

Evaluations in H2020

Peer review or poor
review?

Cecilie Ydemann Hansen
Torben Høøck Hansen

UNIVERSITY OF COPENHAGEN



The trigger... too many 'events' like this

Proposal addressed this scope:

A. [2018] Small organisms, big effects for plants-
Belowground biodiversity* interaction with plants
(RIA)

Proposals will lay the ground for better understanding and applying the benefits of *soil* organisms... Activities will explore the processes and interactions between plants and the different plant and *soil* micro and macro biota. Work will expand knowledge of the impacts of land management on *soil* Findings on the beneficial effects of functional *soil* biodiversity for crop production will feed into the development of strategies and tools for sustainable *plant/soil* management..

*** OED def.: A. adv. Below the surface of the ground; underground**



Photo of a "soil scientist"
source: Wikipedia

What the evaluators wrote...

"The state of the art is well discussed, and main knowledge and methodological gaps are very clearly described. The way the holistic approach is envisaged is really a novel concept, *although the limits of the studied system do not include above-ground biodiversity as a whole, which is a shortcoming.*"

Fun fact: This is what they wrote earlier in the ESR:

"They [the objectives] *show high pertinence with the scope* of the topic, since the project will tackle soil biodiversity in different crop cultivars..

So we had to get some real insight

Our aims:

- The H2020 evaluation system as seen from the evaluators' perspective
- To give Research Support Officers an understanding of the serendipitous nature of proposal evaluations

Why?

- Examples where the evaluations left us wondering what actually happens during the review process

Background

- 114 replies to an online survey among all Danish H2020 evaluators active in the 2014-15 calls
- 27 in-depth interviews with evaluators, following a semi-structured guide
- Report on the findings of the survey and interview

<https://ufm.dk/en/publications/2018/evaluations-and-evaluators-in-horizon-2020-report-on-an-analysis-among-danish-evaluators>

Results – some confirmed ‘old truths’

- **Verbosity, muddled text** = 46 % say it has “critical” or “significant” influence on the evaluations

Or as an expert said during the interview:

*“People think its stories, but it’s **ALL** true – unclear language, use of platitudes, muddled meanings etc., it **ALL** influences the score even though it will never be written down in the ESR”*

Experts are in a hurry = 76 % spend 4 hours or less reading your proposal

.

Less time, less understanding?

- **Interesting fact:** Less time spent reading = less forgiving of verbosity, spelling mistakes etc.

Influence ->	None	Minor	Some	Significant	Critical
Time spend					
Up to 2 hours	0%	13%	35%	35%	17%
2-4 hours	0%	18%	28%	45%	10%
4+ hours	4%	22%	44%	26%	4%

RIA/IA/CSA/ITN evaluators' replies: *Time spend reading / influence of "verbose and/or hard to understand sentences"*

Results – ‘new things’ in H2020

Gender = 51 % state they were briefed on gender related issues
21 % state that gender was not discussed during evaluation
60 % find gender unimportant when it comes to gender distribution of PIs or research leaders in a consortium

As one evaluator said:

“Not something I pay much attention to; it is more for academia”

Gender continued

Interesting fact

Influence ->	Gender of PI not important	Gender of PI can be important	Gender of PI is important
Female	39%	35%	26%
Male	68%	22%	10%

Interesting fact no. 2 = 71 % of survey respondents were men

Q: "In your opinion, how important is gender when it comes to the composition of the group of primary investigators/research leaders in a given proposal?"

Results – ‘new things’ in H2020

Impact = 49% replied they were “somewhat confident” or “uncertain” when scoring impact

Interesting facts

- Almost all interviewed experts found the focus on impact relevant
- Many felt proposals ‘overplayed their hand’
- Different views on how concretely impact should be described:

“[I want] concrete business plan, KPI... more convincing that a lot of promises”

VS

“Quantified impact is very hyped. All talk about indicators, even when it does not always make sense”

When reviewers look for different things under "Impact"

"It's just something that's taken from looking into a crystal ball"

VS

"A concrete business plan, a set of Key Performance Indicators or similar are still more convincing than a lot of promises"

Future perspectives

Some issues based on the interviews and the survey, and the concerns raised by the evaluators

Logistics

What happens to the review process when the programme gets larger?

- None of the evaluators of collaborative projects found that remote evaluation was recommendable
- Some felt that it was a problem with many new and inexperienced evaluators
- Many expressed that already now, the time was insufficient to ensure the quality of evaluations

Interdisciplinary research

- On one hand, quite a positive attitude
- On the other, deemed to be hard to evaluate
- Sometime a lack of coverage of expert fields
- Takes experience to evaluate this type of projects– at odds with many new and inexperienced evaluators

Suggestions to the EU on how the process can be improved?

- Clearer guidance (to make evaluations more efficient and sound)
- Better training of new evaluators
- Some proposal elements could be standardised, e.g. some IPR issues or mandatory dissemination activities
- Short(er) proposals

Reflections on the process and the results

- All those interviewed showed and expressed real dedication to the role as evaluator
- Impressive response rate
- Interviewees were all eager to talk with us, with many giving us far more time than planned
- We got confirmation of some old 'truths' about proposal writing, but were also surprised (e.g. the contradiction of views about how to deal with impact, the negligent role RRI played)

Reflections on the process and the results

- Worrying that 28% answered that the competences present in expert groups only partly meet the needs for a proper evaluation
- Are the mainstreamed issues not remembered by evaluators or are they absent from the briefings?
- Interpretation of results is difficult
- Both the numerical data (limited numbers) and the interviews gives insights, but must be used with care

Thank you for your time