

## Workshop INNO-Appraisal



**September 24 and 25 2009**

**Venue: Room 5A, Breydel Building, Brussels**

### **SUMMARY NOTES**

#### **Purpose**

The **purpose of the workshop** was to discuss with around 20 SELECTED participants

- the findings of the statistical analysis
- the in-depth analysis of selected evaluation issues in RTDI policy.

#### **Participation**

The turnout as for participants was impressive and exactly in line with the planning and expectations. The workshop attracted 19 external participants from 9 different countries (Austria, Belgium, France, Germany, Greece, Italy, Spain, Turkey, UK), mainly from ministries and innovation policy agencies, as well as representatives from EU DG Research and the OECD. Together with the project team (7) and the EU official Alberto Licciardello, the total number was 26 participants. In the build up to the workshop, a number of additional persons who had expressed an interest in attending had to be rejected as the workshop format had to be limited in order to facilitate intensive debate. (See appendix for the list of participants). Four of the 19 external participants were members of the InnoAppraisal steering committee.

#### **Agenda – presentations and discussion**

The agenda was designed in order to achieve the two aims stated above. In each session, a project member presented key aspects and findings, and one participant acted as discussant. Annex 2 contains the full agenda. A short summary is given below while a separate annex containing all presentations has been sent to the Commission.

The first day was devoted to presentation and discussion of the main findings across the whole data set. A first presentation by Jakob Edler and Abdullah Gok focused on the methodology. This involved a live demonstration of the repository on the web. This repository has proved to be popular (as evidenced by hits on the web) and is a major deliverable of the project, supporting the policy intelligence requirements of PROINNO very broadly. In the course of its design, a dialogue with the Trendchart coordinators has led to some improvements in the ways Trendchart deals with the evaluation dimension.

A second presentation by Jakob Edler provided an overview of evaluation practice in Europe, covering the descriptive statistics of the overall data set of evaluations. It was stressed that, for technical reasons, three countries are still missing from the analysis: Sweden, Finland, Netherlands. Ullrike Blankenfeld (BMW, Germany) commented on this presentation. She was supportive of the overall idea to have an overview of evaluation practice in Europe and concentrated on the issue of usefulness. While this issue was to be dealt with the next day, Ms. Blankenfeld already highlighted the importance of context and timing for the usefulness of evaluation reports and exercises.

A second presentation by Michael Dinges and Martin Berger summarised the main hypotheses and correlations across the whole data set. This was intended to convey an idea as to how specific aspects of the evaluation exercise link to each other, i.e. which properties co-exist and how do the properties of the evaluation (use of methods, ToR etc.) impinge upon quality and usefulness. This presentation was a broad *tour d'horizon*, led to an intensive debate and set the scene for the in-depth contribution of day two. The commentator, Alasdair Reid (Technopolis Group, Belgium), stressed the usefulness of the approach, but urged for a clearer linking of evaluation properties to policy types. The aim should be to find out: what works under what circumstances for what kinds of policy measures. He suggested to make better use of the policy categorisation in the Trendchart, which, as the team confirmed, has a number of constraints. It was agreed to categorise the measures in a separate working step (see below). Reid also was interested in the issue of the development of innovation evaluation culture over time: however, the project is not designed to analyse the time development in greater detail at aggregate level. Nevertheless, this will be considered for some country cases that have been done in earlier phases or which will be done towards the end of the study. Furthermore, Reid would like to see what is special with innovation policy compared to other policies. Although this is a very important aspect, unfortunately, the project does not and cannot consider comparisons to other policy fields.

The second day was devoted to three in-depth issues. It was kicked off by a presentation of Stephanie Daimer and Susanne Bühner on impact assessment. Daimer/Bühner presented the main use of impact assessment and how – and under what circumstances – usefulness can be improved. In addition to the statistical data of the database, a set of interviews deepened the understanding of the context conditions and the intentions of policy makers when they commission, conduct and use evaluation in the policy cycle. Further, there is no clear link between impact assessment on the one hand and usefulness on the other (see below, Cunningham presentation). The commentary by Frederique Sachwald stressed the importance of impact assessment and was very positive about the INNOAppraisal function to actually make policy makers more aware of the meaning of methods they use. She gave some insights into their own requirements in the French ministry. At the same time she acknowledged that the expectations of impact assessment has to be modest as regards evidence and attribution. Comments were also given on the need to link the analysis of impact to the type of policy measures. The subsequent discussion stressed the importance of clear definitions and the timing of impact assessments (i.e. ex ante vs. ex post).

The second presentation focused on behavioural additionality (BA). Abdullah Gok and Jakob Edler combined new theoretical thinking with the data on the use of BA across the INNOAppraisal database. The project is lucky enough to mobilise the PHD work by Abdullah Gok on BA. For the first time, this presentation looked systematically at how BA is used, how it links to other evaluation properties and what shortcomings the concept and, moreover, the applications of the concept have. In addition to the analysis of the statistics of the INNOAppraisal, Gok/Edler performed a text analysis of a set of evaluation reports that

claimed to apply BA. Interestingly, the concept is ill defined, applied in very different ways and often not used to its full potential. The major argument is to improve the conceptual underpinning so that the black box of behaviour can be opened and the meaning of policy as a means of influencing behaviour of firms can be better understood. In his comment, Mario Cervantes from the OECD welcomed this deepening of the BA concept, which links to the BA activities of the OECD carried out a couple of years ago. Cervantes stressed the need to have a taxonomy of behaviour to better understand it and the potential of BA to link evolutionary and neo-classical approaches. He also stressed that behaviour as such is not the final aim of policy, rather it is the outcome in terms of innovation, the economy or societal well-being. Thus the linking of behaviour to these final outcomes must be carefully conceptualised. The subsequent discussion demonstrated how important BA is and that for BA, in contrast to most other aspects of INNOAppraisal, some kind of guideline for analysts and policy makers could be defined.

Paul Cunningham then discussed, in-depth, one dimension that had underlay the whole workshop, i.e. what determines the usefulness of evaluations? He set the scene by explaining the dimensions of usefulness and some preconditions for its optimisation. He then presented the database analysis in depth linking, in an exploratory fashion, most of the variables with some exploratory composite indicators for usefulness. While there are few very strong messages, a set of important findings nevertheless emerged out of this exploratory exercise, such as the fact that summative evaluations are regarded as being more useful than formative, planned evaluation being *less* useful than non-planned and that quality itself is not a clear determinant of usefulness (pointing to the meaning of contexts and time windows). He finished by suggesting the need for better indicators of usefulness (as INNOAppraisal data is based on self assessment of policy makers) and a need for a more qualitative analysis on top of the statistics. Rupert Pichler was exceptionally well placed to act as discussant, as he chairs the evaluation platform in Austria and has as policy maker commissioned many evaluation studies. From his experience, the correlations that Cunningham presented are persuasive, the analysis makes sense. But Pichler stressed the importance of the definition of usefulness (for what?, for whom?) and the very purpose of the evaluation in the eyes of the policy makers. For Pichler, the most important determinant is time, the window of opportunity for an evaluation to have an impact on policy making. As other speakers before, he also stressed the importance of linking the analysis with types of policy measures: ideally, he would like to see a typology of linkages between a limited set of programme types and evaluation types.

In a last presentation, Ionna Garefi presented the result of the analysis done together with Effie Amanatidou on the relationship between evaluation needs and practices in the context of structural funds (SF) and the evaluation practice and culture of a set of selected countries (Malta, Poland, Greece). Garefi/Amanatidou presented the requirements that are connected to structural funds and how those requirements are met in the three countries. They showed that, despite specific requirements, the evaluations done in the context of SF do not differ significantly from other evaluations. However, despite the requirements they are lower in quality standards. Still, the SF requirements for evaluation have made a big difference in the countries in terms of the build up of evaluation culture. In the future, SF evaluation should be more flexible and better concrete guidelines for implementing should be given. Nikos Konstantopoulos in his comment welcomed the distinction between SF and non-SF evaluations and confirmed that despite the lower level of quality, the positive effects of the SF requirements in terms of building up awareness and capacity are overwhelming. He stressed the need for a balance between rigorous uniform regulation on the one hand and flexibility for the idiosyncracies on the other hand.

## Wrap up and conclusions

In a last session, Ken Guy and Jakob Edler wrapped up.

INNOAppraisal is an important project. It raises awareness about evaluation practice, its effects and shortcomings in European countries in the area of innovation policy. It delivers a service to the community in terms of the repository, the overall analysis and the in-depth issue analysis. This is a major outcome that we should not underestimate. In addition, the project already links with other activities in the field, mainly the OMC net on evaluation and the OECD activities on Behavioural Additionality, next to a set of national links to evaluation communities (e.g. Austria, Germany). There are indications that the project has functioned as a catalyst and crystallisation point for evaluation discourse in innovation policy in Europe.

It is of great importance to transmit the findings of the project to the next generation. While many of the findings are not spectacularly new, they still make their points on the basis of a new and broad analysis and can be contextualised in a new setting. In so doing, they are valuable as the institutionalised memory of the evaluation community.

On that basis, a strong suggestion was made to continue the work of INNOAppraisal in some form (e.g. within the Trendchart agenda) in order to keep up a repository of evaluations and the opportunities for further analysis along the lines of INNOAppraisal). The extra costs entailed, now the repository is set up, are small, and existing correspondents networks could be used. In addition, INNOAppraisal and potential future activity should link up globally: an evaluation seminar organised by China in the context of an OECD activity would be one obvious opportunity.

The workshop was well-timed, delivering a presentation of the technical approach and its outcomes. This allowed the experts present to comment on the technicalities and to understand the basis of the underlying data and analysis. In the next step, the process would need to be presented the other way round, i.e. start with the major findings and have the technical approaches as back up information.

The set of findings are not only new, but highly interesting and shed some interesting light on the conditions for usefulness and impact of evaluation.

*The following consequences for the project can be defined*

The database needs completion in terms of the three missing countries NL, Finland, Sweden. Once this is done, the major statistical analysis can be finalised

There is a need for linking evaluation practice with the characteristics of policy measures. The Trendchart database is not best suited for this linkage and the team intends to apply its own typology to the various policy measures. On that basis, a necessarily simplistic typology of evaluations would need to be developed to get a clearer picture of what kinds of impact / usefulness are linked to different types of evaluation.

A typology of countries is desirable, with a restricted set of criteria, mainly focusing on the depth and breadth of innovation policy and how this is matched with evaluation practice.

The project does not cover the regional level. The complexities of the national level have so far highlighted that to such an analysis would overburden the approach. In addition, as the

remit and size of regions is so varied,(with a region like Baden-Wuerttemberg being many times bigger and having much more innovation policy measures than most small countries in the EU), the added value of this approach is questionable.

The issue areas that have been analysed appear to be relevant: no additional issues were raised by the participants even when asked in the final round of wrap up. For the individual issues, however, a second statistical analysis is needed once the three missing countries are included, and a more qualitative analysis will be delivered.

The overall analysis and assessment should take the context and characteristics of policy makers into greater consideration - the usefulness of evaluations cannot be understood without this context knowledge. For example, the real purpose of an evaluation can be legitimisation only (accountability only) or it can be learning only. Both are legitimate and justified, but mean very different things for evaluation practice. If this is true, the link of usefulness to the quality of evaluation is not straightforward, as overall quality might be poor, but if those aspects that are needed at a certain point in time are suitable, the evaluation might still be highly useful. It is not possible to control for those aspects in the quantitative analysis but the issue analyses will further deepen this dimension.

## Annex 1: List of participants (steering committee members in bold)

Name	Organisation	Email
<b>Nick Constantopoulos</b>	General Secretariat for Research and Technology, Greece	<a href="mailto:nkon@gsrt.gr">nkon@gsrt.gr</a>
Alasdair Reid	Technopolis	<a href="mailto:alasdair.reid@technopolis-group.com">alasdair.reid@technopolis-group.com</a>
Bart Kamp	INNOVA Europe	<a href="mailto:b.kamp@innova-europe.eu">b.kamp@innova-europe.eu</a>
Peter Teirlinck	BELSP0	<a href="mailto:Peter.TEIRLINCK@belspo.be">Peter.TEIRLINCK@belspo.be</a>
Paul Schreurs	IWT, Flanders	<a href="mailto:ps@iwt.be">ps@iwt.be</a>
Vincent Lepage	SPW/Direction de la Politique économique	<a href="mailto:v.lepage@mrw.wallonie.be">v.lepage@mrw.wallonie.be</a>
Julie Raffailac	CLORA	<a href="mailto:raffailac@clora.eu">raffailac@clora.eu</a>
Miriam Ruiz Yaniz	Technopolis	<a href="mailto:miriam.ruiz.yaniz@technopolis-group.com">miriam.ruiz.yaniz@technopolis-group.com</a>
<b>Ullrike Blankenfeld</b>	BMW, Germany	<a href="mailto:ullrike.blankenfeld@bmw.bund.de">ullrike.blankenfeld@bmw.bund.de</a>
Sonja Kind	VDI/VDE Innovation + Technik GmbH	<a href="mailto:Kind@vdivde-it.de">Kind@vdivde-it.de</a>
Evelina Santa	Spanish Ministry	<a href="mailto:evelina.santa@micinn.es">evelina.santa@micinn.es</a>
<b>Jari Romanien</b>	TEKES	<a href="mailto:Jari.Romanainen@tekes.fi">Jari.Romanainen@tekes.fi</a>
Frédérique Sachwald	Ministère de l'Enseignement Supérieur et de la Recherche	<a href="mailto:frederique.sachwald@recherche.gouv.fr">frederique.sachwald@recherche.gouv.fr</a>
<b>Rupert Pichler</b>	BMVIT, Austria	<a href="mailto:rupert.pichler@bmvit.gv.at">rupert.pichler@bmvit.gv.at</a>
Neville Reeve	EU Commission	<a href="mailto:Neville.REEVE@ec.europa.eu">Neville.REEVE@ec.europa.eu</a>
Mario Cervantes	OECD	<a href="mailto:Mario.CERVANTES@oecd.org">Mario.CERVANTES@oecd.org</a>
Silvia Grandi	Italian Institute for Industrial Promotion	<a href="mailto:GRANDI@ipi.it">GRANDI@ipi.it</a>
Deniz Bayhan	TTGV	<a href="mailto:dbayhan@ttgv.org.tr">dbayhan@ttgv.org.tr</a>
Margaret Dennis	BIS	<a href="mailto:Margaret.Dennis@bis.gsi.gov.uk">Margaret.Dennis@bis.gsi.gov.uk</a>
Thorvald Finnbjörnsson	RANNIS	<a href="mailto:thorvald@rannis.is">thorvald@rannis.is</a>
Jakob Edler	MloIR	<a href="mailto:Jakob.Edler@mbs.ac.uk">Jakob.Edler@mbs.ac.uk</a>
Paul Cunningham	MloIR	<a href="mailto:Paul.Cunningham@mbs.ac.uk">Paul.Cunningham@mbs.ac.uk</a>
Abduallah Gok	MloIR	<a href="mailto:Abdullah.Gok@postgrad.mbs.ac.uk">Abdullah.Gok@postgrad.mbs.ac.uk</a>
Ken Guy	Wise Guys Ltd.	<a href="mailto:ken.guy@wiseguys.ltd.uk">ken.guy@wiseguys.ltd.uk</a>
Michael Dinges	Joanneum	<a href="mailto:michael.dinges@joanneum.at">michael.dinges@joanneum.at</a>
Martin Berger	Joanneum	<a href="mailto:Martin.Berger@joanneum.at">Martin.Berger@joanneum.at</a>
Stephanie Daimer	Fraunhofer	<a href="mailto:stephanie.daimer@isi.fraunhofer.de">stephanie.daimer@isi.fraunhofer.de</a>
Susanne Bühler	Fraunhofer	<a href="mailto:Susanne.Buehrer@isi.fraunhofer.de">Susanne.Buehrer@isi.fraunhofer.de</a>
Ioanna Garefi	Atlantis	<a href="mailto:garefi@atlantisresearch.gr">garefi@atlantisresearch.gr</a>
Alberto Licciardello	EU Commission	<a href="mailto:alberto.licciardello@ec.europa.eu">alberto.licciardello@ec.europa.eu</a>

## Annex 2: AGENDA



*Thursday, Sep. 24*

<b>1.00 – 1.30</b>	<b>Reception</b>
1.30 – 1.45	Commission: Welcome and Introduction
1.45 – 2.30	The overall idea, methodology and the repository (Edler /Gok)
2.30 – 3.30	Evaluation practices in Europe (Edler) <i>Comment: Ullrike Blankenfeld, BMWi, Germany</i>
3.30 – 4.00	Coffee Break
4.00 – 5.45	Key findings, basic messages, and link (Berger & Dinges) <i>Comment: Alasdair Reid Technopolis Belgium</i>
5.45 – 6.00	Wrap up day one (Edler)



*Friday Sep. 25*

- 9.00 – 9.05            Welcome and introduction into the issues (Edler)
- 9.05 – 10.15        Impact assessment (Bührer, Daimer)  
Comment: *Frédérique Sachwald Ministère de l'Enseignement  
Supérieur et de la Recherche, France*
- 10.15 – 10.30        *Coffee break*
- 10.30 – 11.45        Behavioural additionality (Edler / Gok)  
Comment *Mario Cervantes OECD, Paris*
- 11.45 – 13.00        What determines usefulness (Cunningham)  
Comment: *Rupert Pichler BMVIT, Austria*
- 13.00 – 14.00        *Buffet Lunch*
- 14.00 – 15.00        Structural Fund evaluation (Amanatidou, Garefi):  
Comment: *Nikos Konstantopoulos General Secretariat for Research  
and Technology, Greece*
- 15.00 – 15.30        *Wrap up: Ken Guy*  
*Jari Romainainen, TEKES, Finland tbc,*  
*All participants invited to give final statements*
- 15.30 – 15.45        *Conclusion - Commission*