CORE organic

Project no. **ERAC-CT-2004-011716**

Project acronym: **CORE Organic**

Project title:

Coordination of European Transnational Research in Organic Food and Farming

Instrument: Co-ordination of Research Activities
Thematic Priority: Coordination of National Activities (ERA-net)

D 5.2: Scientific evaluation of trans-national projects – Between credibility and national preferences

Prepared by

Thomas Alföldi and Urs Niggli, FiBL
Stéphane Bellon and Julien Blanc, INRA

Due date of deliverable: **30.9.2007**
Actual submission date: **30.9.2007**

Start date of project: **01.10.2004** Duration: **36 Months**

Organisation name of lead contractor for this deliverable:
Research Institute for organic agriculture FiBL, CH-Frick

Revision [draft]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

DISCLAIMER

This publication has been carried out with financial support from the Commission of the European Communities within the ERA-net scheme under the 6th Framework Programme for Research and Technological Development. The support is greatly acknowledged by the authors.

Content

1 Executive summary

The scientific evaluation of organic food and farming (OFF) projects is challenging under two aspects. Firstly, it has been observed in the past that traditional evaluation criteria may not consider appropriately the system-oriented, innovative and interdisciplinary approach of OFF-proposals. Secondly, the success of trans-national research programmes is highly dependent on a transparent evaluation procedure which clearly distinguishes between scientific evaluation and political selection.

Our objective here is to assess to what extent the evaluation criteria and procedure used for the CORE Organic pilot call challenge these perspectives and to provide further recommendations to improve the future call procedures.

The assessment consists of three parts. Firstly, the appropriateness of the chosen evaluation criteria was examined by a questionnaire sent to the different target groups involved in the CORE Organic pilot call (members of the expert panels, national call contact persons and governing board members). Secondly, the results of the pilot call were used to identify the most important evaluation criteria responsible for the rejection or acceptance of proposals. These analyses were supplemented by a literature review of the latest assessment practices in the field of research evaluation.

The survey showed that the proposed set of nineteen evaluation criteria, clustered within six main categories, fulfilled the expectations of most target groups involved with the CORE Organic pilot call. However, some of the respondents wish to have a stronger focus on aspects of interdisciplinarity. Some of the comments of the applicants showed that handling the gap between the scientific evaluation and the final selection of OFF research projects is the main challenge for a successful evaluation of trans-national research in the future. To overcome this conflict, different measures are suggested.

The analysis of the scientific evaluation of the proposals submitted under the pilot call showed that rejected proposals had received insufficient scores regarding methodological criteria. Proposals which were given a high priority showed higher scores for the criterion *relevance*. In order to improve chances to succeed in future calls, proposals should therefore focus on the criteria *methodology* and *relevance*.

The literature review revealed further potential for optimising the evaluation criteria and the call procedure. Most all publications on the topic of peer reviewing or research evaluation address new challenges from a theoretical perspective. Our analysis has shown, however, that the criteria and procedures used in the CORE Organic pilot call only partially address the new challenges which have evolved due to the specificities of OFF. Interdisciplinary and innovative aspects should be addressed in a more appropriate way.

We therefore suggest that the decision-making process should be open to a wider community of experts in order to assess cooperation and trans-national aspects. Furthermore the process should be open to non-governmental representatives in order to provide wider scrutiny. The way national priorities are integrated in the decision-making process should be thought over. The transparency of the procedure should be improved.

Mechanisms should be implemented that make it possible to fund a few “risky” research projects and to facilitate new incomers to enter the arena.

Further assessment steps could be implemented for those projects that were identified as particularly innovative but methodologically less robust. This may include the tutorial of the ongoing research projects and the evaluation of the results.

Furthermore the evaluation criteria should be refined. Particularly, the creation of a main category *interdisciplinarity* is of crucial importance.

2 Aims

The aim of deliverable 5.2 is to provide model Terms Of Reference (TOR) for the scientific evaluation of ERA-NET proposals and to provide checklists ready to use for trans-national research programmes on organic farming. This will also contribute to improve the evaluation procedure at national levels. These aims were achieved through a critical analysis of the evaluation criteria and procedures, which were used for the CORE Organic pilot call. The analysis consists of the following parts

- Part 1 “feedback”: The appropriateness of the chosen evaluation criteria was examined by critically discussing the experiences of the pilot call with the target groups involved.
- Part 2 “analysis”: The results of the pilot call were used to identify the most important evaluation criteria responsible for the rejection or acceptance of proposals.
- Part 3 “literature”: The analysis of the evaluation criteria and process was extended by including common evaluation practice based on a literature review.
- Part 4 “recommendations”: The recommendations for possible improvements of the evaluation procedures and criteria contain the lessons learnt from the pilot call (synthesis of part 1-3) as well as additional suggestions made by the different target groups.

3 Methods

Tab. 1 provides an overview of the methods and related results for the different parts of Deliverable 5.2.

For part 1 “feedback”, a questionnaire was developed together with Work package 7 (improving call procedures). This questionnaire was sent out to the different target groups involved in the pilot call. An overview of the evaluation criteria used in the pilot call is shown in Table 3[1].

Part 2 is achieved by analysing the results of the expert evaluation panel provided by the call coordinator.

In part 3, a literature study provided additional information on existing evaluation criteria and procedures. Recommendations for future calls are based on the results of parts 1-3.

Tab. 2 shows the call phases and target groups relevant to the scientific evaluation criteria and procedures in order to distinguish between aims of Deliverable 5.2 and Deliverable 7.2.

[1] A detailed description of the evaluation criteria and procedure can be found under http://www.coreorganic.org/research/CORE_final_guidelines_060918.pdf (pages 7-11)

Tab. 1: Methods and aims

	Part 1: Feedback	**Part 2: Analysis**	**Part 3: Literature**	**Part 4: Recommendations**
Method	Questionnaire	Statistical analysis of the results of pilot call	Literature study	Synthesis of results part 1-3
Aims	Target groups quantify the importance, the feasibility of the existing evaluation criteria and propose additional criteria Clarify the question of inter- and transdisciplinarity	Performance of different criteria depending on topics, priority Overview on how submitted proposals fulfilled evaluation criteria with recommendation for future calls	Additional aspects regarding evaluation procedure and evaluation criteria Discuss and include additional aspects to the lessons learnt in the pilot call	Recommendations to improve evaluation criteria

Tab. 2: Relevant call phases and target groups for Deliverable 5.2 addressed in the questionnaire together with WP 7

Target groups **Phases**	**Applicants**	**Expert Panel Members**	**Governing Board (GB)-Members**	**National Call Contact Person (NCCP)**
Preparation			X	
Application	X			X
Evaluation	X	X	(X)	X
Selection	(X)		X	
Follow-up				

4 Results and discussion

4.1 Survey on the evaluation criteria of the pilot call

4.1.1 Importance of the different evaluation criteria

Most of the evaluation criteria used for the CORE organic pilot call were judged to be very important or important (Fig. 1). The national call contact persons (NCCP) and Governing Board members generally judged the importance of the criteria higher than the expert panel members and the applicants. Based on a synthesis of the results shown in Tab. 3 and Fig. 2 the importance of the different criteria can be described as follows:

- *Innovative research* and *scientific quality* are seen as the most important criteria, especially for NCCP and members of the Governing Board. However, it is also interesting to see that 3 respondents of the applicant group are not sure whether innovation and scientific quality are important criteria or not. All target groups ranked *scientific quality* above *innovative research*.
- Within the main category methodology, the criterion *choice of methods* is seen to be as important as the *plan for knowledge transfer*. The *plan for publication* is considered to be less important.
- The main category consortium consists of six criteria. *Qualification of consortium* and *true cooperation* are for all respondents incontestably important criteria. Applicants and expert panel members consider *transdisciplinarity of consortium* and *scientific networks* clearly less important than the other two groups.
- All criteria regarding project management are generally considered to be important and no relevant differences between the target groups were detected.
- All criteria regarding relevance are important except the criterion *societal relevance,* which 20 % of the applicants and 30 % of the experts consider of minor importance.
- The criterion *added value for EC research* is important for all target groups. However, for *trans-national aspect,* there is a strong discrepancy between applicants and experts on one side and NCCP and members of the Governing Board on the other side. More than 30 % of the applicants and 44 % of the experts say that this criterion is of minor importance for the selection of OFF-research projects.

These results can be interpreted in different ways: (i) A certain scepticism towards the scientific evaluation procedure itself, (ii) some criteria are indeed not so relevant for the selection of OFF-research projects or (iii) the importance of some of the criteria has not been understood and needs better communication.

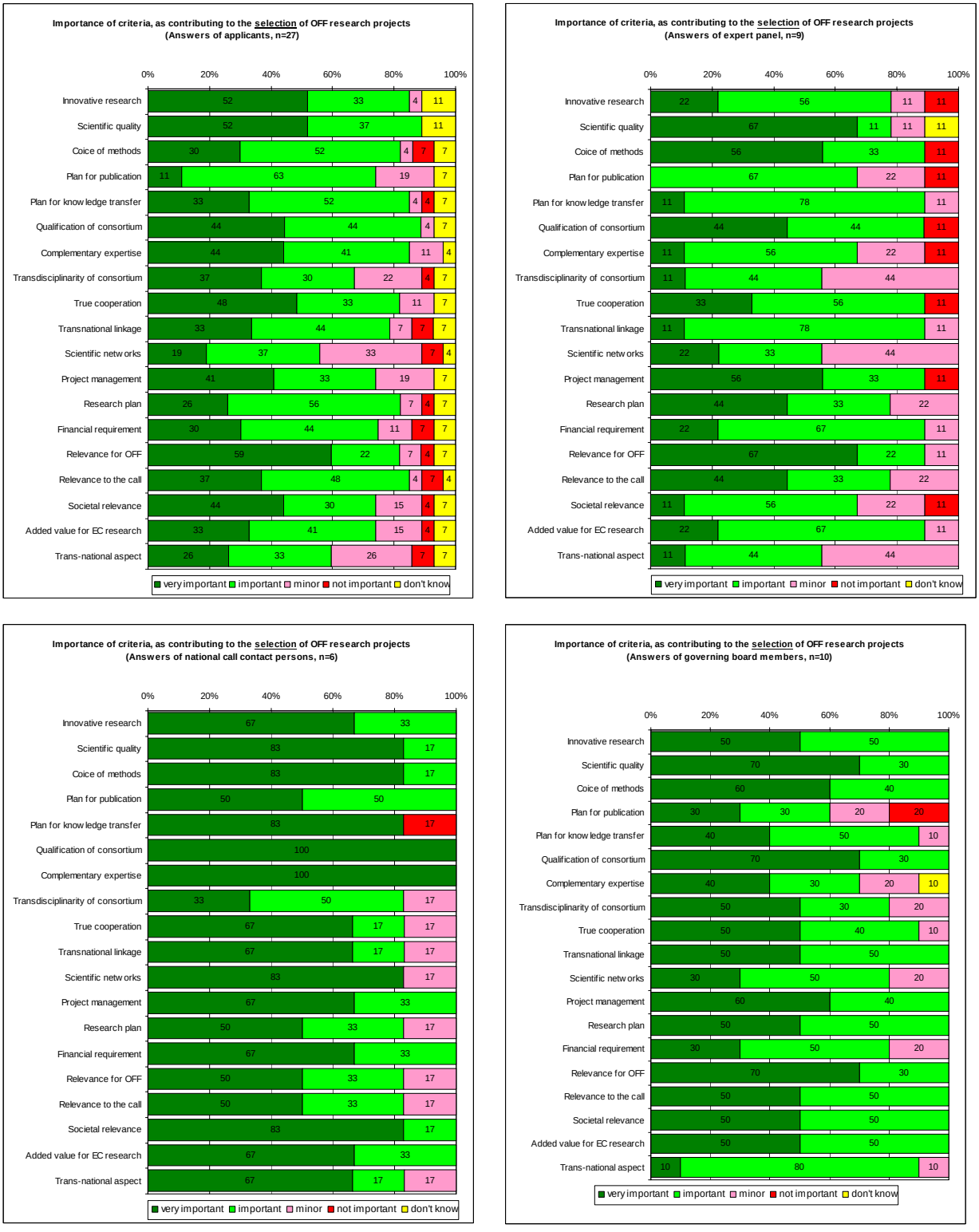


Fig. 1: Importance of evaluation criteria for the selection of organic food and farming projects as assessed by applicants (top left), expert panel members (top right), national call contact persons (bottom left) and governing board members (bottom right).

Tab. 3: Overview of the importance of evaluation criteria for the selection of organic food and farming projects as assessed by different target groups.

Evaluation criteria	Appli-cants	Expert panel	NCCP	GB mem-bers	Remarks
Scientific Innovation					
Innovative research	++	+	++	++	Lower importance of the expert panel compared to the other target groups
Scientific quality	++	+	++	++	
Methodology					
Coice of methods	++	++	++	++	Very important to all
Plan for publication	+	+	++	+	Very important only to NCCP
Plan for knowledge transfer	++	++	++	++	Very important to all
Consortium					
Qualification of consortium	++	++	++	++	Very important to all
Complementary expertise	++	+	++	+	Higher importance for applicants and NCCP
Transdisciplinarity of consortium	+	-	++	+	Criterion is judged from less impor-tant (experts) to very important (NCCP)
True cooperation	++	++	++	++	Very important to all
Transnational linkage	+	++	++	++	Lower importance for applicants
Scientific networks	-	-	++	+	For NCCP and members of the Governing Board more important
Project Management					
Project management	+	++	++	++	Lower importance for applicants
Research plan	++	+	++	++	Lower importance for expert panel
Financial requirement	+	++	++	+	Higher importance for experts and NCCP
Relevance					
Relevance for OFF	++	++	++	++	Very important to all
Relevance to the call	++	+	++	++	Lower importance to experts
Societal relevance	+	+	++	++	Lower importance to experts and applicants
Added Value					
Added value for EC research	+	++	++	++	Lower importance for applicants
Trans-national aspect	-	-	++	++	Very important to NCCP and mem-bers of the Governing Board, less important to applicants and expert panel

++ = very important: criterion is judged to be very important/important by >80% of the target group; + = important: very impor-tant/important 80-60%; - = less important: very important/important 60-40 %

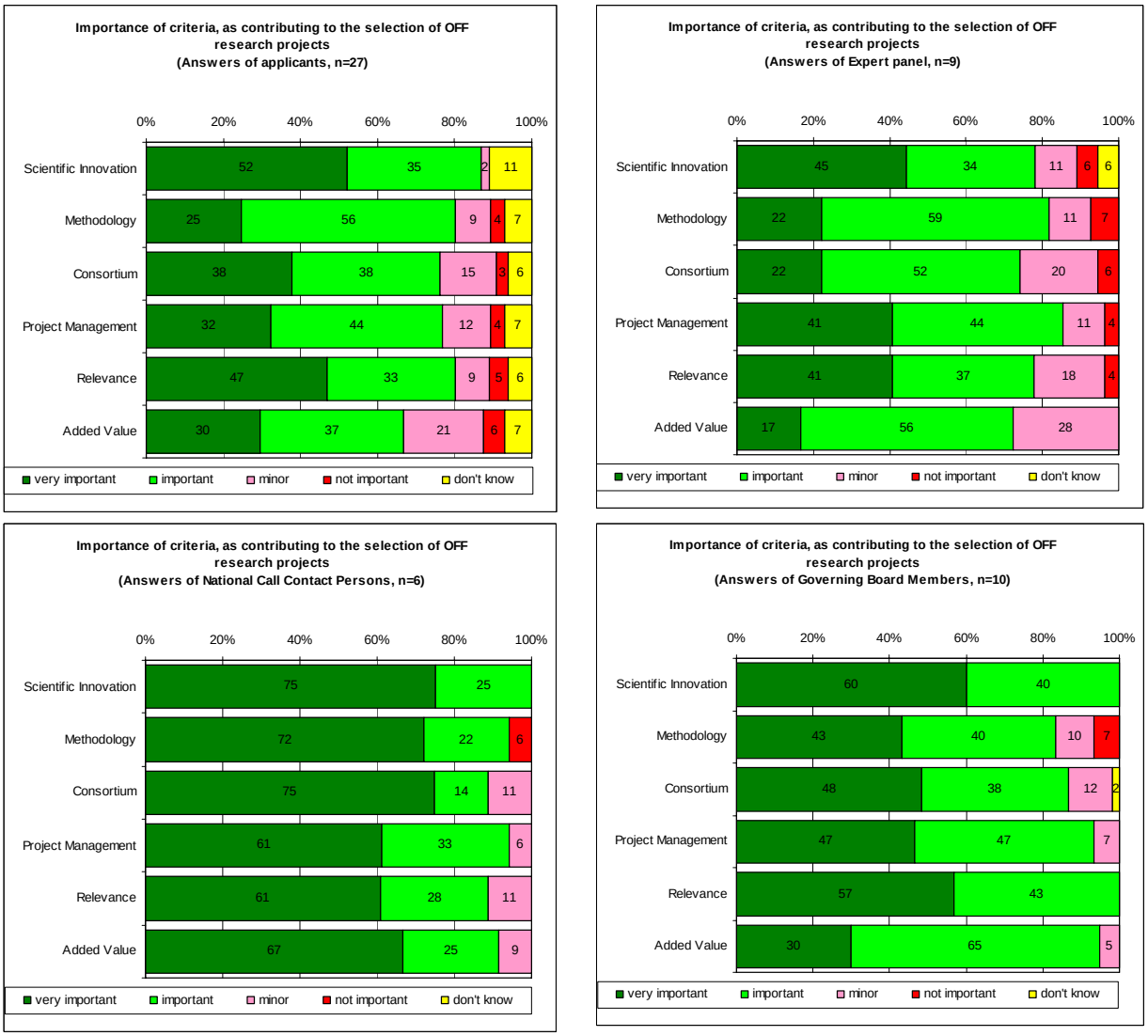


Fig. 2: Importance of evaluation criteria for the selection of organic food and farming projects as assessed by applicants (top left), expert panel (top right), national call contact persons (bottom left) and governing board members (bottom right).

4.1.2 Suitability of the evaluation criteria

The expert panel members were asked about their experiences regarding the suitability of the criteria chosen for the pilot call. Six of nine experts thought that the criterion *relevance for OFF* is easy to assess based on a 6-page proposal (Fig. 3). Also *qualification of the consortium*, *scientific quality*, and *relevance to the call* were mainly judged to be easy to assess. All other criteria are considered to be rather difficult or even very difficult. For six criteria, one expert thinks that an evaluation based on a six-page proposal is not possible.

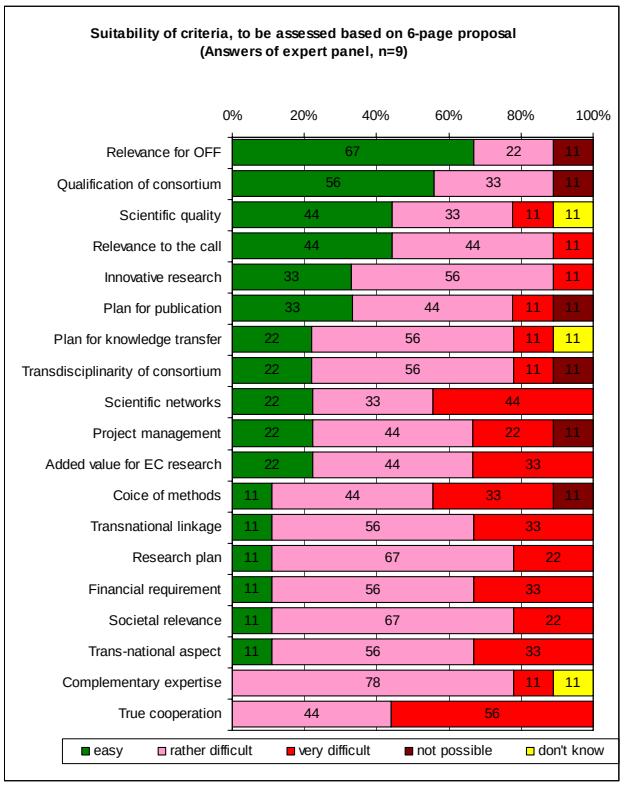


Fig. 3: Suitability of the evaluation criteria as assessed by the expert panel.

4.1.3 Overall impression of the pilot call based on the evaluation criteria

The expert panel members were asked about their impression on how the submitted proposals of the pilot call generally fulfilled the evaluation criteria. The expert panel members had a positive overall impression of the pilot call regarding the criteria *relevance for OFF* and *relevance to the call*. For a second group of criteria four out of nine experts thought the average quality of the proposals was sufficient. Regarding the criteria *complementary of expertise*, *transnational linkage* and *true cooperation,* two experts even had a bad overall impression of the proposals submitted in the pilot call (Fig. 4).

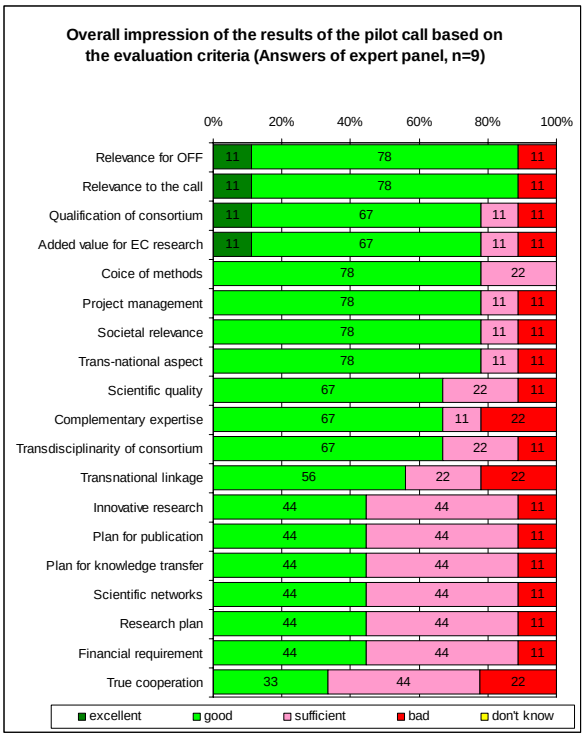


Fig. 4: Overall impression of the results of the pilot call as assessed by the expert panel (right).

4.1.4 Scientific evaluation and final selection of OFF research projects

One of the experiences of the pilot call was a discrepancy of the scientific evaluation and the final selection influenced by national research priorities. Fig. 5 shows that different evaluation criteria became less important when members of the Governing Board had to take the final decision. This difference is especially strong for the criteria *societal relevance*, *innovative research*, *plan for knowledge transfer*, *complementary expertise* and *true cooperation,* whereas the criteria *qualification of the consortium*, *scientific quality* and *relevance for OFF* remain important criteria. However, 2 out of 10 members of the Governing Board consider even these criteria not important anymore when taking the final decision. 50 % say that *trans-national aspects* are playing a minor role.

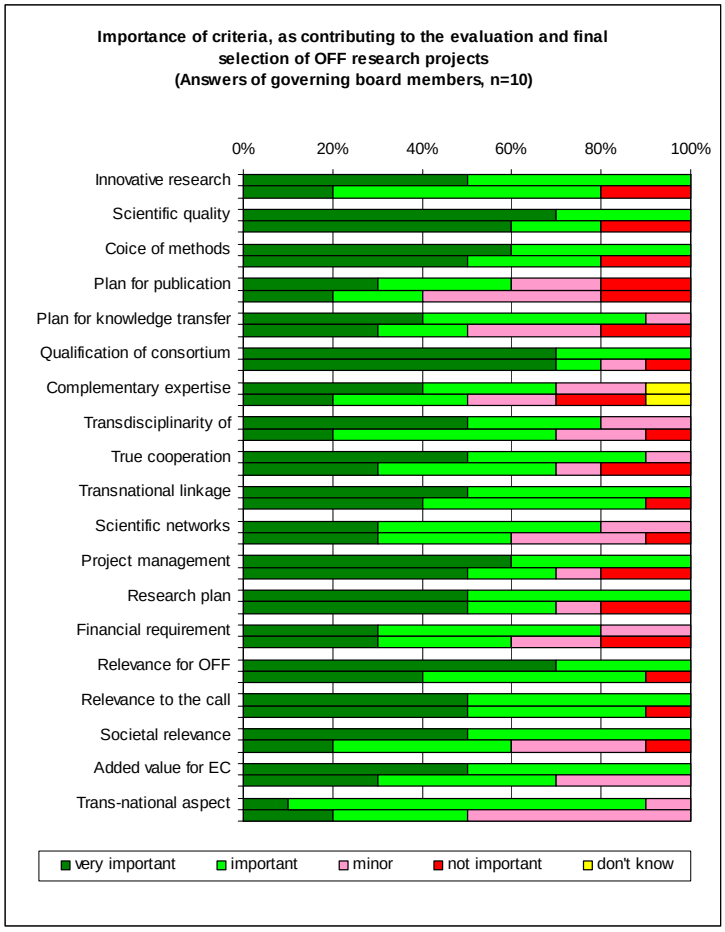
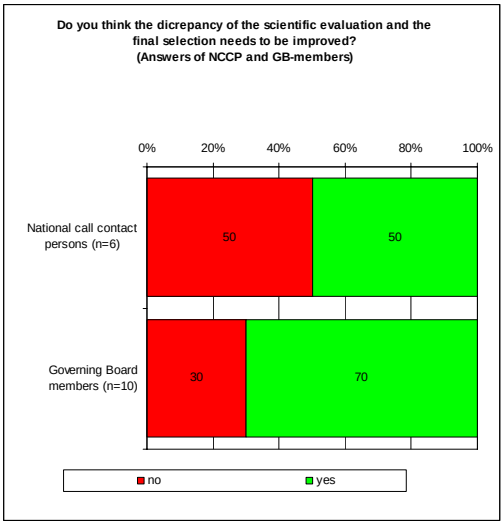


Fig. 5: Importance of evaluation criteria for scientific evaluation (upper bar) and final selection (lower bar) as assessed by the members of the governing board.

70 % of the members of the Governing Board and 50 % of the NCC-Persons wish to improve this discrepancy between scientific evaluation and final selection (Fig. 6). The reason for this conflict is mainly seen in the national research priorities. Some of the experts think that this problem can be overcome with a more detailed call description, less time pressure, a more even allocation of funding between partners or a better communication between the members of the Governing Board and the expert panel. Other representatives even think that a radical change from a national to an international perspective is needed by the decision makers. However, this point needs further clarification as 30 % of the members of the Governing Board and 50 % of the NCC-Persons see no need for action.



Comments by NCCP:

1) Some of the problems which came up in Ede [GB-meeting with final selection, 1-2. March 2007] were caused by the circumstance that we - prior to the application phase and evaluation phase - agreed on three common call topics and two countries immediately after this expressed their disinterest regarding one or two of these topics. That's why the agreement of call topics has to be improved and has to be more binding.~2) We published more or less very general topic themes. Under this general umbrella we received concrete proposals. One example: "Innovative Marketing strategies ..." are of course interesting for our country. After reading the I-Poppy-proposal we realised that parts of the work plan were already done in former national projects. This conflict we were able to realise only after reading the concrete proposal, not at the stage of call topic publication. Our suggestion to improve it would be to look for more detailed call topics. ~3) Due to the short time between the scientific evaluation and the meeting in Ede the time pressure led to insufficient national ranking procedures. This created some confusion during the Ede meeting which could have been avoided.

1) define clearly the individual steps of the whole evaluation process and in particular that of the final selection phase beforehand; ~2) give the details in the applicants guide~3) stick to the decisions made beforehand; ~4) if national priorities can overrule scientific evaluation, state this in the applicants guide (this will prevent annoyance by applicants)

In the selection of the projects the national priorities had too much impact. Also, we did not stick to our statement that all of the countries had to approve an application otherwise it would have been left out. (I personally think that you should be able to leave a country out and then re-write the application, but that should have been clear from the beginning). But the most important issue - projects selected was not always what we had asked for in the call text, i.e. we did not select any "parasite project", although this was a prioritised area.

Comments by members of the Governing Board:

National Research priorities should NOT count at the stage of final selection.

More time, better defined criteria, transparent procedures, all partners should stick to the rules defined by the project team.

A more even allocation of funding among partners.

National representatives should shift to look at international research priorities.

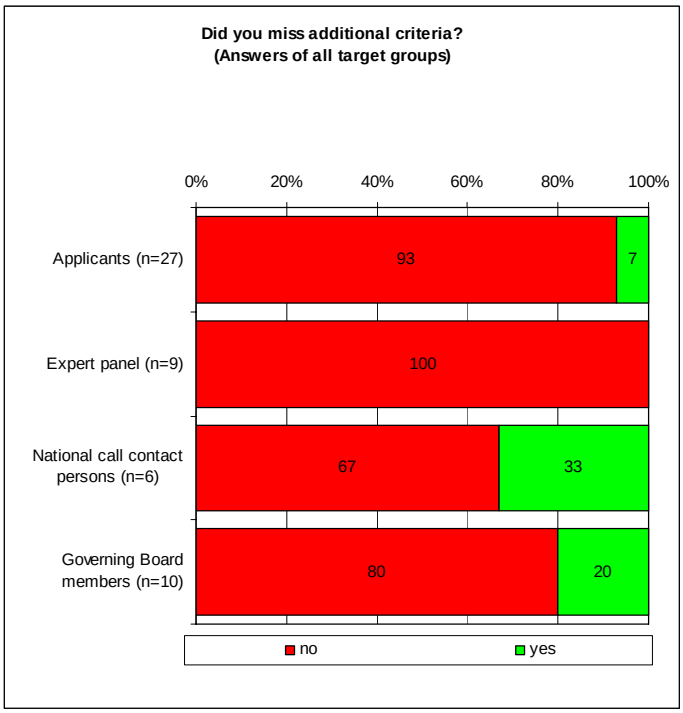
Better communication between expert panel and GB and more time for the discussion and final decisions.

Fig. 6: Responses of NCCP and members of the Governing Board on the question whether and how the discrepancy of the scientific evaluation and the final selection needs to be improved. Comments of those answering with “yes” are listed.

4.1.5 Additional evaluation criteria

It has been discussed elsewhere whether the criteria used for conventional farming research projects do meet the special aspects of OFF research projects. It was one of the aims of CORE Organic to clarify this aspect. Additional criteria especially with regard to *transdisciplinarity*, *true cooperation* and *complementary expertise* were included into the evaluation of the pilot call. In the survey, all target groups were asked whether they think that additional criteria are missing.

None of the members of the expert panel missed additional criteria (Fig. 7). Two persons among the applicants proposed that experts from outside the “classical organic scientific arena” especially practitioners should be involved in the evaluation of the proposals, in order to improve the practical relevance of the projects. We conclude that the question by whom and how the criteria are applied is as important as the set of criteria itself Two representatives of the NCCP and two members of the Governing Board commented that they wish additional criteria. They underline the need for a broader interdisciplinary approach and sufficient criteria for evaluation. However no concrete proposals how to realise this are made. The other two comments also suggest enlargement of the spectrum of the expert panel and are in line with the suggestions by the applicants mentioned above.



Comments by applicants:

Give the chance to experts outside the "organic classic scientific arena"

It is very important that the project is important for organic farmers

Participation by practitioners (organic farmers, farmer organisations, cooperatives, etc.). Their participation would ensure knowledge transfer, taking into account issues important to practitioners and transdisciplinary research approach.

Comments by NCCP:

The holistic claim both of the project goal and the work plan of the consortium should play an important role. We stress out a need of a broader interdisciplinary approach in new transnational research groups. Following this we also need sufficient criteria to evaluate the realisation of this claim in the proposals given in.

I think the Plan for communication and knowledge transfer should indicate more clearly how researchers intend to transfer the research results appropriately to stakeholders (esp. farmers and policy makers) so that they can really benefit from the research results

Comments by members of the Governing Board:

1. Involvement of other researchers from outside the well known "organic community" (e.g. health professionals, soil scientists, climate experts, nutritionists) in order to integrate as much expertise as possible and needed for superior research questions~2. Clear criteria in order to evaluate the real interdisciplinarity of the consortium, the interdisciplinary approach of the work plan and the interdisciplinarity of the cooperation

More emphasis on knowledge transfer, dissemination and implementation of results.

Fig. 7: Responses of all target groups on the question whether they missed additional evaluation criteria for scientific evaluation. Comments of those answering with “yes” are listed on the left side.

4.1.6 Additional comments

At the end of the questionnaire, participants were asked for an overall comment on the CORE Organic pilot call. These comments are listed in Table 4.

Table 4: Additional comments by all participants of the survey.

Target group	Comment
Applicants	To be honest - it never happened to me as a coordinator that I was treated so badly. I have coordinated a lot of national and international projects and was member in many consortia - something like this never happened to me - at least it confirms what I have been thinking about organic and biological farming since years ... no newcomers wanted, no new problems, everything is fine with organic farming ... 6 pages is too little for a good project description for international projects with many WPs. 8 pages would have provided space for illustration(s), and more air to make the text easier to read. Make the proposal evaluation documents and selected referees public, per each proposal, either accepted or rejected. People have the right to know why their work is accepted or rejected. It seems that some motivations were strumental to reduce the scientific impact of the projects.~Due to the lack of space in the proposal, many details were not specified and this did not seem so relevant for the quality of project. I feel it is most useful to combine trans-European and national perspectives, to use European wide understanding but to apply this to specific contexts; the national contexts are never the same and need particular approaches, but it is fruitful to proceed on the European level. My 'radio button' evaluations may not always be that appropriate because I have not followed the process so keenly.
Expert panel	I am very satisfied to take part in this evaluation and I'd like to repeat this experience in the future. It would be helpful to have done this feedback exercise a bit nearer the time when it was still fresh in our minds One starting general meeting could be useful to discuss and give common evaluation criteria to the experts before starting the evaluation procedure. Again, experts would like to know the final evaluation, who is or not funded.
NCCP	National priorities always need to be full filled. We should however have been better to evaluate added value (C2) and thereby have been able to reach relevance on a transnational level. The research community for OF in our country is rather small and needs ventilation. The experience so far, indicate that the continuation of the network would be important for the strengthening of this community. 1) all steps should be clear and transparent (decisions should be laid down in written form)~2) better project management with respect to the timing~3) make a roadmap and set deadlines for individual actions~4) most actions require more time than anticipated!~5) consider procedures and possible consequences of decisions more carefully ~6) check every now and then whether roadmap is still valuable or has to be updated (set milestones)~7) adjust timing of individual WPs (correlation of timing between individual WPs) --> improve networking between WPs (are milestones and timing between different WPs compatible?)
GB-Members	establish commonly agreed clear and transparent procedures and processes (and stick to these)~timlely involvement of all partners concerned~ a success story, well outlined and performed Partly, it was difficult to understand (and accept) the comments of the expert evaluations and there grading (Some proposals were criticised, but they got a Priority 1). ~For the future we recommend a two-step-procedure for the application phase. For the first step we should ask for a 6-page project description as we received in the Pilot Call. There should be a first selection at this stage. As the second step we should ask for a detailed description covering planned work packages, methodology etc. The possibility to use the web based forms and application procedure coming from Sweden was very convenient. In the future we should have forms in English which stay in English in the printed version.

4.2 Analysis of the evaluation of the pilot call

4.2.1 How submitted proposals fulfilled evaluation criteria

A detailed description of the evaluation procedure is provided by Geber *et al.* (2007). All nine experts evaluated all 36 proposals submitted. Nineteen sub-criteria were summarized in six main-criteria (see Tab. 3). However, scores were only given on the level of the main-criteria. Fig. 8 shows that after the first evaluation round, the results between the nine experts differed significantly. Experts 6 and 7 scored the proposals generally higher than did the other experts. On the other hand, expert 3 judged generally lowest and expert 5 covers nearly the whole range of possible scores. This broad range of scoring behaviour of the panel members is considered to be positive and indicates that the proposals have been discussed under many different aspects (see also chapter 4.3.2).

During the second evaluation round, other factors might have been influencing the final list of prioritised proposals. It was observed that project proposals which were presented to the panel by a previously appointed main rapporteur, had better chances for success than proposals which were presented by a vice-rapporteur due to the absence of the main rapporteur.

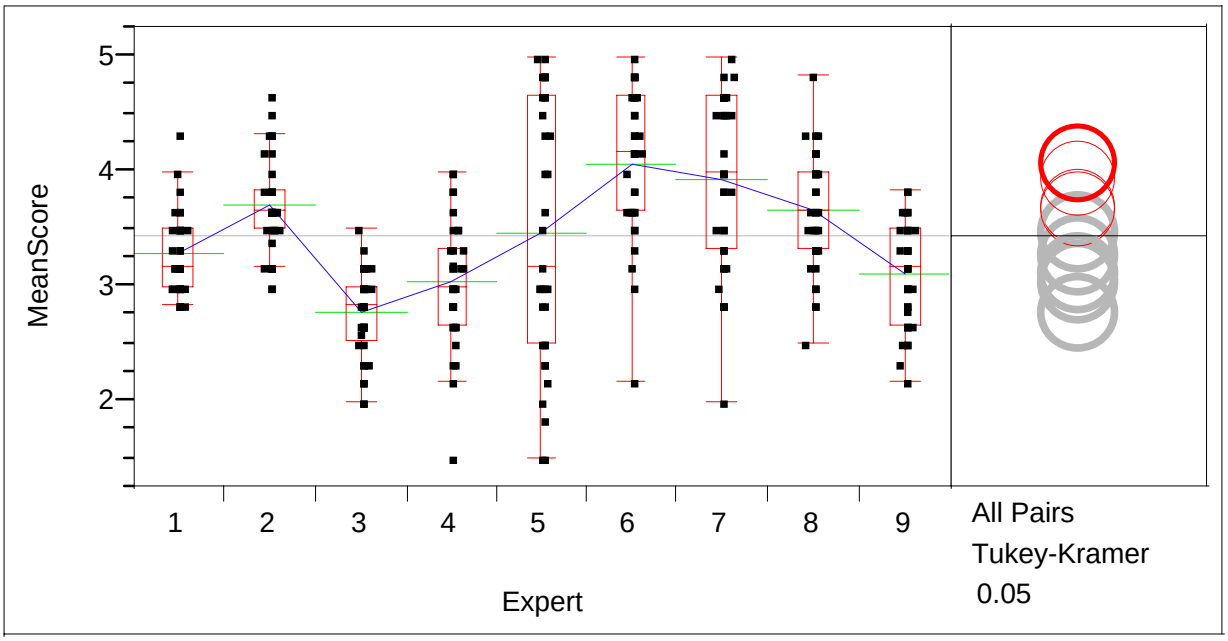


Fig. 8: Evaluation of 36 projects (black squares) of the CORE organic pilot call by 9 experts. (One-way analysis of mean score by expert)

Fig. 9 (left) illustrates the results of the 36 project proposals of the pilot call for each main category of the evaluation criteria. Projects with priority 1 convinced with high scores especially for the criterion *consortium*. This main category includes the criteria *transdisciplinarity of consortium*, *complementary of expertise* and *true cooperation*. As the scores were given only on the level of the main category, no further analyses can be made regarding the important criterion *transdisciplinarity of consortium.* It could be possible, that the criteria covering these interdisciplinary aspects were mixed up with the other important criterion *qualification of consortium* which is in the same main category. Projects in priority 1.5 were scored lower regarding the criterion *methodology* but *added value* was higher. Projects of priority 2 and 3 received lower scores for all criteria. The main reasons why proposals were classified as priority 3 were low scores for the criteria *methodology*.

The animal and marketing projects generally convinced regarding the criteria *consortium*, *relevance* and *added value* (Fig. 9, right). Lower scores were achieved regarding *innovation*, *methodology* and *project management*. Projects on food quality differed much less within the six evaluation criteria.

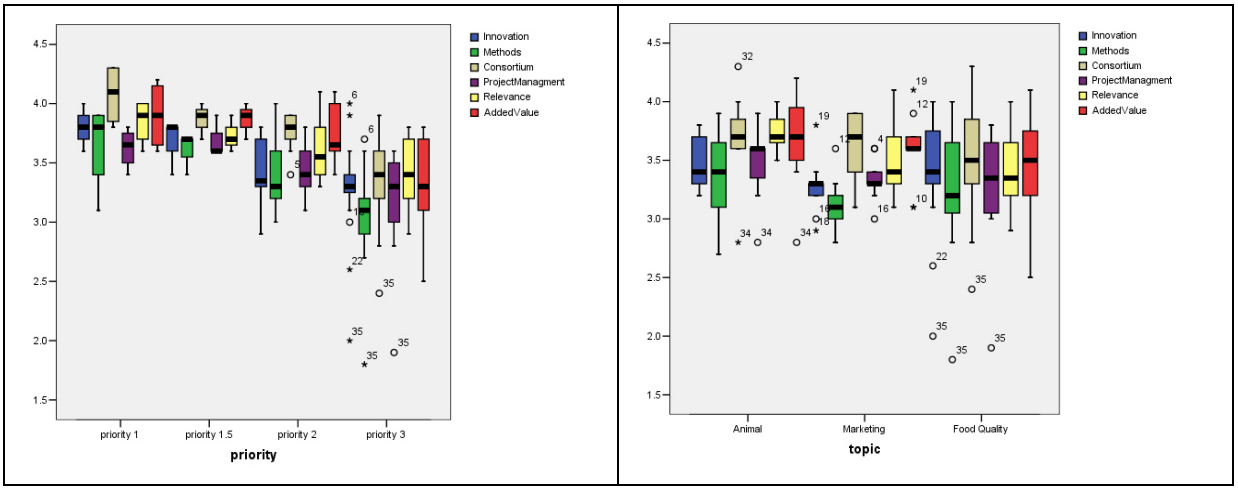


Fig. 9: Box plots of means of the evaluation of 36 projects of the CORE organic pilot call referring to the priority categories (left) and the topic categories (right).

4.3 Review on evaluation criteria and procedure

4.3.1 Introductory remarks

Both the background of scientific evaluation of OFF projects and the main results derived from the analysis of the experts' feedbacks were taken as starting points for this literature review.

The main issues raised are as follows:

1. Interdisciplinarity and innovation are not sufficiently addressed, i.e., the promotion of these aspects and of the assessment procedures should be improved.
2. The concern for cooperation and trans-national aspects, aiming at strengthening European research, should be better addressed.
3. The assessor's provenience and whether the decision process should be open to non-scientific experts and non-governmental representatives should be discussed.
4. The transparency of the procedure should be improved.
5. The discrepancy between the scientific evaluation and the final selection, i.e. the way how national priorities are integrated in the decision process, should be discussed.

4.3.2 The interdisciplinary and innovative research concern

Interdisciplinarity was first discussed in the research evaluation literature in the mid 1980s (Porter and Rossini, 1985; Travis and Collins, 1991) but the concern has particularly grown up the past decade.[2] The main problem is how to ensure that interdisciplinary research is not the looser in the assessment process of research proposals (Laudel and Orrigi, 2006). But what specific criteria and/or specific procedures should be implemented to promote good interdisciplinary research?

Confirmatory bias of reviewers

The difficulty of promoting interdisciplinary research is based on the difficulty of dealing with complex topics and facilitating innovative research. An interdisciplinary research project will usually suffer from a conventional peer review process, known for its conservative and risk minimising aspects. As Hacket and Chubin (2003) recall, a peer review process embodies Kuhn's (1977) "essential tension" between originality and tradition in science, as it is easier to stabilise a research theory in a specific area than to promote a novel one. It is often that novel ideas are judged as impractical, unworkable or implausibly inconsistent with the established body of knowledge (Hacket and Chubin, 2003). This phenomenon is called the "confirmatory bias", i.e. a tendency of some reviewers to accept outcomes that agree with commonly accepted theories and to discredit those that do not (see Hojat, 2003 for example). One of the principal components of the weakness of innovation facilitation through peer re-

[2] The journal "*Nature Sciences Sociétés*" favours such debates and launched specific forums or issues (e.g. Vol. 12 n°1, Jan-March 2004). In April 2006, the journal "*Research Evaluation*" released a special issue on this topic of interdisciplinarity, joining contributions of many "experts" of sciences studies.

view is the difficulty, or even resistance of established scientists to question their accepted paradigms. This may lead to the misconception of novel ideas, may scare scientists for the damages and distortions that innovative ideas may cause within the established theories and research fields. And from a more "social" point of view, it may challenge interest in the monopoly of learning (Atkinson, 1994; Hacket and Chubin, 2003).

We can address the tension between innovation and conservatism from a more technical point of view. A review system that favours rigor would support only those studies that use the strongest research designs and analytic approaches. In contrast, a more responsive system would relax its methodological standards to address "new" problems and question, even if they are ill-structured (Hacket and Chubin, *op. cit*.). That's why Boix Mansilla (2006) promotes to give a temporary credibility to the innovative works.

The experience of the CORE Organic pilot call shows, that interesting ideas were identified by experts in projects submitted under that call, but rejected due to the risk attached or to an unclear approach. Such ideas and topics could be integrated in the identification of further topics and prioritisation of organic food and farming research (see WP6).

Diversity of the Expert panel

Porter and Rossini (1985) already stressed in the mid 1980s that it is difficult to identify peers whose expertise fully encompasses the proposed cross-disciplinary research. Therefore large pluri-disciplinary panels are admitted to be the more efficient in evaluating interdisciplinary research[3]. When the peer panel is well balanced between disciplines involved in the proposals, the panel system, allows a broad representation of divergent judgments and conflicting validation norms. Hacket and Chubin (2003) confirm this view stressing that the combined assessments of several diverse experts are needed to achieve a rounded evaluation of a proposal. It is the only way to take in account the scientific merit of a proposal in all its complexity (Hackect and Chubin, 2003).

As Porter and Rossini say, the use of an expert panel offers the further advantage of allowing open debate about criteria assessments and discussions helping each of the members understand unfamiliar aspects of a proposal under review and enabling them to change their mind after discussion. This matches the view of Langfeld (2001) who compared different reviewing models, and showed that both the rough rating-scales and open-decision making process within the panels usually give ample room for research policy considerations and bring support to innovative project. As he says, "enthusiastic panel members have room to express themselves and manage to change the panel's view on projects that first could be seen as too risky or peripheral or immature".

The existence or a set of identically rated possibly funded proposals thus seem to be a central condition for giving priority to research with special needs, to strengthen pluralism and facilitate funding of controversial innovative and interdisciplinary projects (Langfeldt, op.cit.). In this perspective, low inter-reviewer agreement on a peer panel is not an indication of low validity or low legitimacy of the assessment. It may rather indicate that the panel is highly competent because it represents a wide sample of the various views on what is good and

[3] A panel of 8 to 12 expert is shown to be a good number.

valuable research (Harnard, 1985, Hacket and Chubin, op.cit.; see page 17, Fig. 8). Different experts might properly reach different judgments about the quality of the proposal when their particular area of concern is given central importance and evaluated through their particular set of epistemic lenses. The challenge is to find a diversity of expert that encompasses the various facets of a set of proposals and to avoid duplicative perspectives.

In the identification of country experts evaluating the pilot call, it can be assumed that the following prerequisites were acquired: each expert had a basic knowledge of OFF, had been involved in OFF research projects mobilising interactions with other disciplines (systemic and interdisciplinary approaches), had expertise in at least one of the three identified topics.

Decision process

However, various authors have argued that collective decision process is rarely leading to balanced decision. Press *et al.* (1979) for example, stressed that a group decision process had to be released from the effects of face-to-face interaction pressure and inequity. He developed a Qualitative Control Feedback system to assist policy makers in forming judgements and making decision that reflect the careful interactive reasoning and arguments of all the members of a group. Others argue that voting systems are not fair. In this perspective, Saty developed a method that facilitated the group process to capture preference intensities of individuals and incorporates them into a final group decision (Saty, 1980). Highlighting that the current voting system oversimplifies the representation of voter preferences and “drowns out” the true merit of counter arguments, Saty and Shang (2007) invented a framework for reshaping the group decision process. Finally, many computer based environments were developed to support group work, as Ngenyama *et al.* (1997), for example, who proposed a software to assist facilitators with the task of clarifying positions of participants, identifying openings for dialogue and developing strategies for moving the participant consensus. All these perspective show that if panel reviews systems are thought good assessment implements by allowing open debate about criteria assessments, they are not free of bias due to human factors.

The “consensus building” approach is widely used in the evaluation of EU projects (among others), and implemented through panel discussions. It was also applied in this pilot call and proved satisfactory, following experts judgement (“fair” evaluation).

Another tendency that is more and more taken in research evaluation is the feedback process, or more broadly the “longitudinal involvement procedure, supporting thus a learning process”. Porter and Rossini, in the mid 80s did already suggest that incorporation of feedback in the review process could help remedy the problem of restricted reviewer expertise (Porter and Rossini, 1985).

Such a feedback appears as important not only for the applicants (in order to adjust their project, to re-submit it in another call) but also for the experts (reflecting on how their assessment was taken into account, fine-tuning of their individual and collective report, sharing a common literature base).

Empowerment of applicants

These procedures are, somehow, open practical critics of the peer review seen as a rational decision-making process in which various reviewers apply a set of 'objective' criteria consistently. Rather, as many authors argue, peer review would be a process of collective construction of knowledge claims and reviewers are central actors in the definition and redefinition of "good research" (Langfeldt, op. cit.). Laudel (2006) goes even further and stands that peer review must be analysed as a negotiation and knowledge creation process in a complex actor constellation. Analysing two peer review processes of collaborative networks in Germany, he advocates the empowerment of applicants and the enforced interdisciplinary learning of reviewers through a "longitudinal involvement" process of the actors (Laudel, 2006; Klein, 2006). According to Klein it generates not only competence but also a communication base that increases the number of people capable of conducting interdisciplinary evaluation with interdisciplinary rigor (Klein, op. cit.). G. Laudel and J. Klein are part of these authors who emphasize that assessment of interdisciplinary work needs special institutional rules of assessment rather than special criteria.

This issue relates with possible shifts due to differences in individual experts' interpretation of interdisciplinarity (or transdisciplinarity, which was included as one of the criteria for the consortium) and the feasibility of ex ante assessment of such a criteria, based on a brief project description.

4.3.3 The value laden concern

In a recent study on organic farming grant applications, Rasmussen *et al.* (2007), compared the assessment of grant proposal by peers that have a strong experience in OFF and peers that don't have, i.e., that are more particularly qualified in conventional farming. The authors show that reviewers affiliated with organic farming reviewed grant applications differently than those reviewers without this affiliation and stress that this bias is probably linked (even if not demonstrated in the study) with unshared paradigmatic references, i.e. different knowledge, value and beliefs systems. One of the specific findings that drove them to these conclusions is that within the application assessing criteria, measures of "quality" and of societal relevance were highly correlated, showing thus that social values and beliefs did affect the review process: "It can be hypothesized that the perception of relevance influences the scientific quality assessment." The authors stress that this is a specificity of OFF, emphasizing that in the alternative knowledge and learning paradigm "contextual values are thus believed to enter into the very process of science". Actually, for many authors, OFF is rooted in a "alternative paradigm"[4], opposed to the dominant paradigm which would be the realm of the conventional agriculture and shows a different point of view of the status of science. In the dominant paradigm, objectivity and facts are opposed to subjectivity and values and the role of the scientist is assumed to be unbiased and impersonal. The alternative paradigm recognizes limits to conventional science and stresses the necessity of integrating values and beliefs into the learning processes (Wynen, 1996; Francis and King, 1997; Packham and Sriskandarajah, 2005; Rasmussen *et al*., 2007). Alroe and Kristensen (2002) expressed the

[4] Paradigms are defined as « coherent frameworks of knowledge, values, and beliefs within which experiences are interpreted and made meaningful. Rationality is created within paradigmatic frameworks ».

same ideas, though stressing that agricultural systems research, in general, is inherently framed in a social context, and necessarily involves questions concerning different interests and values in society as well as different structures of rationality and meaning (Kristensen and Halberg, 1997).

Actually the previous viewpoints match a broader opinion that value judgments, including some of an epistemological or methodological nature, are present in any scientific practice. Actually, sociology of science has demonstrated over the last 30 years that scientists are not immune to all external interests (Calvert, 2001). Habermas (1971) already stressed this fact criticising the positivist representations of the pure and interest-free activity of science and expertise. Peer review practice cannot be seen as being the objective, dispassionate process that its advocates. Equally, the notion that science-based regulation can be lifted above politics and ideology through peer review appears seriously misguided (Van der Hove, 2007).

Scientists are not "neutral". Their research and expertise practices are embedded in their scientific paradigmatic realm, in their "culture" with its values and beliefs and, nested to specific interests (political, economical, etc…). It is thus not surprising that the results of Rasmussen *et al.* show that using classical, too simple and usually badly defined criteria as "scientific quality" and "societal relevance" is problematic. As Alroe and Kristensen say research evaluation and expert show "usual difficulties handling separately and independently the criteria of relevance because the process still aspire to the scientific ideal of being value-free and independent of social interests. Research evaluation cannot stay blind to this state of science".

Stressing that, Alroe and Kristensen promote a critical perspective, where "self-reflection" would be very central, and stand for an open communication and description of the value-laden starting point of any research proposal. In order to facilitate peer criticism and the use and critique by different users and stakeholders, they think that research should describe the choices made in research initiation, the delimitations and constraints these choices place on the results, and the areas of ignorance that this implies, as an essential context of the results produced. They underline too that the communication of the full cognitive context is an important precondition for better cooperation between different kinds of science. They match Van den Hove opinion, who advocates that the normative requirement stemming from the fact that scientists cannot be stripped from their values and interest is that knowledge holders and producers should render explicit their values, their ethics and their interests (Van der Hove, *op. cit*.).

There is a growing recognition that the division between science and society is being renegotiated or even vanishing (Salter and Martin, 2001; Nowotny *et al*., 2001). One of its expression is that scientist are no more perceived as value-free and objective truth seekers. From a broader perspective, all the scientific institutions ruling the science process are unavoidably value-laden: "Value choices contribute to orientation of research programmes, they are part of the validation process, they inform the design of education and training systems, and they contribute to the emergence and operations of networks." (Van den Hove, op. cit: 813).

Although this value-laden concern seems to be more prevalent in literature on OFF research, it also appears as relevant for research more generally. If we admit that a technological renewal is at stake, with alternative performance criteria to define for agricultural systems and activities, there is also a challenge in defining new performance criteria for research.

4.3.4 The science-policy frontier is blurred

To acknowledge that the division between science and society is not so clear-cut, and that science practice (and expertise) is part of a broader social and political process, entails that science cannot anymore claim its sole authority in producing knowledge. For Ravetz and Funtowicz, we have entered a "post normal science" era. The socially embedded character of the problems that scientists are dealing with, means that many non-scientists should have substantive expertise on the topic under focus: Scientists have to draw on "extended peer communities" (Ravetz and Funtowicz, 2001). Interactions with this extended peer community can help to ensure that researchers are addressing relevant questions, incorporating the knowledge of these non-academic experts in the analysis and adequately testing the validity and practicality of any prescriptions researchers are proposing. As a result, technological change is embedded with societal concerns and implies new alliances.

This is a perspective strongly advocated by many authors practicing sustainability science[5]. This field which has been constituted around the problems of sustainable development calls for novel forms of research and science-policy interaction through stakeholder participation and the integration of their diverse forms of knowledge and expertise (Blackstock *et al.*, 2007; Lucks and Siebenhuner, 2007). According to Clark *et al*. (2005), promoting the objectives of sustainable development means that science has to develop knowledge in an embedded way with both societal stakeholders and policy makers. Furthermore, these authors stress that non-state actors of different kinds are increasingly crucial for the promotion of cures and innovative solutions to environmental and social problems.

Scott (2007) recalls that quality criteria in evaluation process are now increasingly both internal and external and that some funding bodies have moved from peer review to the broader concept of 'merit' review indicating that they are taking account of non-specialist opinion on the relevance of the research to "socioeconomic problems." We would like to stress here, that this phenomena cannot be isolated from a tendency of science being increasingly driven towards issue-driven approaches and away from curiosity-driven research.

Nowadays knowledge networks no longer include only producers' knowledge, but consumers' knowledge too, and involve more and more different types of organizations—public, private, and hybrid—with different types of goals and changing structures (Callon 1998; Ziman 2000). New organizational forms are emerging in the system of knowledge production. Concepts such as "mode 2 knowledge production" (Gibbons, 2000) and "triple helix" (Etzkowitz and Leydesdorff, 2000) support the ideas that boundaries and connections between academia and society have already been restructured. Since cooperation, transdisciplinarity and boundaries working process are getting crucial in knowledge and science development, improvement is needed in assessing the research environments, the social networks, i.e. to assess the management of connections and boundary dynamics. Consequently, changes in quality control are emerging and this process must continue to be strengthened. They lead and support the development of new criteria and procedure enabling to evaluate collaboration, networks and organizations, and not only individual researchers. They lead and support

[5] Sustainability science has been defined as a "new field (…) that seeks to understand the fundamental character of interactions between nature and society".

to an opening of “expertise” to evaluators that are not only scientific peers but also specific consultants (Hemlin and Rasmussen, 2006).

Moreover, the project level is not identified as an evaluation unit *per se*. If we admit that project consortia share some similarities with research units, a new context is interestingly set in France to renew the evaluation of research units. Three categories of criteria are suggested to evaluate research units: (i) classical outputs derived from scientific and technical activities (intrinsic quality of peer-reviewed literature, patents, software, expert reports, tools to support decisions…), (ii) processes of research in partnership (means-based evaluation including selection of partners, research topic definition and evolution, reporting…), (iii) societal impact and behavioural changes related with research outputs (new regulations, partners organizations, employment generation in research or other domains..).

4.3.5 Conclusions and recommendations from the literature review

As shown in our literature review, the issues summarized in the introductory remarks are strongly interrelated and emerge from global paradigmatic changes in science and research. These changes have brought new complexities and uncertainties degrees in the field of research evaluation. This is one of the reasons why almost all the literature on the topic of peer reviewing or research evaluation addresses these new challenges from a conceptual perspective.

One of the main difficulties of research evaluation is that it has to assume many roles, even more now that it is admitted to stand at the “science-policy interface”. Trying to serve all thee roles peer review tries to solve many tensions at the same time: the tension between tradition and originality, between scientific autonomy and accountability, i.e., scrutiny by wider publics, and between meritocracy and opportunism, i.e., the research field has to stay open to new incomers, avoiding monopolies, but in the same time avoid opportunism of incomers seeking new financial perspective for old ideas.

These tensions may particularly difficult to solve in the Organic Food and Farming research field because of some of its main characteristics:

- Being a field under construction, Organic Food and Farming research has to deal particularly with creativity and innovation. This novelty aspect is even strengthened by the holistic research approach promoted by the OFF “paradigm”. Stressing the importance of existing “interactions” between the different socio-ecological components of our world it claims for a strong integration of disciplinary perspectives and the development of new methodologies to assess new research targets. However getting away from a minimum scientific robustness is not possible, even for the more innovative approaches.
- Being a field heavily loaded with social values and which development relies heavily on social movements, integration of knowledge users, stakeholders, development organization should be of prime concern. Actually, we must admit that this field calls for novel forms of research and science-policy interaction through stakeholder participation and the integration of their diverse forms of knowledge and expertise. At the same time, this point of view may strengthen the tensions between problem finding

and problem solving issues as Organic Food and Farming research field raised on a very urgent problem solving perspective and still strongly leads with this perspective.

Many different “routes”, i.e. specific attempts, have already been taken by the research evaluation practice to enter the new “era” of science-policy interface. The procedure used by the Organic CORE pilot call is one of these attempts. Indeed, the criteria and procedures used in this pilot call witness that some changes of perspective in research has been taken in account, as much as the specificity of OFF. However, criteria and procedures used in this pilot call show evident limitations that should be overtaken to improve the assessment process and, consequently to strengthen the research in this field.

A specific selection procedure, as the CORE organic pilot call one, can be seen as an attempt to solve the tensions above-mentioned within its own set of embedded constraints (from paradigmatic to financial, structural and political constraints). From this point of view, the “distance” of a concrete procedure from its “ideal form” depends strongly on the specific set of constraints it is embedded in. Thus, we can argue that a main challenge for improving a procedure such as the CORE organic one, is to identify the specific set of constraints that drives her away from a consensually (and conceptually) defined “ideal”. This embodies two tasks: to consensually define the “conceptual ideal” and to objectively identify the constraints, in order to get rid of them, as much as possible. Consequently, and according to the broad opinion that collective learning and reflexive attitude are strongly needed in any evaluation processes, we argue that open debates focusing and considering these tensions are needed to improve the proposal’s evaluation in the CORE organic project.

Narrowing this perspective, and according for both the analyses provided and the additional comments by the participant (chapter 4.1 and 4.2) of the survey we can argue that:

To strengthen innovation, specific mechanism should be implemented in order to allow few “risky” research projects to be funded, to facilitate new incomers to enter the arena and to promote curiosity-issue research projects. However, in this context, the challenge is to enable to distinguish between sound innovation and reckless speculation. Assessment process should thus include some “minimum” gate-keeping mechanisms in order to ensure minimum scientific robustness. We argue that basic quality criteria could allow this evaluation, if they are well defined (cf. next section). At the same time, we argue that a latter assessment step could be implemented for the projects identified as more innovative and less robust. This could be made implementing a tutorial of the ongoing research, results evaluation, ect.

The procedure should be driven with more transparency, with a wider interaction between national representatives - NCCP and GB members – applicants and experts. The national priorities should at least be clearly stated at the beginning of the process and articulated with the topics eligible for funding. This step of the process may include external “stakeholders”, which have to be previously identified. Furthermore, the decisions taken at both steps of the selection process - the peer panel ranking and the government funding bodies selection - should be motivated and made accessible to all actors of the process. Actually, we argue that the two-step selecting procedure is far away from the actual quality standards of selection procedures. As experts are expected to confront with one another their opinions and values in a panel meeting, NCCP and GB members should have to do so. If science and society are supposed to be so much linked with another, total transparency of the interest

should be implemented through open debates, both before the call and during the selection procedure. ERA-net, promoting excellence should be on the front line of these transparent and democratic procedures.

Integrating others assessors than scientific expert in the peer panel should be of prime concern. One simple reason is that many of the criteria may not be in the domain of competence of scientific experts. For example, within the main criteria Consortium, the experts admit that the only “easy to assess” sub-criteria is “qualification of the consortium”. One of the possible interpretations could that the experts themselves judge they don’t have the skills to asses these criteria. This would support the idea that others type of experts should be included in the assessment. This opens to a broader concern with including non scientific members in the panel evaluation, concern that appears in the applicants and GB comments of the procedure and that address too criteria “societal relevance” and “relevance for the OFF”, i.e. criteria that explicitly implies social values and beliefs. The same remarks can be made for the criteria block “project management and resource utilisation” which sub criteria are considered difficult to judge by the experts, even thought of high importance as underline in the guidelines of the call. There is here a high discrepancy between the “solidity” of the judgment and the importance of the criteria. This advocate the inclusion of special experts (management experts, organizational experts) in the process or/and the obligation of participants to orally defend their project management implementation.

5 Experts ranking and literature review

In the light of the literature review, we suggest that the repartition of the sub-criteria in the main criteria blocks should be rethought. Furthermore some additional criteria may be necessary to address the problematic issues (interdisciplinarity, transfer of knowledge, value laden). At the same time it seems that some criteria are un-clearly overlapping one another, and that more precision should be given to some of them. Refine criteria and give them more clear boundaries should be of prime concern. A review criterion by criterion shows that:

The main criterion Scientific innovation, with its two sub-criteria *Novel, innovative research* and *Scientific quality of the proposed project*, is the place to assess aims, hypotheses, novelty, new ideas, bold, cross disciplinary approaches, and knowledge of the literature. We do think that these categories are a little bit a lumber-room in front of the problems related in the literature, to assess interdisciplinarity and innovation or non conventional projects. The experts have to operate their own ratios and means between all these aspects and summarize them into two simplifying criteria. More sub-criteria should be included in this category, e.g., if the project is problem solving orientated or not.

The main criterion Choice of method contains diversified criteria, *methodology* corresponding somehow more to scientific quality, and others linked to dissemination, included the *plan for communication* and *knowledge transfer*. The choice of methods could be included in the scientific quality, which then would be more a “scientific robustness” criteria. The criteria *plan for communication* and *knowledge transfer* could be considered apart, and may include more specific items, for example: are there non scientific partners and knowledge consumers or any kind of others stakeholders included in the project? How do they take part in the project?

The main criterion Consortium expertise shows again heterogeneity and overlapping definitions of the sub-criteria:

In the guidelines, an important emphasis is given on the *qualification of the consortium*: antecedents of the individual researchers, and of the group, expertise of individuals and groups, *complementarity of expertise*, capacity to handle the project…, which globally refers to the scientific capacity in term of "skills" of the individuals and groups to handle the research.

Another group of criteria are emphasizing the "practical" capacity of the consortium to handle the project: *true cooperation*, *transnational links*, *scientific networks*, i.e its real capacity to implement the linkages it assumes.

Finally, there is a *transdisciplinarity of consortium* which aims to assess the level of transdisciplinarity. But transdisciplinarity is not explicitly defined. We underline that many debates are going on the questions of interdisciplinarity, pluridisciplinarity and transdisciplinarity, and that it should be made clear, what is intended here: the balance, the sum, or the integration of disciplinary approaches within the consortium.

We believe that these criteria have different epistemic natures, and recommend that they be separated in three sub-categories, some of which may have to be addressed by other type of experts.

Regarding the assessment of the criteria block project management and resource utilisation, additional experts (management and organizational experts) should be included. As already mentioned, the sub-criteria of this category are considered difficult to judge by the experts, even thought of high importance as underline in the guidelines of the call, stressing the need of other type of skills being included in the evaluation.

The main criterion Relevance with its sub-criteria *societal relevance* and *relevance for OFF* should include assessments of knowledge users (not necessary for the more technical sub-criterion *relevance to the call*). This is supported by the literature review and by the feeling of the experts that the assessment of Societal Relevance is difficult.

Finally, the criterion Added Value for EC is a transversal criterion, including evaluation of scientific networks quality, true cooperation and trans-national linkage, skills expertise of the consortium. Difficult though to address such a criterion that tries to assess, merely, the "emergent" components of the partnership, based on sub-criteria considered mostly difficult to be assessed themselves.

6 Conclusions and recommendations

Evaluation criteria

The survey showed that the proposed set of nineteen evaluation criteria, clustered within six main categories, fulfilled the expectations of most target groups involved with the CORE Organic pilot call. However, some of the respondents wish to have a stronger focus on aspects of interdisciplinarity. The actual list of criteria contains already three criteria dealing with different aspects of interdisciplinarity. It is suggested that these criteria could be regrouped into a new main category called *interdisciplinarity,* which otherwise tends to be underestimated in relation to the criterion *qualification of the consortium*. A similar splitting could be made for the main category *methodology*, in order to encourage researchers to apply innovative methodological approaches.

How submitted proposals fulfilled evaluation criteria

The analysis of the scientific evaluation of the pilot call showed that rejected proposals had received insufficient scores regarding methodological criteria. Proposals with high prioritisation showed higher scores for *relevance*. In order to improve chances to succeed in future calls, proposals should therefore focus on the criteria *methodology* and *relevance*.

Scientific evaluation and final selection

Handling the gap between the scientific evaluation and the final selection of OFF research projects has been identified to be the main challenge for a successful evaluation of transnational research in the future. This is expressed by some of the comments of the applicants. Furthermore members of the Governing Board argue that scientific evaluation criteria are less important for the final selection. However, only two members of the Governing Board clearly state that several criteria are not important at all anymore for the final selection. To overcome this conflict, different measurements are suggested such as (i) a more precise description of the call topics, (ii) a two step application procedure, (iii) a commitment of all participating members in the call to fund all topics, and (iv) a more transparent procedure defined beforehand, and a more confidential evaluation and selection procedure.

Additional aspects from the literature review and the lessons learnt in the pilot call

In addition to the survey with the actors involved in the pilot call, the review on the literature revealed further potential for optimising the evaluation criteria and call procedure. However, almost all the literature on the topic of peer reviewing or research evaluation addresses new challenges from a conceptual perspective. These new challenges are paradigmatic changes in science and research and have brought new complexities and uncertainties degrees in the field of research evaluation.

Further work should particularly focus in refining criteria, giving them clearer definitions and boundaries. It would allow to develop new and more suitable sub-criteria to better balance “scientific quality, or robustness” in one hand and “interdisciplinary and innovation” in the other hand.

We suggest to include additional experts (management experts, research users) in the process to deal with project management concern, and at least OF association representatives to participate in the evaluation of future calls.

We suggest to “invent” mechanisms to be implemented in order to allow the funding of a few “risky” research projects, to facilitate new incomers to enter the arena and to promote curiosity-issue research projects. Other assessment steps could be implemented for the projects identified as more innovative and scientifically less robust. This may include the tutorial of the ongoing research projects and results evaluation.

7 References

ALROE, F.H. and S.E. KRISTENSEN (2002) Towards a systemic research methodology in agriculture: Rethinking the role of values in science. Agriculture and Human Values 19: 3-23

ATKINSON, M. (1994) Regulation of science by “Peer review”. Studies in History and Philosophy of Science. 25 (2): 147-158

BLACKSTOCK, K.L, G.J KELLY and B.L. HORSEY (2007) Developing and applying framework to evaluate participatory research for sustainability. Ecological Economics 60: 726-742

CALLON, M. (1998) The laws of the market. Oxford, UK: Blackwell.

CALVERT, J. (2001) Goodbye blue skies? The concept of basic research and its role in a changing funding environment, in SPRU, 2001, University of Sussex, Brighton, also at /www.centres.ex.ac.uk/egenis/staff/calvert/publications.phpS.

CLARK, W.C., CRUTZEN P.J and SCHELLNHUBER. H-J. (2005) Science for Global Sustainability: Toward a New Paradigm. CID Working Paper, vol. 120. Harvard University, Cambridge, MA.

ETZKOWITZ, H. and L. LEYDESDORFF (2000) The dynamics of innovation: From national systems and “mode 2” to a triple helix of university-industry-governments relations. Research Policy 29: 109-123.

FRANCIS, C. and J. KING (1997). Impact of personal values on agricultural research. Society and Natural Resources 10: 273–282.

FREDERIKSEN, L.F., F. HANSSON and S.B. WENNEBERG (2003) The Agora and the role of research evaluation Evaluation 9(2): 149-172.

FUNTOWICK, S. and J. RAVETZ (2001) Post-normal science. Science and Governance under conditions of complexity, in: M. DECKER (Ed.), Interdisciplinarity in Technology Assessment, Springer, Berlin, 2001, pp. 15–24.

GEBER, U., M. KIENEGGER and A. SILMBROD (2007) CORE Organic Final Report – evaluation of the pilot call, deliverable 7.3

GIBBONS, M. (2000) Mode 2 society and the emergence of context sensitive science. Science and Public Policy 27: 159- 163

HACKETT, E.J., and D.E. CHUBIN (2003) Peer Review for the 21st Century: Applications to Education Research Prepared for a National Research Council Workshop Washington D.C. February 25, 2003. Accessible at www7.nationalacademies.org/core/HacketChubin_peer_review_paper.pdf

HARNAD, S. (1985) Relational Disagreement in Peer Review 10: 55- 62.

HEMLIN, S. and S.B.RASMUSSEN (2006) The shift in Academic Quality Control. Science, Technology & Human Values 31: 173- 198.

HOJAT, M. and J. S GONNELLA (2003) Impartial Judgment by the “Gatekeepers” of Science:Fallibility and Accountability in the Peer Review Process. Advances in Health Sciences Education 8: 75-96

HABERMAS, J. (1971) Towards a Rational Society. Student Process, Science and Politics, Beacon, Boston.

KALBERER, T.J. Jr. (1985) Peer review and the Consensus Development Process. Science, Technology& Human Values 10: 63-72.

KLEIN, T.J. (2006) Afterword: the emergent literature on interdisciplinary and trandisciplinary research evaluation. Research Evaluation 15(1): 75-80.

KRISTENSEN, E.S. and N. HALBERG (1997) A systems approach for assessing sustainability in livestock farms in J.T. Sorensen (ed.), Livestock Farming Systems. More than Food Production. Proceedings of the 4th International Symposium on Livestock Farming, EAAP Publication n°89: 16-30. Wageningen, The Nederlands: Wageningen Press.

LANGFELDT, L. (2001) The Decision- Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. Social Studies of Science 31: 820-841.

LANGFELDT, L. (2006) The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments." Research Evaluation 15(1): 31-41.

LAUDEL, G. (2006) Conclave in the Tower of Babel: how peers review interdisciplinary research proposals. Research Evaluation 15(1): 57-68.

LAUDEL, G. and G. ORIGGI (2006) Introduction to a special issue on the assessment of interdisciplinary research. Research Evaluation 15(1): 2-4.

LUKS, F. and B. SIEBENHÜNER (2007) Transdisciplinarity for social learning? The contribution of the German socio-ecological research initiative to sustainability governance. Ecological Economics 63: 418-426.

MANSILLA, B.V. (2006) Assessing expert interdisciplinary work at the frontier: an empirical exploration. Research Evaluation 15(1): 17-29.

NGWENYAMA, K.O., N. BRYSON and A. MOBOLURIN (1996) Supporting facilitation in group support systems: Techniques for analyzing consensus relevant data." Decision Support Systems 16: 155- 168.

NOWOTNY, H., M. GIBBONS and P. SCOTT (2001). Re-thinking science. Knowledge and the public in an Age of Uncertainty. Polity Press, Oxford.

PACKHAM, R. and N. SRISKANDARAJAH (2005). Systemic action research for postgraduate education in agriculture and rural development. Systems Research and Behavioral Science 22(2): 119–130.

PORTER, L.A. and F .A. ROSSINI (1985) Peer Review of Interdisciplinary Research Proposals. Science, Technology& Human Values 10: 33-38.

PRESS S.J., M.W. ALI and C-F.E. YANG (1979) An empirical study of a new method for forming group judgments: Qualitative Controlled Feedback Technological Forecasting and Social Change 15 (3):171-189.

RASMUSSEN, J., V. LANGER AND H.F. ALROE (2006) Bias in peer review of organic farming grant applications. Agriculture and Human Values 23: 181-188.

SAATY, T. L. (1977) A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15: 234-81.

SAATY, T. L. and J.S. SHANG (2005) Group decision-making: Head- count versus intensity of preference Socio-Economic Planning Sciences 41: 22-37.

SALTER, A.J. and B.R. MARTIN (2001) The economic benefits of publicly funded basic research: a critical review. Research Policy, 30 (3): 509-532.

SCOTT, A. (2007) Peer review and the relevance of science. Futures 39: 827-845.

TRAVIS, G.D.L. and H.M. COLLINS (1991) New Light on Old Boys: Cognitive and Institutional Particularism in Peer Review System. Science, Technology& Human Values 16(3): 322-41.

VAN DEN HOVE, S. (2007) A Rationale for Science-Policy Interfaces. Futures 39: 807-826.

WYNEN, E. (1996) Research implication of a paradigm shift in agriculture: the case of organic farming. Resource and Environmental Studies No. 12; Centre for Resource and Environmental Studies of the Australian National University. Archived at http://orgprints.org/00003053

ZIMAN, J. (2000) Real Science. What it is and what it means. Cambridge University Press.