'external' data, which are not contained within the unit (programme, project, institute, …) itself but which serve their assessment. Here one should think of economic data, scenario data etc.

It appears that the first two types of information are mostly used for *ex post evaluation*. The third type of information appears to be most often used when *ex ante evaluation* is at stake and when new priorities are to be defined. Furthermore, in all countries the results of ex post evaluations serve as an important basis for ex ante evaluations.

5.2 ISSUES ADDRESSED

The names given to the different issues addressed in evaluations are highly diverse. Upon closer scrutiny however, *seven* major groups can be distinguished, of which the first four are most frequently mentioned as being important issues that are addressed during evaluation. The remaining three are used to a far lesser extent. These groups are given in Table 11. This table is based upon a screening of all relevant sections of the national reports where the issues addressed were at stake, including the 'datasheets'. No distinction between countries has been made since most of the issues appear in all of them (apart from the last three). The general names in the second column have been given by the Topic leader.

Table 11. The issues addressed in evaluation, in decreasing order of appearance in national reports (exhaustive list)

	Issue addressed by evaluation	***Summarises the following terms found in national reports:***
1.	Scientific and technical quality	the scientific and technical analysis of the results of the 'unit'; research quality; the contribution to the national science and technology base, to technical innovations and to other RTD
2.	Programme operation	programme management; conformity between effective work performed and contracts; timing; organisation; quality of staff; staff capability; administrative procedures; administration; programme operation; costs for the achieved benefits on project level; capabilities of project co-ordinator in managing projects; allocation of resources; goal fulfilment; achievement of initial objectives
3.	Market opportunities	market expectations; market benefits; economic benefits; economy; interest for export; business opportunities; exploitation of results; commercial exploitation
4.	Collaborations	partnerships; collaborations; links between partners; international collaboration
5.	Users	social factors; utility; use; users; utilisation of results; industrial users
6.	Energy issues	impact on specific energy issues; impact on energy efficiency; impact on energy consumption; absolute savings
7.	Other sectoral policies	conformity with and contribution to environmental policies, health policies, geographical distribution etc.

The table shows the issues which are addressed by evaluation of energy RTD in **decreasing order of appearance in the national reports**. Of course one can always argue whether the terms in the right-hand column can all be shared under the general heading of the left column (goal fulfilment for instance might concern several of the other issues). Even so, given the choices made for this table, the (by far) most frequently addressed issues within evaluation are **scientific and technical quality** of research performed under the unit, as well as **the way in which the programme has been operated** from an organisational point of view. Both are mentioned equally often. They are closely followed by both market opportunities and partnerships. It should be noted that 'market opportunities' here include both anticipated markets and the specific exploitation of results. The assessment of partnerships does not limit itself to scientific collaborations but also concerns science-industry collaborations. The remaining three issues are mentioned to a far lesser extent. Assessment of the specific involvement of users (whether these are scientific users, industrial, technical ones, end-consumers or others) is hardly mentioned explicitly (often by Finland, sometimes by France, once by Greece and Austria). It is surprising that contribution to energy policy issues is not mentioned very often. This observation might however be an artefact in the sense that it is such a natural issue in the case of energy RTD evaluation that national teams might simply have forgotten to mention it. Contribution to sectoral policies other than energy policy is hardly mentioned. Finally, contribution to European policies is never mentioned as an issue with regard to which energy RTD is evaluated.

Table 12. The issues addressed, per type of evaluation

Type of eval.	***Issue addressed***						
	1	2	3	4	5	6	7
Ex ante	x	-	x	-	-	x	x
Ex post	x	x	x	x	x	x	x
Monitoring	-	x	-	-	-	-	-

'-' = hardly or not (e.g. yielding: « in ex ante evaluation, collaborations are not/hardly addressed »)
'x' = if the issue is mentioned in a national report it is addressed for this type of evaluation

In summary, one observes that all types of issues are addressed in ex post evaluation, though the extent to which they are addressed varies from country to country.[18] Ex ante evaluation is mainly concerned with scientific and technical issues, market opportunities, and to a far lesser extent with energy issues and other sectoral policies. Monitoring, as said, is in all but three cases exclusively concerned with the operational aspects of a 'unit', not with its content-related issues.

SENSER's methodological setup, using national agencies as entry point into national energy RTD policies and management has been a deliberate and natural choice. However, this made it difficult to assess the role of authorities other than those responsible for energy RTD (e.g. environment, or transport). It has also been difficult to assess the relationship with European programmes. Generally there are two ways of deciding a national strategy: either a country can decide to mainly follow its own strategy and mobilise European research for this purpose, or look at the European research strategy and define its own in line with this. In both cases evaluation of national research would need active assessments relating to the European level. Alignment with European research programmes has, however, been mentioned only a few times as being one of the criteria according to which national research programmes are evaluated.

5.3 METHODS AND INSTRUMENTS USED

Peer review and Expert Panels are *by far* the most used methods for evaluation, experts and peers mainly coming from the field of non-nuclear energy research. 'Professional' evaluators are hardly used anywhere, though sometimes consultancies perform evaluations. Sometimes peers come from abroad. The involvement of peers from research fields other than energy is not mentioned in the national reports.

Most of the time no instruments are developed specially for performing evaluation. Where they are, they concern databases (often for the administration of projects), (postal) questionnaires, spreadsheets or datasheets.

5.4 PARTIES INVOLVED

Parties involved with evaluation are mainly public authorities who initiate and manage evaluation of energy RTD. More precisely, these are:

- those responsible for energy RTD policy itself (ministries for instance);
- managing agencies (being delegated by the first group).

Apart from the above, evaluations are also often performed by panels of independent experts or consultancies. The distribution of different types of parties performing different types of evaluation per country is given below.

[18] Only Finland explicitly mentions coherence with other sectoral policies; only Greece mentions 'geographical location' as an issue which is taken into account; only Finland and France explicitly mention 'users' whereas far more countries mention 'markets.'

Table 13. Types of parties involved, per type of evaluation

Country	***Ex ante***	***Ex post***	***Monitoring***
Austria	no information	external experts; external bureaus/consultancies	no information
Belgium - Wallonia (1)	university department	ministry	ministry
Denmark	-	consultancy; external experts	-
Finland (2)	national and international expert panels	national and international expert panel; in one case a consultancy	('interim evaluation') national and international expert panels
France	ad hoc working groups co-ordinated by managing agency or (for all energy issues) by the Plan[19]	independent experts; consultancies; professional evaluators; university groups	managing agency
Germany	expert teams; ministry, in collaboration with managing agency	ministry, in collaboration with managing agency (ex post eval. on project level)	ministry, in collaboration with managing agency
Greece	ministry	ministry; consultancy	ministry
Italy (3)	managing agency	scientific experts from national research institutes	-
Netherlands	councils/managing agencies, often aided by research groups, consultancies, ad hoc working groups, etc.	councils/managing agencies, often aided by research groups, consultancies, ad hoc working groups, etc.	managing agency/councils
Norway	specifically established working groups with experts from industry, academia, users, etc.	'egenvurdering' by programme management; 'evaluering' by external experts	by managing agency (NFR)
Portugal	government	government; international reviews (OECD/IEA)	
Spain	funding bodies; managing agency	funding bodies; managing agency	-
Sweden	managing agency with the aid of consultancies; peer review	'at end' evaluation performed by managing agency	managing agency
United Kingdom	sponsoring department or agency	independent experts, and/or independent steering and advisory committees; national audit committee (financial eval)	managing agency (e.g. ETSU)

Notes to the table:
(1) For Belgium (Wallonia) only one programme is concerned.
(2) Finland makes the distinction between 'final evaluation' and 'evaluation in retrospect', which are both ex post but the first deals with how the programme reached its objectives, how effectively resources were used, etc. (and thus evaluates what was earlier designated with 'unit operation'), whereas the second deals with actual uptake of programme results by other parties.
(3) Although, according to the national report, no evaluation is carried out in Italy, the report mentions that the national agency ENEA proposes a plan to government which is consolidated in the programme agreement and financially controlled by the Interministry Committee for Economic Planning, whereas scientific review and approval of the programme is performed by a scientific and technical committee composed of experts from the National University Council of CNR (the national research organisation) and the National Institute of Nuclear Physics. In other words, even if the report states that there is no systematic evaluation in Italy some sort of *de facto* evaluation exists, as referred to in the table.

[19] Commissariat Général au Plan.

The following remarks can be made with regard to these results.

The manifold mention of managing agencies as performing evaluation might be a bias from having representatives of national agencies involved in the SENSER project.
Evaluations are mostly carried out or supervised by public authorities or managing agencies. Ex post evaluation is sometimes conducted by consultancies. In one case, Finland, an evaluation initiated by the Ministry of Trade and Industry has been carried out by a trade organisation (the Finnish trade organisation for Ventilation and Air Conditioning Industries).

5.5 MONEY SPENT ON EVALUATION

As regards the money spent on evaluation (cf. datasheets), two main observations can be made. First, in half of the cases it appeared difficult to determine exactly which budget had been reserved for the evaluation of a research programme or other [units]. Therefore, budgets were often not mentioned on the datasheets. Second, when funding specifically dedicated to the evaluation was known, it generally never exceeded 0.7% of the budget of the [unit] evaluated and was mostly around 0.2-0.3%. Only three cases were observed where evaluation cost more than 1% of the total budget of the [unit].[20]

[20] In Norway, 1.2% for the evaluation of the programme for heat pump implementation; in France, 3% for the evaluation of JITEX (French Technology Watch Organisation in Japan), and 3% to evaluate the procedure for participation of higher education students in industrial research.

6. HOW EVALUATION RESULTS ARE TAKEN UP

In the countries which use evaluation for their energy RTD, it appears most often to be used in order to **(re-) define** and sometimes to **stop** research programmes. This means that, either explicitly or implicitly, evaluation of energy RTD is used to define research priorities.
If performed, evaluation does not have an ad hoc character. It is typically meant to be an integral part of energy RTD management on the different levels at which energy RTD is executed (national, agency, programme, project, institute, etc.). This embedding means that in all cases the results of evaluation exercises are used (see also Section 0). Therefore, the level of entrenchment of evaluation into different layers of policy making is probably a better indicator of the robustness of a country's evaluation practices, which appears fairly independent to the precise formal aspects of the evaluation methodology used.

7. NO LINK WITH TC&F, BUT STRONG LINK WITH PRIORITY SETTING

It is striking to observe that an explicit link between TC&F and energy RTD evaluation is hardly mentioned in the national reports, but that the existence of a direct link between evaluation and priority setting is mentioned without exception by all national teams. For instance, in Britain: « ... 'ex-ante' evaluations are generally concerned with developing the rationale, objectives and priority actions for a new phase of a particular programme, ... », or in Norway « *Evaluation ex ante* [...] is not a tool or method, but a process which leads to the establishment of new research programmes... ». As mentioned before, evaluation, at least evaluation ex ante, is a strategic activity and serves to define new priorities.
Where then does TC&F eventually come in? It was not possible to investigate this question in detail using the present material. However, it is the results of TC&F exercises, together with results of ex post evaluations, market assessments of the type discussed under Topic 3, and (more implicit) expert opinion into the process of ex ante evaluation which aims, *de facto*, to set new priorities by starting new research programmes.

8. SOME ORIGINAL APPROACHES

Finally, it is worth pointing out some interesting and atypical approaches found in the national reports:

- in Austria, a 'conceptually' interesting evaluation seems the one which follows the *historical progression* (1973-93) of solar water heaters in this country and the different parties involved in this evolution;
- in Finland, a specific working group has been established to give its view on the evaluation of energy research. This working group designed a detailed set of *guidelines* according to which different energy RTD programmes have to be evaluated (together with Norway and France, Finland mentions the highest number of different evaluations in the field of non-nuclear energy RTD);
- the UK has taken *'foresight' as case study*. At first sight this looked surprising ('should this not be part of Topic 2?'), but, as the author of the report later explained, the national team took the foresight example since it is probably the most extensive ex ante evaluation ever conducted in the UK in the field of energy RTD. Together with the observation made in the previous section, this calls for a reinterpretation of the relationships between foresight and evaluation. Whereas they are often seen as separate, in practice this does not seem to be the case;
- evaluation performed by ADEME in France is strongly driven by the concept of *socio-technical networks*. This concept allows the agency to represent the different parties involved with technical innovations in the energy field (research institutes, firms, users, regulation, local authorities, etc.), as well as what they produce (scientific papers, technical objects, norms, standards, etc.). The agency tries to foster innovation by improving and strengthening the relationships between different parties. Consequently, evaluation is directly geared toward an assessment of these links and their dynamics;
- finally, Greece and Spain take the evaluation of *research institutes* (mentioned further only by France as [unit]) as a case study). This means that special attention is given to the role and position of energy research institutes in the broader energy landscape, and that institutes are seen as one way, among others, to organise research (parallel to programmes, procedures, etc.).

In other words, whereas to a certain extent evaluation approaches show similarity and convergence across countries, national specifics also exist, linked to the type of research performed, the more conceptual views on how evaluation should be shaped, and, as suggested at the beginning of this report, the tradition which exists in national evaluation practice. These national approaches will not be discussed in further detail here since they are extensively described in the national reports.

9. CONCLUSIONS

Evaluation of NNE-RTD is not accountancy, but a strategic activity: in all cases it helps in defining new actions by learning from past experience. Most of the research in the countries under scrutiny is organised in research programmes. Evaluation of NNE-RTD is most often performed around such research programmes. In most countries the tendency is moving from ad hoc NNE-RTD evaluation practices toward a certain standardisation of evaluation practices, even if these differ from one country to another. Apart from the countries in which no NNE-RTD evaluation is carried out at all (Italy, Belgium), only in Denmark is formalisation of evaluation practices not explicitly sought. There is a very unequal experience with NNE-RTD evaluation across countries. The UK, France and Finland have a tradition dating back to the beginning of the 1980s, if not earlier, whereas countries such as Austria or Greece are just starting.
The political levels on behalf of which evaluation is carried out are different across countries; the national (government) level appears to be the most important one, eventually through an agency. The number of different policy levels on which evaluation is carried out depends on the maturity and tradition a country has with evaluation.

The national reports mention different reasons for carrying out an evaluation. The three main reasons are:

- justification of public expenses;
- assessment of goal fulfilment;
- setting of new priorities.

Evaluation revolves mainly around the following four types of 'unit':

1. the definition of the entire national non-nuclear energy programme: mainly ex ante;
2. the assessment of research programmes: mainly ex post;
3. projects: both ex ante (project selection) and ex post;
4. but (to a far lesser extent) institutes, industrial sectors, research areas, procedures, technologies and partnerships.

With regard to the methodologies deployed to perform evaluation, with very few exceptions, evaluation of NNE-RTD relies mainly on peer review and expert opinion. There are usually no evaluation instruments developed.

Evaluation addresses different issues. The issues addressed by non-nuclear energy RTD evaluation are, in decreasing order of appearance, in the sum of national reports:

1. scientific and technical quality
2. programme operation
3. market opportunities
4. collaborations
5. users
6. energy issues
7. relation with other sectoral policies

The first four issues are addressed by all countries. The last three appear to be addressed only to a marginal extent. 'Relationships with European policies or research programmes' is not explicitly mentioned as being an issue addressed by NNE-RTD evaluation at national levels.
As concerns the relations between the different issues and the three different types of evaluation distinguished, the following can be observed. Ex post evaluation normally addresses the first four issues and the last three – but to a lesser extent. Ex ante evaluation addresses issues 1 and 3, and sometimes 6 & 7. With three exceptions, monitoring focuses on programme operation only (issue 2).

With regard to the types of parties carrying out evaluation, these are:

- ex ante evaluation: by public authorities and government, managing agencies, sometimes by independent expert panels but rarely by consultancies only; often it is performed through specially established ad hoc working groups involving several of these parties;
- ex post evaluation: most often by expert panels, sometimes by consultancies or managing agency;
- monitoring: mostly performed by managing agency or public authority involved.

Although one can observe similarities in the way the different countries evaluate their energy RTD, differences and some original approaches also exist. Austria for instance has evaluated the 'technological progression' of

solar heaters in that country. Finland established (some 10 years ago) a special working group to establish the way in which energy RTD should be evaluated, resulting into an extensive set of guidelines. France uses the approach of 'socio-technical networks.' This to identify and monitor the relationships between parties and intermediaries built up through the research (and other) programmes of the agency. More recently, the approach has been used to perform assessments with a more strategic character. Greece and Spain mention the evaluation of non-nuclear energy research institutes as an important evaluation activity. Finally, the UK cites the huge national foresight exercise in the field of energy as one of the biggest ex ante evaluations ever undertaken in this field.

Finally, an important conclusion from the national reports on evaluation of non-nuclear energy RTD is that no explicit links between Technology Characterisation and Foresight (cf. Topic 2) and evaluation exist. TC&F results might however be mobilised, like the results of ex post evaluation, within *ex ante* evaluation.

10. INTERPRETATION & RECOMMENDATIONS

Topic 1 aimed to investigate whether, why and how evaluation of non-nuclear energy RTD is performed in European Member States, and subsequently to derive indicators and evaluation methodology on a European level.

This report has reviewed Member States' practices in evaluation of non-nuclear energy RTD. Also, the European Commission's own experience has been briefly reviewed (annex 3). There, a certain methodological 'decline' was observed going from the NNE3 evaluation to the recent Five-Year Assessment. Still it must be concluded that the Commission's evaluation practices in energy RTD have often been more sophisticated than that of individual countries. Therefore, producing a list of 'indicators for evaluation' for the European Commission seems redundant. In fact, the results discussed in this report demonstrate that the question is not merely to find the 'right' indicators once and for all. Good indicators are local and temporary, depend strongly on the piece of research at stake and the precise objectives stated, and that both can be subject to change. Instead, the broader challenge is to set in motion the process which allows evaluation to become more widespread, and connected to other elements in the strategic management of non-nuclear energy RTD. It is through the establishment of such learning processes that 'right' indicators will be found (and modified when new important issues come to the fore!) - not the other way around.

Therefore, instead of providing indicators *stricto sensu*, this last section will present the 'challenges' evaluation of energy RTD faces if one takes such a view. First, unproblematic issues in evaluation in Member States and European level programmes will be discussed. These are the assessment of scientific and technical quality of the research, and the evaluation of programme operation. Next, the problematic issues in evaluation, both for Member States and – given the recent assessment of the non-nuclear energy programmes – also for the European Commission, will be described.

10.1 WHAT IS UNPROBLEMATIC IN ENERGY RTD EVALUATION...

Given the practices described in national reports and the current state of the art in European evaluations of non-nuclear energy RTD (see Annex 3), the authors feel that two issues are no longer problematic in the evaluation of energy research: (1) the assessment of scientific and technical quality of research, and (2) the evaluation of how programmes are operated.

10.1.1 Scientific and technical quality of the research

The evaluation of scientific and technical quality of non-nuclear energy RTD is felt to be the most unproblematic part within the evaluation practices on both national and European levels. Procedures for selecting research projects are established in all countries (peer review project selection). Ex ante evaluation serves, as a process involving different (types of) parties, to identify the promising broader areas for the future. Scientific and technical ex post evaluation of 'units' does not seem problematic either, at both national and European levels. In fact, it is the most important issue evaluations of Member States appear to deal with. One can thus expect that cumulative experience gained for this issue is satisfactory.

10.1.2 Programme operation

Evaluation of 'programme evaluation' is the second most well-developed issue in evaluation of energy RTD as performed by Member States. This covers the assessment of whether the selection procedures were followed, where and how money was spent, whether deadlines were kept, whether cost statements were filled in according to the rules etc.[21] In summary, this is *administration*.
The ability of evaluators to perform such assessments well mainly depends on the bookkeeping quality of the management of the [unit] under scrutiny. The more data gathering (on project progress, expenses, results, dates etc.) is formatted beforehand, the more this type of evaluation can become routine. Such is the case of, for instance, ADEME's EVAL database but it also becomes increasingly common for European RTD programmes. Although the evaluation of programme operation is an important aspect of evaluation, and has to remain so if

[21] A rather exhaustive list of 'Programme operation' items, appearing to cover the different items mentioned in national reports, can be found in the document established by CREST in view of the 1996 monitoring exercise of the 3rd and 4th Framework Programmes of the European Union.

only to see whether programmes are run without squander, from a conceptual point of view, this type of evaluation does not represent any difficulty.

10.1.3 Recommendations

Evaluation can always be improved. Nevertheless, the Topic Leader feels that optimising evaluation methods and practices for scientific and technical assessment of research [units] has become a fairly unproblematic matter. Scientific and technical quality of research in both Member States and at the European level are mostly performed by expert panels. These are often supported by information on the [unit] evaluated. Several methodologies for gathering such information exist. Similarly, programme operation ranks second in the list of issues which are addressed by energy RTD evaluations. Neither at Member States level, nor at European level, do barriers seem to exist to carry such evaluation out as desired – it simply has to be done.

In summary, evaluation of scientific and technical quality of research, and evaluation of [unit] operation seems to be performed in a satisfactory manner. This current state of affairs should be maintained at both national and European levels. Member States – in particular those where evaluation does not (or hardly) exist – might consider performing it more often and more systematically. In all cases, evaluation could eventually benefit from more systematic use of methods developed specially to assess scientific and technical quality (documentation, and other). Overall however, serious improvements of evaluation *in general* will in our view not come from further improvements of methods and practices for evaluation of scientific and technical quality of research, nor from further improving methods and practices for the evaluation of [unit] operation. Therefore, it is recommended that the priority for further development of evaluation methods not be placed on these two issues.

10.2 WHERE THE FOCUS SHOULD BE

Whereas scientific quality and programme operation are thus not priorities in further methodology development for evaluation purposes, there are several issues in energy RTD evaluation which deserve more attention. These are:

- socio-economic benefits of energy RTD;
- monitoring of energy RTD;
- circulation of ‘good practices’ between Member States;
- the taking into account of European policies in Member States' evaluation;
- how long-term vs. short-term issues are accounted for.

The present section discusses these six ‘problematic’ issues in turn. For each item, a ‘challenge for a new evaluation approach’ will be derived, from which both European and Member States’ research should be able to benefit. This will allow us in the next section to make a proposal for an evaluation framework giving special attention to co-ordination between the European and Member States level, and taking into account the results from SENSER’s other topics.

10.2.1 Socio-economic benefits of energy RTD: the blind spot of evaluation

The assessment of the *socio-economic benefits of research*, in both ex ante and ex post evaluation (let alone in monitoring: see below), appears still highly problematic. The authors of the present report suggest that the priority for further development of evaluation methods should be on this issue.

Funding of energy research on a European level emerged as a direct response to the oil crisis. Therefore, from the outset, energy research in Europe was to lead to specific results. It thus had to lead to at least one clearly defined socio-economic benefit, that is, to provide reliable energy to European citizens. Whether European research indeed managed to do so, is less clear. This is well illustrated by the following citation, taken from the recent report of the Five-Year Assessment Panel (p.18):[22]

[22] The Five-Year Assessment Panel evaluated the non-nuclear energy research programmes under the Third and Fourth Framework programmes in 1996. For an overview see Annex 2, on European evaluation practices.

> *It is impossible to gauge how much of this success [in energy conservation in Europe since the first oil crisis] can be directly attributed to EU and Member States' RTD programmes. Clearly, since these programmes supported the development of new technologies, it is reasonable to claim that some of these benefits are due to EU and Member States' programmes.*

However, as the citation well expresses, this is only a *claim*. What would have happened if no European research had been carried out? The recent report of the Five-Year Assessment Panel states that the JOULE programme has remained too 'fundamental' and that THERMIE results have not been replicated to a satisfactory extent.[23] However, nobody would argue that European non-nuclear energy RTD as a whole has remained without any effect. Given the current state of evaluation practice, however, nobody knows precisely how to assess or evaluate these effects, let alone what these effects – on future markets, on society – precisely are. And, following the country reports, very few of the Member States seem to have a clear-cut opinion on how to deal with this topic. Most of them look for market opportunities in their evaluations – these rank third in the list of issues addressed by evaluation. However, compared to the earlier issues, they are looked upon to a lesser extent. Moreover, broader societal issues are hardly addressed explicitly through evaluation according to the national reports – and if they are, mostly through ex ante evaluation. Also, the related issues of 'collaborations' and 'users,' both directly related to the adoption of research results, are seldom addressed in evaluation at both national and European levels.

The challenge. Due to its specific nature, research results can generally not directly be characterised by their monetary value. This is especially true for research where markets, and thus monetary benefits, simply do not (yet) exist. Although several attempts have been made to evaluate research with regard to its socio-economic value (methods for internalisation of environmental costs are an example), such evaluations are based upon expectations rather than on hard facts. Inversely, scientific and technical assessments can be based on observations, are consequently easier to perform, and, unsurprisingly, currently constitute the main focus of non-nuclear energy RTD evaluation in Member States.
For much of non-nuclear energy research in Europe markets barely or do not exist. If one compares it to research for which markets do exist it is easy to understand that the latter will theoretically always 'win' – and no new research would be funded. Therefore, it is also not surprising that none of the country reports describe a formal method to directly derive financial benefits from funded research.[24]

With regard to the country reports, it is thought that *general* socio-economic indicators applying to every research [unit] across Europe, do not exist. Precise socio-economic indicators appear to be specific to different pieces of research and in the first place depend on results and benefits envisaged by the promoters of the research themselves. However, the *expected* socio-economic value of research can be addressed more generally. But only *indirectly*. The country reports (in particular from Finland and France) suggest three main routes to address the (anticipated) socio-economic benefit of research actions from this more general point of view:

- **by analysis of the type of collaborations established through research projects**.

Collaborations in research give a very strong indication of the interest and trust different types of partners put in the project (industrialists, research organisations, etc.). An analysis of collaborative patterns ('vertical' or 'horizontal')[25] will indicate how 'close to the market' the parties involved consider research to be;

- **by analysis of the specific and/or anticipated involvement of *users* in research projects**.

The involvement of users in research is a good indicator for the evolution of a project, for at least two reasons. First, it testifies to the interest put in, and thus the experienced usefulness of, a given technology. Second, it simultaneously gives an indication of whether some potential uses might have been ruled out by the past technical choices made in the project;

- **multiplier ('snowball') effects of research actions, and the adoption of research results by parties other than direct participants to the research**.

As stated above, the Five-Year Assessment Panel put big question marks against the actual diffusion of research results from the European non-nuclear energy programmes. It is clear that the adoption of research results by parties other than direct participants to the research is only a potential indicator of success. In fact, several

[23] The evaluation of 100 JOULE projects, achieved in March 1997 and not yet published is somewhat more positive about this issue.
[24] In the discussions around the country reports (cf. meeting of October 15, 1996) some representatives were even against it.
[25] Where in 'horizontal collaboration' industrialists of the same type work together (also termed 'pre-competitive,' or 'basic technological' research), and in 'vertical collaboration' they do not (but work only with research groups, supplier firms, etc.).

methods for following through the diffusion of research results exist or can easily be developed by looking at the different types of outputs a research project delivers.

Of course, evaluating these issues becomes interesting only if they are looked at from a dynamic point of view (cf. the Austrian case study). Thus, the evolution over time of collaborations, user involvement, the adoption of research results, and, ultimately, the establishment of markets, should be looked at over broader time periods rather than being considered statically.

10.2.2 Monitoring the progression of a technology or research project

The issue of adopting research results brings us to the current role of monitoring non-nuclear energy research in Member States. Before (and after) being adopted by an eventual user, a technology, a technological research project or a programme follows a *particular route*. This feature is present in nearly all national reports through the idea of 'monitoring.' However, with some exceptions, monitoring is in most cases equivalent to regularly checking administrative matters. This is not the case in the Finnish and Swedish idea of interim evaluation which aims to evaluate research results and their users. In the French network approach, a technology's 'progress' is assessed in terms of new parties (researchers, firms, public authorities, etc.) who become interested in, and link themselves to it. Finally, the idea of progression is explicitly dealt with in one of the two Austrian case studies, regarding the diffusion of solar water. In other words, monitoring does not have to limit itself to administrative matters only. It can also assess socio-technical and socio-economic aspects of a technology or technical research project.

The challenge for developing a new evaluation approach is to design a type of monitoring which looks beyond programme operation only, as is presently done in most of the cases. It might differ from ex post evaluation in scale (that is, it could be a smaller exercise), but should preferably be similar in scope (that is, address the same issues).

10.2.3 Can the countries learn something from each other?

The initial aim of the SENSER project was to make recommendations and indicators for the *European* level. However, it is felt that the participating countries may also benefit from the analysis of each other's evaluation practices.
It must be concluded that the ways in which evaluation practices have been shaped over the past few years in the different countries which participated in the project greatly depend on the local political circumstances. Therefore, it is not the purpose of the present report (even apart from the question of subsidiarity) to prescribe certain evaluation methods and tools. Nevertheless, countries with a more long-standing tradition in evaluation share some similarities which might be interesting for the other countries involved. In particular:

- the multiplication of levels on which evaluation is carried out as evaluation practices become more long-standing;
- developing guidelines (cf. the Finnish case) for national evaluations.

Apart from this, it is recommended that 'good' practices be circulated, to allow comparisons of the different evaluations relating to similar themes, research programmes or energy technologies, and to exchange experts in each others' evaluation exercises – a practice observed in some Scandinavian countries.

The challenge for developing a new evaluation approach is to make use of 'good practices' in Member States, as analysed through the national reports. In some cases, given its long-standing reputation in evaluation of energy research, these 'good practices' might well be the evaluation practices of European programmes!

10.2.4 Taking European policies into account

Since the Maastricht Treaty (art.130H) the European Commission is now explicitly obliged to co-ordinate the research it funds with that of the Member States, although no such clear obligation exists in the other direction (although co-ordination always comes from more than one side). Such an obligation would not be sensible insofar as the European programmes are established and approved by the Member States, and thus a *de facto* co-ordination is already expected to exist. This notwithstanding, it was surprising to see that in the evaluation of national energy RTD programmes, co-ordination with European policies – R&D, energy or other – was mentioned by only one country as an issue which was systematically investigated through E-RTD evaluation. Even if countries have their sovereignty in choosing their national RTD, it is strongly advised that Member States consider co-ordination with research carried out at the European level as an issue in national evaluations.

The challenge for developing a new evaluation approach is first that it has to take note of what happens in the Member States. Symmetrically, however, given the close complementarities that exist, it is suggested that, in their evaluation, Member States consider more seriously the links between their own national energy research programmes and European energy research and other policies. More generally then, a new evaluation approach should be able to consider European and Member States' energy RTD together.

10.2.5 Long-term/short-term research and evaluation

The two previous issues become all the more significant if one takes into account one of the main results of SENSER's Topic Areas 3 and 4.[26] Both analyses point to a decrease in 'long-term' research at national levels, and simultaneously at an (implicit) entrustment of this type of research to the EU level. In other words, Member States support the short and medium term (for reasons of globalisation, privatisation and liberalisation).
In the national reports no explicit distinction was (asked to be) made between different types of research (long, medium, short term) and their relationships with different types of evaluation. The analysis of the use of different indicators as mentioned in the national reports allows, however, for the following observation. On the one hand, if a programme has a more 'scientific' character, evaluation appears to concentrate more on the scientific and technical aspects of the research. Such assessments, as outlined above, could always be improved by making more and better use of methodologies which over the last decades have been established for this purpose. Overall, such assessments seem relatively well developed and established.
On the other hand, if a programme is clearly market driven and should lead to usable products and technologies, the problem appears not to be so difficult either, since assessments then simply analyse whether the envisaged products have been delivered and whether they are adopted, or sold. This is done by 'market assessments' of research results – which can be very successful assessments provided that the programme of concern is indeed close to the market.
However, non-nuclear energy research as performed at both European and national levels over the last decade(s) has never been entirely 'scientific' – making use of technical-scientific assessments only is insufficient. Nor are they based upon markets which are entirely established – the case in which evaluation methods would be available, but in which *public* research programmes would not necessarily be needed anymore. Energy RTD at both European and national levels is technological research *par excellence* and thus involves both scientific and socio-economic or market aspects. This however necessitates new types of evaluation which might borrow from existing scientific (long-term) evaluation and from market (short-term) evaluations. It should itself be new. Neither in the Member States nor at the European level has this type of evaluation been invented yet, though in dispersed fashion various elements do exist.
The challenge for developing a new evaluation approach is that it has to be able to take into account, again without problem of continuity, the short, medium and long term.

[26] Topic Area 3 concerns market drivers for RTD. Topic Area 4 concerns Areas for EU intervention, and specifically yields a list of national RTD priorities and expenses.

ANNEX 1

D A T A S H E E T F O R E N E R G Y R T D E V A L U A T I O N S
(u s e k e y w o r d s a n d s h o r t d e s c r i p t i o n s)

n° … of …

Name of the evaluation: ..... The evaluation applies to a ..... (fill in [unit]) General domain of application of the [unit]: ..... Type of evaluation: .....(ex-post / ex ante / monitoring) Duration of the evaluation: from … / … / 19… to … / … / 19… At which moment/period in time is it located in the course of the development of this [unit]: /... /... /19...(start date of [unit]) (end date of [unit]) .../.../19...\ mark moment: Budget of the evaluation ..... (approx kEcu 1995) Total Budget of the [unit] ..... (approx kEcu 1995) (ratio = …)
What were the reasons for *initiating* the evaluation?
Who initiated the evaluation / who asked for it to be carried out?
What were the initial *objectives* of the evaluation?
Who carried out the evaluation (expert panel, consultancy bureau, the management of the [unit] itself, ...)?
Which *issues* are/were addressed by the evaluation to achieve its objectives (e.g. economic feasibility, technical assessment, acceptability - feel free to add items)?
Which *types of information* are/were used in the course of the evaluation (e.g. potential market shares of products, project descriptions, etc.)?
How is/was this *information processed* (e.g. through peer review, documentation analysis, etc.)?
Was the evaluation specially *structured*, and if so, how? (e.g. a specific database set up, …)
Which are the *main results* yielded by the evaluation (e.g. specific data, specific knowledge on the evolution of the research, etc.)?
Have these results been *visibly used* or otherwise been taken into account? If so, by whom and how (e.g. definition of future action)?
How/by whom was the evaluation process *managed* (e.g. specialised bureau, national authority)?
Further comments (e.g. hidden objectives of the evaluation, new objectives which became clear during the course of the evaluation, other ...)

ANNEX 2. Three original approaches to the evaluation of energy RTD (excerpts from national reports)

United Kingdom: the ROAME approach

In the UK the initiation, implementation and final assessment of government-sponsored R&D programmes is based on the 'ROAME' principle. i.e.

Rationale	An explicit statement of why government funding should be used for a programme
Objective	This is a statement of what is planned to be achieved and by when. An objective should always be expressed in such a way that it will be possible later to tell whether it has been achieved
Appraisal	The process of identifying required benefits, defining rationale, setting objectives, examining options and weighing up the costs and benefits before the programme starts
Monitoring	This concerns regular checking of progress against plan.
Evaluation	This is the process of checking, at the completion of a programme or project, to what extent and how the objectives were met

As a consequence of this approach the criteria for both ex ante (Appraisal) and ex post (Evaluation) evaluations of R&D areas and programmes are set at the Rationale and Objectives phases. It also means that the evaluation criteria are fairly generic, although the approach to evaluation against these criteria may vary depending on the nature of the area or programme. Important items, which are apparent in the examples given in the annex to this chapter are:

- value for money (commonly measured as a cost to benefit ratio);
- extent to which declared objectives will/have been met;
- efficiency and effectiveness of programme management;
- continued relevance of objectives;
- effectiveness of technology transfer.

In contrast, the evaluation of individual projects tends to focus on performance criteria such as meeting objectives, delivery to time and cost, quality of work, etc.

Finland : guidelines for the evaluation of sectoral research

In its statement from 1987 the State Council for Research and Technology noted that sectoral research activities in Finland in particular should draw their attention to the development of evaluation activities. Following this, during autumn 1988, the Ministry of Trade and Industry published guidelines for the evaluation of sectoral research within the Ministry's responsibility area. This includes both non-nuclear and nuclear energy RTD.

Energy RTD in Finland was organised into so-called national energy research programmes in 1988, based on recommendations by the Committee on Energy Research. The first generation of energy RTD programmes was planned for the period between 1988-92 and the programmes were supposed to provide the first programme level interim reports during 1990. This was also to provide a suitable setting for the first evaluation of the programmes. For this purpose the Ministry appointed a Working Group for the Evaluation of Energy Research Programmes (WGEERP) in September 1989. The WGEERP was to draw up a proposal on methods suitable for evaluation of energy RTD and to produce a plan for implementing evaluations of the national energy research programmes in a way that would serve both the management of the programmes themselves, the assessment of their international standard and the planning of a new generation of programmes.
The WGEERP delivered its report in March 1990. It defined the objectives and criteria of energy RTD evaluation on three distinctive levels:

Level of evaluation	Objectives	Criteria
General impact of energy RTD programmes and the activity of the funder	• assessment of the impact of energy RTD programmes from the viewpoint of general objectives of energy policy and the energy economy • assessment of the significance of energy RTD programmes as a tool for research and technology policies	• benefits of the programme to the national energy economy • distribution of the benefits to different sectors of the economy • compatibility with the objectives of science and technology policies in general • impact of RTD programmes on energy research and research institutes • the appropriateness of the programmes as a tool for administering RTD funding
Scientific level and significance of energy RTD programmes	• evaluation of the scientific and technical level of energy RTD programmes in comparison to international standards • assessment of the social and energy impacts of the programmes on their own application field, and the impact of individual programmes on technological development and innovation	• international scientific level of research, increase in technological know-how in the programme area, training of researchers • the appropriateness of programme objectives • the expected benefits and innovations by potential user group • private sector interest in financing the activities
Internal implementation of the programmes	• evaluation of the progress of the programmes according to their objectives • productivity, results and quality of the programme	• materialisation of the objectives • scientific level and methods • focus • productivity • internal organisation, efficiency of resource use • publications and information dissemination

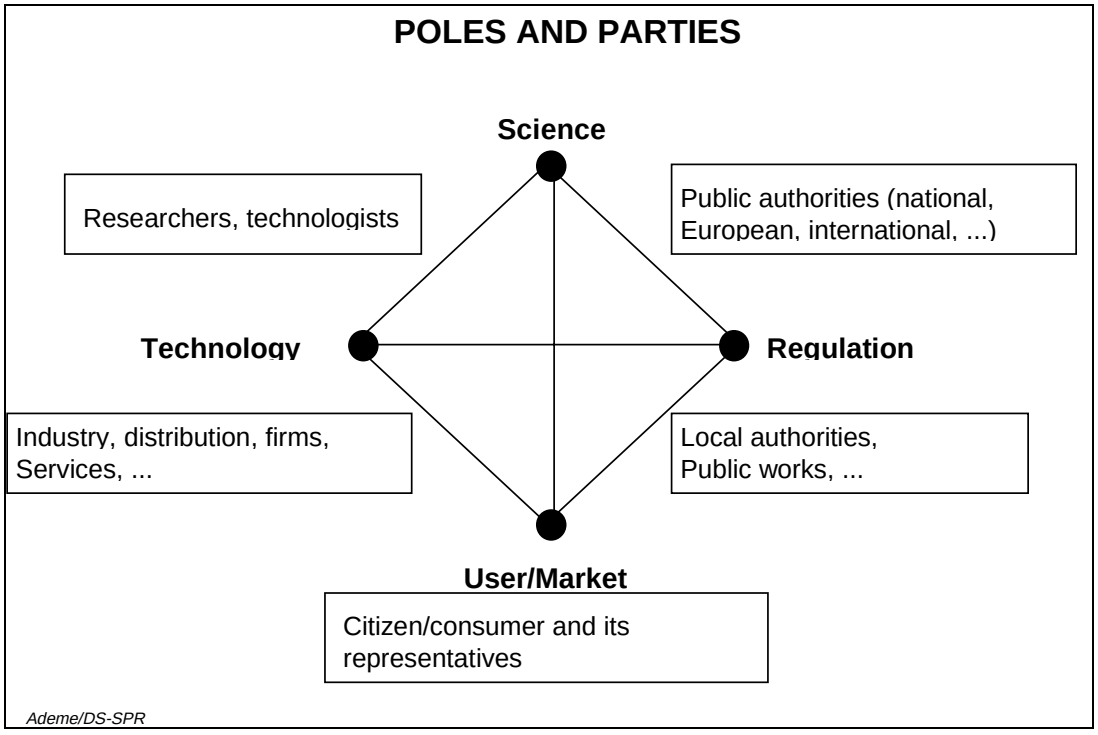
In its report the WGEERP made a distinction between “evaluation ex ante”, “evaluation during implementation”, “final evaluation” and “evaluation in retrospect”. From the terminology and semantics point of view these differ somewhat from the definition of “evaluation ex ante”, “monitoring” and “evaluation ex post” used later in this report. The WGEERP’s definition of “evaluation ex ante” corresponds to the one used in the SENSER project. “Final evaluation” and “evaluation in retrospect” are both forms of “evaluation ex post” with the distinction that “final evaluation” is concerned with questions such as “did the programme achieve its initial objectives?” and “how effectively were programme resources used?”, whereas “evaluation in retrospect” deals with evaluating the actual industrial or other uptake of research results after an appropriate period of time has elapsed after the termination of the programme.

The WGEERP put strong emphasis on the “evaluation during implementation” or “interim evaluation” of the energy RTD programmes. As a concept this is actually a combination of evaluation ex ante, monitoring and evaluation ex post. It corresponds to evaluation ex ante in the sense that one of the main objectives of programme evaluation is to provide guidelines, recommendations and plans for the implementation of the remainder of the unit evaluated or of a new programme/set of programmes. Interim evaluation is also monitoring in the sense that interim evaluation by definition concerns the checking of an ongoing RTD programme or programmes. And finally, interim evaluations are also a form of evaluation ex post, since the majority of information upon which the evaluation is based comes from individual completed projects.

France: the socio-technical network for the evaluation of non-nuclear energy research

ADEME's 'network'-approach, extensively discussed in the report of the Panel project,[27] appears a robust methodology to evaluate energy-, and other types of RTD. The methodology (or rather the concept) allows not only for evaluating the technical aspects of RTD, but, because of its focus on the links between the heterogeneous parties in the innovation process allows, in an integrative manner, also for a full-fledged analysis of the socio-technical character of RTD and innovation processes in general. It provides good heuristics for performing an evaluation, with regard to the type of parties (research, 'societal' etc.) one should investigate as well as with regard to the data one has to analyse (e.g. in one case it will be more appropriate to analyse patents applied for in a certain research field, whereas in another it might be better to analyse the involvement of SMEs in publicly funded research). The approach is constantly improving. Recently it has been broadened to take into account not only energy but also the environment field. Furthermore, research has been undertaken to incorporate more ex ante aspects in the model, which until now was mainly focused on ex post evaluation.

A schematic representation of a socio-technical network is given below.



The idea behind the network concept is that an innovation becomes more successful when the network of parties becomes:

- *longer* (extending from one (not necessarily science !) pole to another;
- more *complete* (that is, they involve more (different) 'poles', without 'holes' between them, e.g. presence of science and market poles, but absence of technology pole);
- and more *convergent* (parties sharing objectives).

The role of the managing agency (here ADEME) is to ensure that a network is established around a certain technology. The approach gives clear clues for intervention, e.g. when a pole is only weakly developed, more effort should given to this; when all parties exist but do not converge, the effort should be on assembling them, trying to define common objectives, or organise exchanges; finally, when the network is considered to be too 'short' (with respect to science or the market), extension should be sought.

[27] Cf. Marsh et al, 1996.