

Lessons to learn from evaluation systems set up in the context of Structural Funds: impact on evaluation culture and institutional build up

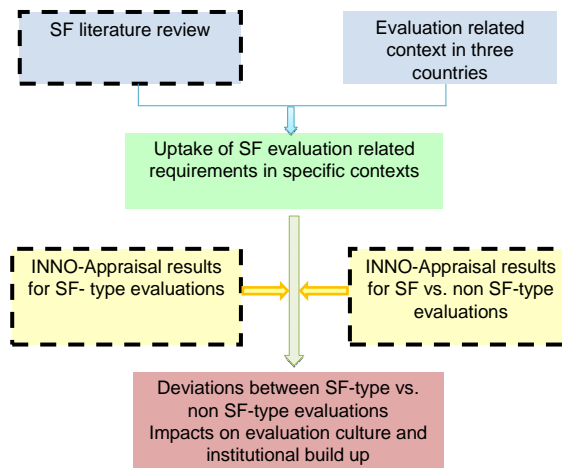
Ioanna Garefi, Effie Amanatidou

garefi@atlantisresearch.gr,
amanatidou@atlantisresearch.gr

Contents

- Methodological Approach
- SF evaluation requirements
- Delivery & practice in SF vs. non SF type evaluations
- SF evaluation requirements in different contexts
 - Feasibility
 - Appropriateness
 - Impacts
- First Conclusions
- Key points to discuss

Methodological Approach



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF evaluation requirements (1/4)



- Evaluation, monitoring and financial control: a prominent position under SF framework
- SF evaluation requirements relate to structures, evaluation design and execution, evaluators' selection, publicity and dissemination of results and quality assurance.



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF evaluation requirements (2/4)



- **Structures:** Authorities for Management, Certification, Auditing and a Follow-up Committee.
- **Evaluation design:** request in the documents for negotiations for description of monitoring & evaluation system, info on appropriations and resources needed and ex-ante evaluation.
- **Evaluation execution:** Community structural assistance is subject to ex-ante, mid-term and ex-post evaluation.
- SF provisions specify the aim, type, and nature of evaluations and provide suggestions on how the different types should be carried out. Reference to existing evaluation guides is made.



SF evaluation requirements (3/4)



- **Evaluators' selection:** explicit mention to independent bodies or experts selected via a competitive tendering process.
- **Publicity & Dissemination:** refers not only to activities supported under SF but also evaluation results; good practice to make public the entire evaluation report.



SF evaluation requirements (4/4)



Quality assurance: suggested standards for:

- **Evaluation report:** Meeting needs; relevant scope; open process; defensible design; reliable data; sound analysis; credible results; impartial conclusions; clear report useful recommendations.
- **Evaluation process:** Coherent obj.; adequate ToR; tender selection; effective dialogue feedback; adequate info; good management; effective dissemination.



Inno Appraisal Workshop, Brussels, 24-25 September 2009

The principal hypothesis...



- The type and content of SF requirements related to evaluation may lead to specific characteristics in the delivery & practice of evaluation under the SF framework in comparison with non SF type evaluations. These may refer to evaluation topics, data analysis & collection methods, type of evaluators, quality characteristics, dissemination.
- SF demands high standards on structures and processes that inevitable need some institutional learning and structure building.



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF vs. non SF type evaluations (1/8)



Evaluation topics: less striking differences than expected across the two groups overall

- More differences when looking at topics per evaluation type within each group, for example:
 - SF / ex-ante: consistency, complementarity, policy / strategy development & relatively more gender and minority issues;
 - SF/ interim: effectiveness, internal consistency, programme implementation efficiency;
 - SF/ ex-post: goal attainment and outputs, outcomes and impacts.
- These differences are clearer in SF group which may be partly attributed to SF guidelines.
- Across the two groups the issues addressed under each evaluation type, e.g. by interim or ex-post evaluations are more or less the same. Difference: non SF focus more on additionality aspects while SF group turns more to programme implementation efficiency.



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF vs. non SF type evaluations (2/8)



Evaluation topics: differences per evaluation type

Evaluation topics	'SF' sample (% of users / total no of answers under each option)				'non SF' sample (% of users / total no of answers under each option)			
	Acc	Ex-ante	Ex-post	Interim	Acc	Ex-ante	Ex-post	Interim
C.7.a. External Consistency	9,68	35,48	9,68	25,81	14,56	6,80	9,71	40,78
C.7.b. Internal Consistency	10,00	33,33	10,00	30,00	16,19	9,52	10,48	40,95
C.7.c. Coherence/Complementarity	9,68	35,48	6,45	16,13	11,22	6,12	9,18	30,61
C.7.d. Goal Attainment/Effectiveness	11,11	3,70	11,11	33,33	17,48	6,80	17,48	49,51
C.7.e. Outputs, Outcomes and Impacts	9,09	18,18	15,15	27,27	18,27	8,65	18,27	48,08
C.7.f. Quality of Outputs	11,54	7,69	11,54	11,54	13,27	3,06	10,20	22,45
C.7.g. Value for Money/Rol/Cost-Benefit	8,33	8,33	8,33	8,33	2,06	0,00	4,12	12,37
C.7.h. Programme Implem. Efficiency	7,41	25,93	22,22	29,63	13,86	4,95	9,90	38,61
C.7.i. Project Implem. / Efficiency	7,69	7,69	7,69	19,23	8,91	0,99	10,89	19,80
C.7.j. Input Additionality	7,41	7,41	3,70	22,22	7,84	3,92	15,69	14,71
C.7.k. Output Additionality	10,34	6,90	10,34	13,79	10,78	3,92	13,73	17,65
C.7.l. Behavioural Additionality	3,33	3,33	10,00	13,33	12,50	3,13	11,46	19,79
C.7.m. Policy/Strategy Development	6,67	26,67	6,67	23,33	15,24	9,52	11,43	40,95
C.7.n. Gender issues	6,67	33,33	0,00	13,33	6,06	2,02	1,01	4,04
C.7.o. Minority issues	0,00	20,00	0,00	0,00	1,04	1,04	1,04	0,00



Inno Appraisal Workshop, Brussels, 24-25 September 2009

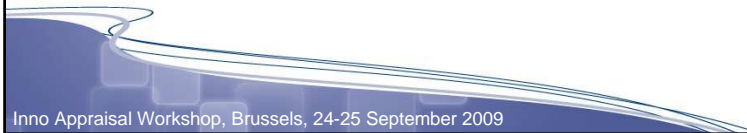
SF vs. non SF type evaluations (3/8)



Data Analysis and collection methods: not surprising differences across the two groups

- **Data analysis:** in general 'descriptive statistics' and 'context analysis' stand out in both SF and non SF groups. 'Document analysis' comes next. 'Case studies' more in non SF group.
- **Data collection:** 'monitoring data' and 'existing surveys' first in both cases. 'Document search' among the first preferences in SF group but superseded by 'interviews' in non SF. 'Participant surveys' a dominant method more in non SF.

Clearer differences across the evaluation types within each group than across the two groups



SF vs. non SF type evaluations (4/8)



- The INNO-Appraisal survey evidence does not strongly support the hypothesis that the relevant SF requirements make a difference in terms of evaluation topics, data analysis and collection methods applied.
- They do make a difference though in guiding the topics to cover and analysis & collection methods to use across the evaluation types (ex-ante, interim, ex-post) within the SF group.



SF vs. non SF type evaluations (5/8)



Quality standards: certain criteria feature the highest in the SF group but interestingly it is the non SF group that generally presents higher scores. Yet, the SF group presents a quality mean >3 in 10 out of 12 criteria.

Quality standards	SF type*		Non SF type	
	Mean	Rank	Mean	Rank
Suitability of methods chosen	4,18	1	4,16	3
Well documented information sources	4,07	2	3,90	6
Address of ToR	4,06	3	4,13	4
Analysis based on data	3,96	4	4,19	2
Design appropriate to objectives	3,78	5	3,90	5
Conclusions based on analysis	3,76	6	4,35	1
Satisfactory application of qualitative methods	3,61	7	3,81	7
Coverage of broader context	3,59	8	3,16	11
Satisfactory application of quantitative methods	3,52	9	3,71	8
Discussion within government circles	3,12	10	3,28	9
Discussion with participants / stakeholders	2,89	11	3,25	10
Usefulness of recommendations	2,47**	12	3,04**	12

(*) Number of responses very low (<20) – should be considered with caution

(**) Average mean of options selected under the specific group of questions (E.2.c – E.2.h)



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF vs. non SF type evaluations (6/8)



Quality standards:

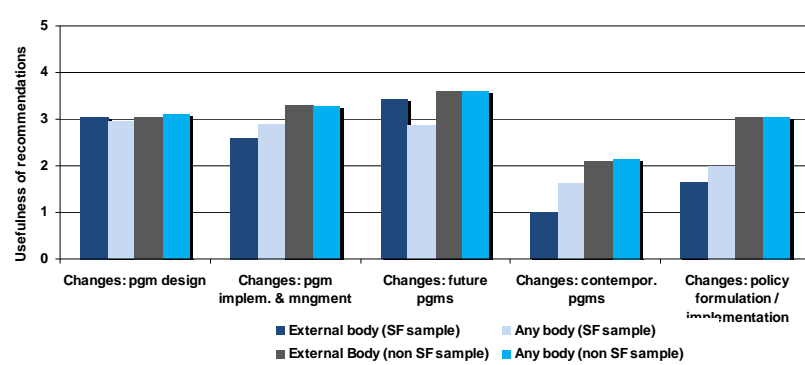
- One might expect that appraisals with high scores in quality may result in high impact in terms of dissemination and usefulness of recommendations. However, the evidence is not clear enough to support this argument in the SF case.
- Picture is clearer in non SF: High quality appraisals seem to lead to more discussion within government & stakeholders and also to more useful recommendations especially in relation to policy formulation and implementation.



Inno Appraisal Workshop, Brussels, 24-25 September 2009

SF vs. non SF type evaluations (7/8)

Use of external evaluators & usefulness of recommendations: no striking differences across SF and non SF groups.



Within the SF group, however, external evaluators do seem to make a difference in the usefulness of recommendations referring to changes in future programmes.

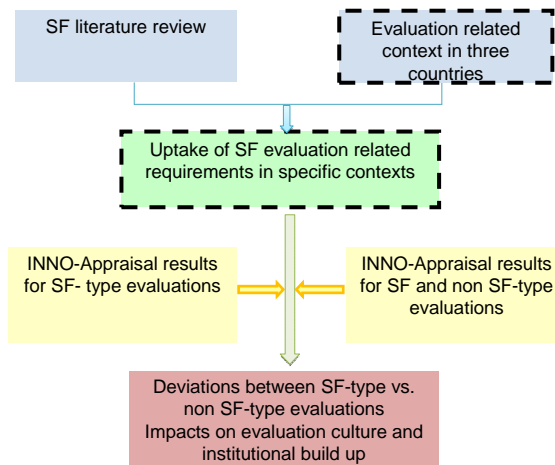
SF vs. non SF type evaluations (8/8)

Publicity & Dissemination of results:

- Results are published by 100% in SF vs. 82% in non SF
- But while non SF may be available slightly less than SF type evaluations they are discussed with relevant audiences more, despite the fact that this is explicitly mentioned as a quality criterion in SF evaluation guides.

The impact of SF evaluation on evaluation culture and practice – exploratory country analysis –

Greece, Poland, Malta



Innovation system in Greece



- Public sector is the main performer of innovation whereas the business sector is the weakest part of the Greek science system
- Greek businesses are rather followers than innovation leaders and thus more involved in innovation transfer than innovation creation.
- Weak links between business sector and public research sector.
- Lack of focus in the research sector linked to absence of industrial actors
- Incentives to private operators (S&T parks, incubators, technology brokers, venture funds)



Inno Appraisal Workshop, Brussels, 24-25 September 2009

Greek Evaluation System/Culture



- Programme/ project evaluation was first introduced by GSRT in the late 80s but it mainly concerned ex-ante.
- Innovation policy measures are being evaluated following the rules of the Structural Funds
 - **Ex-ante**: implemented on an overall operational programme level
 - **Interim**: focused on results and impact in only a few specific schemes, only two of which aim to promote innovation
 - **Ex-post**: Lack of systematic ex-post evaluation of the implemented programmes & measures
 - Lack of reference to the overall impact of programme's priorities on innovation and innovativeness

Evaluation for Evaluation

Even though the uptake and delivery systems of each scheme are well investigated, there is a lack of impact assessment and appreciation of each scheme's qualitative achievements



Inno Appraisal Workshop, Brussels, 24-25 September 2009

Interview Results so far... (still in progress)



Feasibility?

- System too fragmented: certification authority is set up whereas a management, auditing authority and follow up committee are less active.
- Public administration faces problems while trying to be consistent with SF requirements -> low quality, lack of know-how and fading out of transfer of knowledge.

Appropriateness?

- The first thing that applies are the values and idiosyncrasy of the country, whilst SF guidelines come second.
- Lack of appropriate and qualified people.

Impacts?

- Production of better indicators but not creation of clear policy lines.
- SF guidelines can assist in a certain way as reference tools.
- Impacts are forthcoming and positive but not always identifiable and direct.



Innovation System in Poland



- Polish NIS was described as fragmented in the past without adequate coordination mechanisms between key ministries –the situation now is improving gradually.
- Main operators dealing with innovation, science and technology policies are the Ministry of Economy and Ministry of Science and Higher Education.
- Major change in the Polish governance system: PAED – Polish Agency for Enterprise Development.
- Insufficient science- industry cooperation.



Polish Evaluation System/Culture



- Underdeveloped evaluation policy until their accession period
- National Evaluation Units - Well structured evaluation system (6 different evaluation units for the different OPs)
- Resources for evaluation have increased over the years both in personnel and financially.
- Important evaluation capacity but still a challenge
- Two major problems: insufficient budget and timing.

Inno Appraisal Workshop, Brussels, 24-25 September 2009



Interview Results so far... (still in progress)



Feasibility?

- Need for more evaluation guidelines (without though being obligatory) and experts who give recommendations on how to implement them.
- Monitoring system requires improvements so that evaluations could be conducted on a more efficient basis.

Appropriateness?

- Compliance with SF requirements with only minor adjustments.

Impacts?

- There are some spill over effects from implementing the SF regulations which are difficult to measure.
- Some of the evaluations are executed internally as a capacity building exercise.

Inno Appraisal Workshop, Brussels, 24-25 September 2009



Innovation System in Malta



- Main operators dealing with research and innovation initiatives are:
 - Malta Council for Science and Technology (MCST)
 - Malta Enterprise (ME)
 - University of Malta
- A better policy coordination was achieved through the improved and tight linkages of these public agencies.
- The late restructuring of the governance system (2008) had an impact on the innovation and research structure
- A cluster of multinational firms is the main contributor to the national expenditure on R&D.
- Lack of a stand-alone innovation strategy showing significant overlaps in innovation and research policy measures



Inno Appraisal Workshop, Brussels, 24-25 September 2009

Maltese Evaluation System/Culture



- No mid-term evaluation for the structural funds programme 2004-2006 is available as this was a short programme but only ex-ante.
- A relatively weak if not infant evaluation culture.
- Formal reviews and evaluations are undertaken occasionally attributing to a limited or inexistent assessment of the effectiveness of policy measures
- Certain evaluation reports are not published but are kept internally to the respective ministries



Inno Appraisal Workshop, Brussels, 24-25 September 2009

Interview Results so far... (still in progress)



Feasibility?

- Need for more flexibility in the SF requirements in order for changes in the local innovation system to be achieved.
- Evaluations should be carried out according to criteria defined on a country basis and on the type of project.

Appropriateness?

- Relevance and fine tuning to the local context is taken into careful account during ex-ante evaluations.

Impacts?

- The uptake of SF requirements introduces a certain financial rigour in policy design and gives greater attention to the monitoring of indicators and profiling of participants.
- Need for more transparent evaluation procedures.
- SF requirements may lead to an improved evaluation system.
- Requirements are important for drawing sound lessons and ensuring more relevant and effective policy-making.

Inno Appraisal Workshop, Brussels, 24-25 September 2009



Inno Appraisal Workshop, Brussels, 24-25 September 2009



First Conclusions (1/2)



Do SF regulations make a difference?

- Albeit some exceptions, evidence does not in general suggest that SF requirements lead to specific characteristics in the delivery & practice of evaluation. They do make a difference in guiding the execution of the different evaluation types under the SF framework.
- Despite the quality standards promoted by SF provisions it is the non SF type evaluations that present higher quality scores.
- It is high quality non SF (than SF) evaluations that may be discussed more and lead to more useful recommendations.

Are they appropriate for the context's studied?

- The up-take of SF evaluation requirements faces challenges ranging from particular idiosyncrasies and fragmentation of actors in the national innovation system, to lack of internal capacity or inability to form clear policy lines.

Inno Appraisal Workshop, Brussels, 24-25 September 2009



First Conclusions (2/2)



Are they feasible?

- SF evaluation requirements need more flexibility to take into consideration country and programmes' specificities

What were the impacts from their implementation so far?

- Impacts from implementing SF evaluation requirements are slow and difficult to identify.
- SF evaluation requirements have introduced though a certain rigour in policy design improving the quality of the programme and related policy making.
- SF guidelines can assist as reference tools.
- They can also help towards more transparent evaluation procedures and more effective policy-making.
- But they are not enough for establishing a real 'evaluation culture'.

Inno Appraisal Workshop, Brussels, 24-25 September 2009



Key points to discuss



- Evidence does not suggest that SF evaluation requirements make a difference in evaluation quality but it does show that the quality of the SF type evaluations is not that low.
- If the context is also taken into account then the conclusion becomes: *SF requirements help produce relatively good quality evaluations in contexts of low evaluation traditions and evaluation systems facing several challenges and still in infant phases of development.* So, is it that SF regulations do make a difference in the end?
- What are other impacts on culture, inst. build up, learning?
- How else can SF regulations help improve evaluation systems and nurture an 'evaluation culture'? What else is needed?
- What needs caution in the implementation of SF regulations?
- Recent changes in regulations leave SF evaluations more at the hands of Member States. How will this affect the implementation and impacts of SF requirements so far?
-(any other?).....



Inno Appraisal Workshop, Brussels, 24-25 September 2009



Thank you!



Inno Appraisal Workshop, Brussels, 24-25 September 2009