Evaluations and evaluators in Horizon 2020

Report on an analysis among Danish evaluators

June 2018





Danish Agency for Science and Higher Education





Published by Danish Agency for Science and Higher Education Bredgade 40 1260 Copenhagen K Denmark Phone: +45 3544 6200 sfu@ufm.dk www.ufm.dk/en

Publication can be downloaded at ufm.dk/en/publications

ISBN: 978-87-93706-15-6 (UK-print) ISBN (electronic publication): 978-87-93468-94-8 (UK-web)

Evaluations and evaluators in Horizon 2020

Report on an analysis among Danish evaluators

June 2018

Contents

1. Introduction	7
2. Survey	9
2.1 Basic statistics – background, gender, age and areas of expertise	9
2.2 Distribution between H2O2O instruments evaluated	11
2.3 Level of experience	12
2.4 General questions (e.g. reading medium, average time spend on a	
proposal etc.)	12
2.4.1 How much time do evaluators spend reading each proposal?	14
2.5 Questions relating to the composition of expert groups and panels	14
2.5.1 Main problems of group composition	15
2.6 Questions relating to evaluation of impact	16
2.6.1 What type of impact?	17
2.6.2 What is the probability of the impact actually happening?	17
2.7 Questions relating to SME participation, gender and geographical	
spread of consortium	17

3. The interviews

3.1 How do evaluators read proposals and what did they say about the	
quality?	21
3.1.1 Project abstracts	22
3.1.2 Layout of the text	22
3.1.3 Graphics & Illustrations	22
3.1.4 Gantt charts	23
3.1.5 Verbosity vs. simplicity in the text	23
3.1.6 A summary of what the evaluators in general like when read	ing
a proposal is worth considering during proposal writing	24
3.2 Evaluators point of view - specific proposal types	25
3.2.1 Proposals for the European Research Council (ERC)	25
3.2.2 The collaborative proposals (RIA, IA and CSA)	27

4. Why become an evaluator?

35

21

4.1 Reasons for signing up as an evaluator	35
4.2 Workload of evaluators	37

4.3 Briefings before the evaluation?	37
4.4 Would you do it again?	38
5. Comments on the H2O2O evaluation procedures	41
About the report	45

The Covent Garden Building, Bruxelles. Designed by architects Montois Partners in association with Art & Build.

1. Introduction

How do Horizon 2020 (H2020) evaluators read proposals? What are the pitfalls that makes H2020 evaluators subtract points? This report sheds on several of those pertinent and often unresolved questions that proposers ask themselves when writing proposals for H2020.

The Danish Agency for Science & Higher Education, the EU Office at University of Copenhagen and the Research Support Office at Aarhus University all have a common and strong interest in providing the best possible support for H2O2O proposers. The three partners, on a daily basis, answer questions on how best to present ideas and how best to convince the deciding evaluators to fund projects. This report is an effort to refine and supplement advice in this field with new insights that are both quantitatively and qualitatively based, and that are, so to speak, coming "directly from the horse's mouth", that is, from the evaluators themselves.

The original idea for this report came from the University of Copenhagen. After reading a number of H2O2O (H2O2O) Evaluation Summary Reports from proposals that University of Copenhagen (UCPH) participated in during the 2014-15 H2O2O calls, it seemed that something had changed from previous framework programmes, things were not 'business as usual'. Some of the remarks in the Evaluation Summary Reports baffled researchers from UCPH and their consortium partners as well as research support staff. Discussions with colleagues in Denmark and in other EU member states led to a wish to better understand the role the evaluators have, and to get a better understanding of the strengths and weaknesses of the H2O2O evaluation process.

This subsequent joint analysis and report hopefully provide insights into the evaluation system, as well as the values and motivations of evaluators, which can be of use to future proposers for H2020.

Thanks is due to the more than 100 Danish H2020 evaluators who took the time to fill out the online survey, and the 33 evaluators who volunteered to give a follow-up interview about their experiences. Without their valuable contributions, given with much openness and generosity, this report would not exist. Following the many positive and constructive reactions to our questions, and after reading all the notes and comments, the authors are left with an overall impression of dedication, skill and honest effort to do a difficult job, sometimes in somewhat strenuous circumstances.

A word of warning: Although 114 evaluators from Denmark participated in the evaluation of the 2014 and 2015 calls covered in this report, many thousand experts have been involved in the evaluation of H2020 proposals so far. Therefore, what is presented are the experiences and insights from a small cohort acting as a critical case. However, if one accepts our claim that the Danish evaluators can be representative of what many H2020 evaluators experience and think, this report can be a useful guiding instrument for future proposers, as well as a starting point for discussions on the H2020 - and perhaps Horizon Europe - evaluation procedures. Marianne Zandersen, Senior researcher, Department of Environmental Science, Aarhus University. Evaluator on "Research and Innovation Action" applications under "Climate Action, Environment, Resource efficiency and raw materials".

2. Survey

The evaluators included are all from the official public list of experts on the Participant Portal¹, and it was possible to find contact information on 215 evaluators based in Denmark who had participated in the 2014 and 2015 calls². The 215 persons were invited to answer an online survey. The survey was subdivided into a general section focusing on experience, habits when evaluating, time spent reading a proposal, cross-cutting issues etc., and an instrument-specific section focusing on the theme and type of projects evaluated. The evaluators had the option to add comments to almost all questions, and many did so. In total 114 replies were received, an overall return rate of 47 %³. At the end of the survey, the evaluators had the option of volunteering to participate in a follow-up interview, 33 did so. The survey is 100 % anonymous without any links between a given individual and this person's answers.

2.1 Basic statistics – background, gender, age and areas of expertise

Figure 1

Employment of evaluators participating in the survey



¹ http://ec.europa.eu/research/participants/portal/desktop/en/funding/reference_docs.html#-

² While it was straightforward to find the evaluators from the universities and other public institutions, it proved far more difficult to find the evaluators from private companies and consultancies. Some companies do not list any contact information of staff, apologies are in order to those evaluators that could not be included due to this.
³Not all 114 replied to all questions in the survey. E.g. when asked about impact, only 86 provided answers.

Figure 2

Gender of evaluators participating in the survey



Figure 3

Age of evaluators participating in the survey



Figure 4

Main area of expertise of evaluators participating in the survey



2.2 Distribution between H2O2O instruments evaluated

Figure 5

Types of proposals evaluated



Note.: Other kinds of multi-partners projects cover collaborative projects in all of H2O2O including FET Open proposals, Infrastructures and Fast Track to Innovation.

2.3 Level of experience

Figure 6

Experience of the evaluators (i.e. number of proposals evaluated so far)



2.4 General questions (e.g. reading medium, average time spend on a proposal etc.)

Figure 7

How much time do you normally spend reading a proposal?



Figure 8

How do you normally read the proposals?



Figure 9

Questions relating to the proposal as text



2.4.1 How much time do evaluators spend reading each proposal?

Many researchers, and most research support staff, have for many years mused over how evaluators read proposals; how long time they spend reading and what they like and do not? When it comes to proposals to the framework programmes, this has often been based not so much on facts, but on titbits of information from colleagues, informal talks with EU evaluators and what insights the European Commission itself has revealed.

As expected, the average time spent reading a proposal varied quite a lot but the most common was between 2 and 4 hours or 4+ hours (total of 70 %), but it is noteworthy that 30 % of the evaluators spent less than 2 hours reading a proposal. Part of the difference is of course due to the variations in the length of proposals (10 pages for a SME instrument proposal vs. a 100-page infrastructure RIA)⁴.

2.5 Questions relating to the composition of expert groups and panels

Figure 10

The composition of the expert groups



⁴ Length of the Part B, sections 1-3. In addition to Part B comes the A forms, description of beneficiaries (section 4) and ethics and security (section 5).

Figure 11

In your opinion, was the composition of competences in the group/panel adequate? (e.g. for evaluating all cross-disciplinary elements of the proposals)



In H2020, the European Commission has strived to increase the share of evaluators from non-university organisations⁵, as well as taking on-board "brand new experts"⁶. In our survey (figure 1), 39 % of the responding evaluators work outside universities, a figure in line with the European Commission's overall aim of having at least 40 % of non-university evaluators. It should be noted that even though the number of non-university experts has risen, some of the comments from our survey state that the academic component is too dominant.

When asked about adequacy of the composition of competences in the expert groups (figure 11), a majority of 68 % found the composition of the groups/panels adequate for the evaluation task. It should be noted that 28 % responded only "Partly" so, and four % said "No". In essence, 32 % of respondents have – to a lesser or higher degree - experienced the composition of evaluation groups as not ideal for the purpose. In the follow-ing, some of the problems that the 32 % have experienced will be outlined, since those replies may represent insights on what may be improved upon.

2.5.1 Main problems of group composition

Groups were sometimes lacking cross-disciplinary expertise, or there was not an equal mix between universities and the private sector. In addition, some evaluators said that there was too little participation from industry. Some among the 32 % expressed that in some cases other evaluators had insufficient knowledge, insufficient academic understanding and/or problematic communication skills, which hindered the effectiveness of the group as a whole. One of the evaluators from this group of respondents stated in the comments that there is sometimes also what could be perceived as a cultural barrier between different regions of Europe, and that this perceived need for geographical

^s Universities have traditionally supplied far the majority of experts in the framework programmes, also in the top-down strategic themes/areas.

^e https://ec.europa.eu/programmes/horizon2020/en/horizon-2020-statistics

diversity often hampered effectiveness. Another evaluator mentioned that the European Commission are so focused on avoiding potential conflicts of interest that the level of competence is neglected when appointing experts for a given evaluation; this to such a degree it is making it impossible to gather high-quality experts, as these tend to collaborate and/or publish together. Finally, an evaluator found that *"many evaluators are unexperienced and narrow-minded"*, and speculated that this might be the result of the experienced evaluators having reached the limit of permitted evaluator working days in H2020.

2.6 Questions relating to evaluation of impact

Figure 12

How confident are you as an evaluator when scoring the expected impact as stated in the proposals?



Throughout H2O2O, many perceive the impact section as both the hardest part to write, and the hardest part to evaluate, mainly because it deals with estimations and predictions of a future beyond the project itself.

It was specifically asked how confident the evaluators were in relation to evaluating the impact sections, to test whether evaluators feel they have the necessary skills. This can also point to the question whether the European Commission has given sufficient guidance to evaluators to evaluate the impact section properly.

Figure 12 shows that more than half of our respondents are "Very confident" when evaluating the presentation of impact. This figure is somewhat reassuring, but on the other hand 44 % profess to being only "Somewhat confident" with five percent being "Uncertain". The last two figures - together representing almost the other half of respondents clearly shows that there is still very basic uneasiness in many evaluators, when it comes to evaluating impact. The comments received for this section revolve around two basic questions: What type of impact? And: What is the probability of the impact actually happening?

2.6.1 What type of impact?

Impact can be on many levels and affect many different and specific target audiences, (i.e. broader societal impact vs. impact on business/economy, or impact on young vs. impact on elderly people). This means that in the evaluation situation, it is imperative for the evaluators to have a clear vision of which type of impact is specifically important for the EU for the topic evaluated, also in order to able to judge and choose between different types of impact for winning proposals.

This points to the importance of carefully crafted and very specific guidance for the evaluators presented by the European Commission before the actual evaluation of the proposals. Our analysis also shows that to a large extent the European Commission already takes on this responsibility, but it could still be an area open for improvement.

2.6.2 What is the probability of the impact actually happening?

Another issue raised in a number of comments is that some evaluators feel uneasy when judging the probability of promised impact actually happening. The recurring theme of the comments is along the lines: "Is this realistic?", "How can I judge if this is realistic?" as comments on the credibility of the trajectories put forward by the proposers. The onus is on the proposer to efficiently put the evaluators' minds at ease by providing bullet-proof and referenced documentation, and in addition not to promise too much, thereby losing credibility.

2.7 Questions relating to SME participation, gender and geographical spread of consortium



In your opinion, to what extent did SME participation influence the scoring of the proposals?



Figure 14

To what extent was gender an issue in the evaluations you took part in?



Figure 15

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



Figure 16

In your opinion, to what extent did the geographical spread of the consortium influence the scoring of the proposals?



Evaluation meeting in Covent Garden Building, Bruxelles.

Photographer: David Beaugnet, European Research Council Executive Agency (ERCEA)

3. The interviews

All 33 volunteers were contacted to arrange interviews, and 27 interviews were conducted. The interviews were carried out following an open interview guide divided into a general section and an instrument/theme specific section. We designed the interview guide to provide nuances and more in-depth views on the themes from the survey, but also covered a few completely new areas, such as a question into the reasons for signing up as an evaluator in the first place. Depending on logistics and preference, the interviews were either face to face or by phone. In some cases the interviews were recorded, but in most, the interviewer took notes. The length of each interview varied but mostly took between 45 minutes and one and a half hour. As the interviews were open, the evaluators had the opportunity to take the lead and put emphasis on subjects they found interesting to focus on. This led to some variations, e.g. some paid a lot of attention to crosscutting issues like Responsible Research and Innovation while others did not mention it at all.

Readers should note that for the sake of anonymity some quotes have been slightly paraphrased in order to secure 100 % concealment of identities. Rewording is of course kept to a bare minimum and nowhere is the substance changed.

3.1 How do evaluators read proposals and what did they say about the quality?

Most evaluators start by familiarising themselves with the proposals by looking at the content list, Part A forms and the abstract, or by a more unstructured skimming of the text to get an overview and a first impression of the quality before a more in-depth reading from "A-Z". Most go through the proposal more than once, spending the second or third time reading looking for things such as coherence between ideas and methods, match between objectives, tasks and deliverables, or checking if the project is "balanced and well-structured".

Interviewed evaluators who estimated their time spent on reading and evaluating proposals indicated that they used at least three hours on each.

Only one of the evaluators used a software tool to track specific keywords throughout the proposals. Almost all evaluators admitted that they quickly get an impression of proposal quality (after a few pages of reading or skimming various key elements of the proposal). At the same time, they assured the interviewers that all proposals get a thorough and fair treatment, as they are fully aware of the large amount of time the proposers use on writing their proposals. The fact that you have to take an independent position on each of the proposal elements also mean that the evaluation must be thorough, as one of the evaluators said.

When asked if there were elements in the proposals that they always or often go over lightly, the evaluators' answers vary. Some mention comprehensive "background chapters", other mention the "annexes": CVs, partner descriptions or WP descriptions, "standard text" or "copy-paste" text used in for example the implementation part of the proposals. It does not mean that it is ignored, but rather that it is read less thoroughly, or only parts of it is read to check the quality or feasibility of these parts of the project. As one RIA evaluator stated "A list of excellent CVs does not guarantee a good project, young researchers might as well have great ideas". Another RIA evaluator argued for not reading the WP descriptions in too much detail because the scientific questions can be solved in different ways; it is therefore difficult to assess whether one or the other approach should be preferred as long as the objectives are met and the challenges will be solved. The ERC evaluators generally read through all parts of the proposals very thoroughly which means that in step one only Part B1 is read, while in step two both B1 and B2 were read in depth.

3.1.1 Project abstracts

Based on the interviews, the abstract seems to matter quite a lot, both as a helping hand to understand the project idea, and as a way to enchant the reader. The evaluators generally find a high quality abstract very appealing. High quality means that the abstract is clear, precise and concise. If it embraces all three main sections; excellence, impact and implementation in a clear manner it also demonstrates from the beginning that the proposers have control over the project. A good abstract also highlights "what's new" in the proposal, and what it is that makes a given project "a winner". Furthermore, the abstract should signal "enthusiasm and excitement", leaving the impression that reading the proposal will be "exciting and a pleasure", as one of the evaluators said.

A good abstract is especially important as many projects span a range of subjects and specialities, and an evaluator might not be equally knowledgeable in all. Here the abstract must provide an explanatory overview. Even when the evaluators are specialists on the subject, the abstract provides them with an overview when dealing with long complicated texts. So proposers need to pay attention to the abstract, and this takes time. Make sure you present your idea, overall goals and level of ambition very clearly from the start.

3.1.2 Layout of the text

Today's technology enables evaluators to make cheap colour printouts or read onscreen (see figure 8), making the choice of layout, colours etc. complex. The interviews indicate that layout, design of tables and figures play a role. If readability is problematic, proposers runs the risk of aggravating evaluators, or as one said during the interview, "Small print in bad colours is really hard to read, very annoying", and one said very directly that "Bad graphics, hard to read, overdone complexity, small print, sloppiness, or just plain repetitions – BAD impression!"

3.1.3 Graphics & Illustrations

When asked about their views, the evaluators almost unanimously agreed that it can be very useful with graphics in order to illustrate concepts, methodology etc. As one evaluator said: "a good illustration also shows that the consortium agrees on the direction of the project". A 'good' graphic presentation is of high technical and conceptual quality. Illustrations and graphics should be simple, but at the same time create an overview that supports the reading. It should of course be highly relevant, and use of graphics should not be "overdone". Some of the evaluators stressed that especially when you are

not an *"expert in the field"* a good illustration can be helpful. One of the evaluators suggested that it might be a good idea to get feedback from non-experts on drafts before graphics goes in the proposals. One evaluator mentioned that diagrams sometimes were used as the basis of discussions during consensus meetings.

3.1.4 Gantt charts

There seems to be a divide between evaluators when it comes to the Gantt charts⁷; some hardly glance at them, others really dig into them. In the interviews, there is a slight tendency towards academics being the most casual Gantt chart-reader, while evaluators from the consultancies and industry place weight on the charts delivering meaningful information. The evaluators mainly use the Gantt charts to validate other parts of the proposal, in particular the implementation section in RIAs or other collaborative types of projects. "Realistic" and "feasible" are the key-words here, as the Gantt chart is seen as the overview that can reveal if e.g. the time planning is faulty, if the coherence between work packages is lacking or if the project seems to be lacking inherent progression.

3.1.5 Verbosity vs. simplicity in the text

When asked about language, complex vs. simple text etc., all interviewees replied that they really appreciate to see things presented clearly, without too many words and without it being more academic than needed. The data from the on-line survey confirm this. 46 % of the Danish evaluators stated that verbose and/or hard to understand language have significant or critical influence on the evaluation. As one evaluator stated:

"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 evaluation report."

⁷ A Gantt chart shows the timing of the different work packages and their components. Gantt charts are a mandatory component in many types of H2020 proposals. Since their importance are often discussed during the writing of a proposal, and this part at times either becomes a time-consuming ordeal or made in the last second as an afterthought, it was decided to ask all interviewees about it.

3.1.6 A summary of what the evaluators in general like when reading a proposal is worth considering during proposal writing

What do the evaluators like?

- The first pages should be exciting. Do not start with 'Adam and Eve', pitch your ideas immediately and answer the questions 'why is it important' and 'how will your concepts solve the problem?'
- A good abstract that pitches the idea, embraces excellence, impact and implementation, and creates curiosity and excitement.
- Make only short background descriptions that convince the evaluators that you are the right consortium for answering the questions and bring the research beyond state-of-the-art.
- Proposals should be well structured, covering the right areas under the different criteria in the proposal template.
- A strong focus on relevance for the project. Nothing even slightly irrelevant should have a place in the proposal.
- Clear and convincing objectives.
- High quality graphics that illustrate the concepts in a simple manner.
- That figures and Gantt chart show the project has a clear idea about how the individual parts and tasks are interconnected and timed intelligently, so the reader can see a well-thought out project plan.

Other useful evaluator points of view

- Long proposals are not necessarily good proposals.
- A poorly structured proposal will inevitably leave an impression of a poorly structured project.
- Layout and readability is important, so make sure you have time to get it right.
 You might think it very superficial, but how the proposal looks can influence how the evaluators perceive it.
- Some proposers have problems in formulating true objectives and confuse objectives with tasks or deliverables.
- A good graphic visualising the concept is easier to keep in mind than two to four pages of text, when discussing the proposals at the consensus meetings.

3.2 Evaluators point of view - specific proposal types

In the following sections, evaluator perspectives on the most important aspects of different H2020-proposal types are highlighted.

3.2.1 Proposals for the European Research Council (ERC)

This section includes statements from evaluators working with all three major ERC project types: Starting Grants (StG), Consolidator Grants (CoG) and Advanced Grants AdG). All respondents agreed that what makes a proposal persuasive is first and foremost that the research question or problem has a fundamental, ground-breaking character, with significant scientific weight. They also stressed that the suggested methods must have the quality and strength to answer the research questions and meet the objectives of the project. It should be clear what the scientific impact is expected to be.

ERC projects were expected by the evaluators to be beyond state-of-the-art, build on new concepts and not be incremental in nature, in other words not be a continuation of ongoing research activities. The proposal should convince the reader that the project is feasible, and be supported by a strong CV and track record of the proposers.

The evaluators stressed the importance of the proposers being able to explain how new the idea really is, and why other current approaches or explanations are not chosen. It is also important that the proposer explains that there may be issues that are unclear or uncertain, but that the proposer has strategies for managing it in the project. "Do they master the methods convincingly is something we among other things test in interviews", as an evaluator said. A good project has clear objectives and goals as well as a realistic time-line/schedule. It does not need to be a complicated or complex project but it needs a clear focus and there should be a 'red thread' throughout the project.

When it comes to the candidates' profile, evaluator expectations depend on whether they are in a StG, CoG or AdG panel. For all StGs, a CV should be well prepared and show that the candidate has real potential. This means that it is clearly above average, includes a few recent high quality publications in high-ranking journals, and at least one as the primary investigator and author.

CVs of more experienced researchers, applying for CoG and StG at the higher end of the eligibility window, should preferable also have published as primary authors in high ranking, more general, scientific journals. It is furthermore seen as very important for the candidates applying for CoG or AdG prove that they are 'still in business', e.g. by having new strong scientific publications showing that they are still able to think in new concepts and get new ideas.

Finally, also solid merits regarding education of PhDs and postdocs will count, as well as having been invited as a speaker at important and prestigious scientific conferences. One of the evaluators said national funding merits were not that important since most panel members or evaluators do not really understand the national funding structures. However, if a proposer has been awarded a prestigious international research prize or similar, this was definitely noted. All evaluators commented that the H-index in itself was not an important criterion for a researcher's merit. The evaluators find in general that the integrity of the panels and of the review process is very high. The panel members will be very aware of e.g. biased scientific judgement by panel members or the external evaluators used in the second stage of the evaluation process.

A summary of what evaluators like when reading ERC proposals

What do the evaluators like?

- A clear and easily readable proposal, with a clear hypothesis, strong objectives and persuasive methods.
- When a project is groundbreaking and at the same time realisable. If possible also that the project description contains a bit of the solution to problems that are investigated.
- That the candidate has established an international network of research collaborators and has been working with the best within the field of research.

Other useful evaluator points of view

- Get to the point quickly; do not waste space on long introductions, repetitive and vague text.
- Remember to address all aspects of the project convincingly in the B1 part including the choice of method and the feasibility of the project.
- The proposal should not 'drown' in too many details, you are not writing a scientific paper.
- An impressive CV will not at all do it alone. The project needs to be outstanding and convincing.
- If you have manuscripts that are still in a review process it might be wise waiting until they are accepted for publication before submitting a ERC proposal.
- Using a strong methodology to answer the research questions is of utmost importance. Great ideas not supported by strong methodologies will very likely fail to convince. In this context, it may be a problem that some fields of research do not have the tradition of using methods that are strong enough to be rewarded with an ERC grant.
- Interdisciplinary approaches are fine, especially when combining new concepts or bringing new methods into the field of research. It is important though that the outcome of the project addresses a well-defined/known basic research community.

3.2.2 The collaborative proposals (RIA, IA and CSA)

Structure of consortia, participant types and profiles

With regards to the often discussed subject amongst proposers of having or not having partners from specific parts of the EU, the majority of evaluators stated that participation from specific countries (e.g. from new member states (EU-13)) in itself was not an issue during evaluation, and was in most cases not even a topic taken up by the experts. A few of the interviewed evaluators mentioned that they had met other evaluators expressing a positive attitude toward EU-13 member state participation, but at the same time stressed that it was still scientific excellence which was the determining evaluation criterion in the panels. Although geography per se was not a criterion, most found a "sound and well thought out" representation of diversity of countries in the consortia important. As one said: "It might not be very convincing if a project with a consortium of only a few Northern European countries are going to provide global solutions". The added value each of the consortium members brings to the project generally needs to be considered very carefully, according to the views expressed. The evaluators will typically be sceptic of consortia that seem to have what looks like "pro forma" partners, e.g. having notably small budgets or no clear role in the project. Some also found it negative if consortia seemed to have a biased representation of partners from one of the participating countries, making it look more like a national project than an EU project.

Concerning SME participation, most evaluators stated that having SMEs or industry in the consortia could be an advantage, but their role should be significant and clear. One said that if it is unclear how the SME/industry partner will benefit from the project, or contribute to the project, the evaluator would check the company descriptions and eventually the company homepages for clarification.

When asked if collaboration on previous (EU) projects is an advantage or not, evaluators almost all say that it can be positive if it is justified, and if it is not just a *"more of the same"* type of project. A few added the comment that the proposed consortium should also include new partners and not only the old ones, as otherwise it appears very unlikely that the call *de facto* will be answered properly.

With regard to the coordinators' qualifications, the evaluators said that scientific experience counts the most, or the coordinator being from a highly esteemed research community. One found it positive if the coordinator came from a strong research institution, as the evaluator might not necessarily know much about the person coordinating the proposed project. Most of the interviewees mentioned that advance knowledge about the proposers did not really matter, as most of them are more or less unknown to the evaluators. Only a couple of the evaluators explicitly mentioned that knowledge about the proposer potentially might affect the evaluation.

Excellence section

Most evaluators expected some sort of introduction in proposals, but at the same time stressed a preference for a short and precise opening. The general view was that the introduction should pitch the central idea and highlight the *"red thread"* that goes from the challenge addressed in the call and the context behind, to the expected outcomes and impact of the project. The evaluators in general found that the wording of the objectives is very important. Objectives should be clear and convincingly described, be

measurable, and should give the direction of the project through all the sections. Most of the evaluators said that it is enough with only a few but high quality objectives which the proposers refer to throughout the proposal.

The state-of-the-art section should primary convince the reader that the consortium *"knows what they are doing and dealing with"*. Although the evaluators generally consider it important, a long description of state-of-the-art is not needed. Quality is preferred over quantity and the section must be understandable to those evaluators that are not specialists in the specific scientific area. One said that the state-of-the-art section should primarily be an identification of knowledge-gaps in need of clarification in order to solve a given challenge, thus being an introduction to a description of how the proposal are going to go <u>beyond</u> state-of-the-art and close the aforementioned gaps.

When asked whether they check references⁸, most evaluators confirmed that they do, but to various degrees. It is often the expectation that at least some in the evaluator group check this. Some evaluators also consult institutional web pages and the researcher's appearance on these websites and other websites. When checking references some use data repositories like Web of Science and PubMed, others rely on their own deep insight into the area and even expect to recognise at least some of the 'keyreferences' cited in the proposal. Some, but not all, also check if references on the latest results are in the proposals, as they expect proposers not to overlook what is new. Others do not check to this degree, mainly due to lack of time, or say they trust the proposers to have a good grip on existing state-of-the-art. Some evaluators simply do not find new results all that crucial if the topic addressed is very specific and/or e.g. asks for more applied type of research.

It was highlighted by evaluators that a good excellence section should include a motivation (why the project should be done, why it is important?), address a problem or challenge, and point to how it could be solved. In addition, the evaluators needed to be convinced that the solution to the problem was innovative and superior compared to known technologies.

Impact section

A good impact section, according to the interviews, should convincingly highlight the importance of outcomes generated by the project. Description of impact, whether on society or business, must be concrete and specific, arguing convincingly for the effects of any future implementation. It is important that the impact section focus on the European "added value". A good proposal addresses how the project will reach the end users and how proposers will ensure that the target groups will benefit from the project.

Almost all interviewed evaluators found the strong focus on impact in H2O2O both important and very relevant in the context of H2O2O. Quite a few of the interviewees stressed the need for the proposers to take the impact section seriously, as it normally carries the same weight as the excellence section. For some of the evaluators, the emphasis on impact was seen as a clear step in the right direction by pushing research results into e.g. better welfare and public health, new innovative products and increased security.

⁸ In Research and Innovation Actions, many applicants strive to demonstrate insight into existing state-of-the-art by referring to articles and other scholarly works

However, the way some proposals addressed impact was found wanting by many of the evaluators. Quite often, they felt that statements on the expected impacts were too unrealistic in light of objectives, or the actions and tasks suggested by the proposers. One evaluator said it seems sometimes like "*it is just something that is taken from looking into a crystal ball*". From some evaluators, the advice was that it is best to be honest and say directly that the expected impact was a "best estimate". The question about defining what 'good impact' looks like was by some seen as difficult. One simply said that it was somewhat impossible to define, as it varies according to the specific context. On the other hand, another evaluator said "a concrete business plan, a set of *KPIs* [Key Performance Indicators] or similar are still more convincing than a lot of promises".

The evaluators were generally in favour of quantitative impact measures - if it is at all possible to produce and present them realistically. At the same time, most of them agreed that it is far from all types of projects that produce outcome and have impacts of a quantifiable nature.

The evaluators were generally not very specific in their comments about expectations to the dissemination and exploitation part of the proposals. However, the more well thought out and realistic the better. One gave the advice that the proposers should "remember to include more and more varied types of instruments to reach the target groups". A policy brief is not enough. Making a dissemination plan that targets the different specific audiences so it actually creates value for them is much more convincing.

Implementation section

A good implementation section from the evaluators' point of view can be summed up as describing the coherence between work packages (WP), convincingly explaining the different roles of the consortium partners in the project, and demonstrating the competences of the beneficiaries and their ability to implement the proposed project.

When asked how they use Gantt⁹ charts and Pert charts¹⁰ in their assessments some find them very important while others do not spend much time on them. They all agreed though that they are important in the sense that the diagrams should create overview, illustrate the coherence (between objectives, tasks, milestones and deliverables) and show progression in the project. They also said that "complex and overcomplicated" diagrams are a sign of a project that is fragmented, not sufficiently coherent and thus unrealistic and unfeasible.

"Good project governance is a must", one evaluator said, "in order to secure a wellsteered project that can handle the various types of reporting as well as risks management". Some evaluators do check the coordinator's and partners' management competencies, while others also focus on the size of budget allocated to project management, simply to assure that funding is sufficient to guarantee a successful project. Once more, the evaluators expressed a liking for keeping it simple but convincing; one evaluator offered a warning, saying that although a lot of "management stuff is pretty standard you should not just fall for the temptation to do a copy paste solution".

 $^{^{\}circ}$ A Gantt chart shows the timing of the different work packages and their components

¹⁰ A Pert chart is a graphical presentation of the components of work showing how they inter-relate

Keep it simple and convincing is also the two keywords when it comes to the evaluators' recommendations concerning the "Consortium as a whole" description. This section of a RIA/IA should illustrate the coherence of the group, and clearly show where the partners complement each other – in short, how each partner brings an added value into the project. Importantly - for projects with lead- and end-user involvement - the section should also describe where and how the users take an active part in the project.

The evaluators generally expect work package descriptions to be pretty detailed, with one saying the norm should be 1-1½ pages per work package. Evaluators also stressed the importance of work packages being linked in a coherent manner, each reflecting its place and role in a logical way when seen from an overall project perspective. At the same time, it was stated that each work package must make sense on its own. A few of our interviewees drew attention to the fact that evaluators are not necessarily specialists on the often very specific scientific descriptions provided in the work packages, making it important for the proposers to ensure that non-specialists can understand a work package at least at a general level. Deliverables must also make sense and reflect the objectives' level of ambition.

As with the rest of the proposal, deliverables must make sense and they should be ambitious. Normally you need quite many, easily 20-30 or more. The description of deliverables should not be overdone and too detailed though. One of the evaluators offered the advice, especially to new proposers, that it is better to have a fairly high number of deliverables, spread over the duration of the project rather than only a few (big) deliverables at the project's finish. Deliverables should be submitted throughout the whole period and not only at the end of the project.

When asked how much they look at the connections between milestones, critical risks and contingency plans, several evaluators found these links very important, and therefore scrutinised them closely. They saw the often-complex projects as high-risk projects and stressed the need for a well-developed risk and contingency plan, which e.g. differentiated between low, middle and high-risks cruxes, clarified how resources were available in case of mishaps etc. Many evaluators stated they also checked how well balanced the budget was in view of allocated tasks, the distribution of person-months and tasks as part of analysing the feasibility of the project. Some admitted that, in honesty, they only looked superficially at the budget.

Partner description and ethics sections

Partner descriptions (section 4) are by most of the evaluators considered quite important as they document that the right competences are present in the consortium. Some even start by reading the partner descriptions in order to get a first impression of the consortium's strength. Others do not read all but select a few and read them in depth. With regard to commitment letters, the evaluators were somewhat divided in their opinions. Some found them useful while others did not consider them much. It seems to be dependent on the context. If it is a letter from an end user that is expected to benefit from the project, it might strengthen e.g. the impact part of the proposal and thus in the end make a difference. If the commitment letters do not really refer to the project and its deliverables it is more seen as an annoying element.

Concerning ethics, (section 5) the evaluators said that this section is a necessary (and thus important) part of the proposal that can easily be filled out in a table-like check-list. One gave the advice that it should not be based on personal judgements.

The cross-cutting aspects of a RIA/IA (RRI, gender and sex)

Most evaluators had limited experiences with the "responsible research and innovation" (RRI) aspect in the sense that not many of the projects evaluated had paid much attention to it¹¹. Only one evaluator stated that RRI was a key issue, but this was due to the nature of the programme evaluated.¹² Some answered yes, when asked if they were briefed on the subject, but not all of the respondents found it an important aspect, with one directly stating "*it* [RRI and gender] *is not something I mind much…it's more for academia*".

Gender and sex, another mainstreamed crosscutting aspect of H2O2O, seems to have had a somewhat overlooked role in the evaluation process. When asked, most evaluators mentioned they had been briefed on gender being one aspect to include in the evaluation, but as one said when asked about the briefing "yes and no, it was mentioned but without emphasis... like it was in FP7". When it came to the issue of gender balance in the consortium, experiences varied, but a single evaluator did experience once that gender could give "an extra point, if the named persons in a proposal are balanced gender-wise."

When it comes to the gender issue, some did look at it, but generally, the impression was that it was not a big issue and therefore not discussed in-depth at the meetings. One evaluator mentioned that the gender aspect is part of getting your homework done and this aspect can also be the determining factor that at the very end makes the difference if proposals are of similar quality in all other aspects.

A summary of what evaluators like when reading collaborative proposals (RIA, IA and CSA)

What do the evaluators like?

- A good balance between countries and between academia and industry/SME in the consortium (but see below).
- A few (ideally between 1-5) strong measurable objectives that function as the "red thread" throughout the proposal.
- A high quality (preferably short) and convincing "State-of-the- Art" chapter that can act as an introduction to the "beyond state-of-the-art" part of section 1.4.
- Quantitative impact measures (if it makes sense).
- Coherence across work packages.
- A simple but convincing management part that clearly describes how to manage the reporting and risk contingency.
- Partner descriptions that convincingly show that each partner have the strength, experience and competencies to deliver the tasks described.

¹¹ One RIA evaluator found it a shame that RRI was not taken more serious as a norm. RRI was never evaluated as "an integrated part of the projects, unless it was part of the scope".

¹² The "Science with and for Society" theme has RRI as a defining component in parts of the topics.

Other useful evaluator points of view

- Geography does not matter unless stated explicitly as a criterion in the call text (but see above).
- It seems that the evaluators prefer goals for impact that are explicit regarding context and balances realism with ambition.
- When it comes to "measures to maximise impact", the evaluators prefer realism and a well-developed plan (as opposed to a long list of un-connected activities).
- Management structure and the description of it should be tailor-made to fit the project.
- Having worked together on projects before is considered positive as long as it is not "old wine on new bottles".
- A very complex management structure also indicates that the project can be high risk (i.e. lack feasibility).
- Do not have too many partners (often 5-9 is enough) in a RIA/IA project and avoid having too many work packages. Each partner should have a clear role.
 Remember to cite yourselves if you are a researcher.
- A Gantt chart should be precise and clear. Especially evaluators from non-academia look carefully at Gantt charts. A Gantt chart should show that the pro-ject is coherently organized i.e. 1) Is the distribution of tasks and work reasonable and realistic? 2) Are the deliverables distributed properly among the partners?

Evaluations and evaluators in Horizon 2020

Enrico Cappellini, Associate professor, Natural History Museum of Denmark, University of Copenhagen. Evaluator on "Research and Innovation Action" applications under "Nanotechnologies, Advanced Materials, Biotechnology and Advanced Manufacturing and Processing".

4. Why become an evaluator?

This section is included to present the evaluators' reasons for being evaluators and what they perceive to be the benefits and drawbacks of being an European Commission evaluator.

The reason is simple; It is important to have far more Danish experts amongst those evaluating the many types of projects in H2O2O and future framework programmes. On this note, it seemed important in this report to cover the actual working conditions of the evaluators.

Overall, being an evaluator is a task that takes time and effort and while your expenses are covered and you are paid for your time, it will not cover you in gold. On the other hand, *all* of the interviewed evaluators mentioned a number of advantages associated with the work.

So, hopefully, this part of the analysis may help encourage experts and researchers to enter their names in the European Commission's expert database (http://ec.eu-ropa.eu/research/participants/portal/desktop/en/experts/index.html).

4.1 Reasons for signing up as an evaluator

Many of the interviewed evaluators state that they were asked to consider evaluating EU proposals by colleagues, their superiors or by the European Commission itself, but the most prevalent reason was - for almost everyone - that they felt they could get something worthwhile out of it.

The evaluators interviewed have vastly different levels of experience in evaluating EU projects, but they all agree unanimously that they benefit themselves – in some way or other – from participating in the evaluation work. Not all 27 interviewees elaborated, some simply said, "Yes" or "Yes, I learned a lot" and moved on, but others gave in-depth answers on different types of benefits. The three most-cited benefits are:

Technical and academic knowledge (20 evaluators cite this specifically)

A recurring theme is the knowledge you get on future developments in your field of expertise.

One interviewee put it like this:

"By evaluating upcoming projects I get deep insight into which way the research field is moving, a fantastic perspective on the future. And I feel my professional range is expanded" And another:

"Academically, I get new knowledge on what is beyond state-of the art, about what is the latest trends and the cutting edge, and that gives me new ideas and inspiration for my own research".

Finally, one respondent succinctly put her technical and academic benefits of participating in the evaluation like this:

"Very much indeed; it is the best paid continuing education you can get!"

Network

The networking of a crowd of knowledgeable persons together during the evaluation week in Brussels is something that is a benefit-driver for many respondents. Fourteen interviewees specifically cite 'the network' as a very positive factor. Some respondents talked about the international nature of the new network and some about the broadness of the network, e.g. meeting people from academia, business, government and NGOs. A few elaborated further on some benefits of their specific new network, with people from that network turning into partners. One very enthusiastic respondent said:

"I have gotten such good contacts through this. Today, I am coordinating H2O2O projects, where project partners are coming from the evaluator network I have established. It is excellent!"

Knowledge on process and proposals

It should be noted that not all respondents (evaluators) are in the business of writing proposals or want to participate in EU projects. Nevertheless, the third most-cited benefit was about how the knowledge gained on the evaluation procedures and from reading and debating proposals had helped them write better proposals themselves. Ten evaluators specifically mentioned this type of impact. One respondent called it a revelation:

"In the beginning, it was an eye-opener, to see how the process was unfolding and it was very good for my own proposals". Another mentioned his/her personal success rate:

"For me, personally it was specifically useful for two reasons: It gives you an idea of how good proposals are structured and how the evaluation process takes place. It has affected my success rate in a very positive way!"

Other comments on benefits for the evaluators

Other benefits noticed, albeit on a much smaller scale, was "the opportunity to influence how EU-funds are distributed", "seeing the effort of bridging between different professions and areas of expertise" and exercising skills in "making good summaries, remembering, arguing, getting good at reaching a consensus".

One single respondent that stands out from the rest, was in stark opposition to this and replied "surprisingly little", when asked about academic and technical insights gained from evaluating. However, the same evaluator still found it "rewarding and exciting to be part of the process" and "learned a little from the proposals". Another person stated a specific disincentive to participate, namely that the pay seemed very low for all the hard work.

In conclusion, an overwhelming number of interviewed experts cited either one, two or all three of the benefits below as central:

- 1. Getting an overview on the forefront of academic and technical knowledge in their fields.
- 2. Building/expanding a broad, international network.
- 3. Improving skills on proposing and getting funding yourself.

4.2 Workload of evaluators

According to the interviews, the standard evaluation workload differs from call to call with an average of around 10-12 proposals per call. Most of the respondents felt that the evaluation itself entailed quite an intense workload and that they had to work hard under time constraints. One respondent specifically noted that the evaluation process at one point had been cut down by a whole day, which meant the panel was working very long hours, but only received standard pay without compensation. Nevertheless, the vast majority, and notably, all who commented on this, felt that the evaluation process itself was an efficient, transparent and fair process, and expressed an overall positive view of the evaluation process.

4.3 Briefings before the evaluation?

Our questions about the level of briefing received prior to evaluating tried to ascertain whether the briefings were adequate and fulfilled their role of preparing the evaluators to participate fully in the process. All respondents that replied to this question had participated in a briefing and most found it very helpful. In addition, the majority of briefings seemed to have both a more generic part (H2O2O generic) and a more specific part, relating to the precise topic in question. Some of the more experienced evaluators noted that the briefing beforehand of evaluators had improved in H2O2O, and there was a notable overall satisfaction with the briefings and the support from the European Commission in this regard. However, a few respondents felt that briefings were mostly for newcomers, and on the other hand, some newcomers still felt uneasy after the briefings.

4.4 Would you do it again?

The vast majority of the experts said outright that they would be happy to take part in more evaluations, but some had practical reservations of it not colliding with other important work and the topic being central to their field of expertise.

Evaluations and evaluators in Horizon 2020

Signs inside the Covent Garden Building, Bruxelles.



Photographer: David Beaugnet, European Research Council Executive Agency (ERCEA)

5. Comments on the H2O2O evaluation procedures

Based on the survey and interviews, this section offers some recommendations for ensuring the evaluation process can retain the generally high regard amongst proposers.

Composition of expert groups – It is somewhat worrying that 28 % of the respondents found that the competences present in expert groups only partly meet the needs¹³, all the more since a number of the comments expanded on this in very direct and critical ways. While a little less critical, some of the interviews too revealed reservations about the quality of the peer review they had experienced in H2020. This is especially valid for the evaluations in the Industrial Leadership and Societal Challenges pillars, whereas ERC evaluators found the evaluation process *"as fair as it can be"*, as one evaluator summed it up. The European Commission is aware of some of the critiques that are voiced, so the recommendation from this report is only that these issues are not neglected.

Remote consensus meetings and remote evaluation - While it was seen as understandable and commendable that the European Commission try to keep the cost of the expert evaluation low, the introduction of remote evaluation and consensus meetings as a money-saver is a novelty to which some of the Danish evaluators objected. In general, the physical meetings, where experts can discuss "across the table, which is a must when we are talking about large projects", was mentioned several times as better than on-line meetings. Some interviewees felt that given the time and effort spent by the proposers, the evaluation process should match this in quality, and this quality could so far only come with in-person meetings.

If remote evaluation and consensus is to become more of a success, it also demands that the IT systems available to the experts are of high quality and easy to use, and backed by excellent guidance. From both survey and interviews, there is a number of critical remarks concerning both the quality of the instructions and the IT systems themselves, which is why it is hoped that the European Commission will continue their improvement measures.

Evaluation of Impact sections - Almost all evaluators found the strong focus on impact in H2O20 both important and very relevant in the context of H2O20. Some of the evaluators mentioned explicitly that the emphasis on impact was a clear step in the right direction because it was seen to be pushing research results into e.g. better welfare and public health, new innovative products and increased security.

¹³ See figure 11. In addition, four % of the respondents replied "no" to the question "In your opinion, was the composition of competences in the group/panel adequate?"

However, 49 % of the evaluators responding to the on-line survey stated they were not completely confident evaluating the impact of proposals. Given the nature of the concept of "Impact" and its dependency on the context of the specific project, it is perhaps somewhat understandable that almost half of respondents were not "completely confident". On the other hand, it could also easily be interpreted as far too many 'less-than-confident' evaluators, which would be a problem, especially in light of the significance of Impact for the goals of H2O20. Overall, the sum of 49 % seems to underline the importance of the efforts of the Commission in establishing and maintaining easily accessible and high quality specific guidance for the evaluators on Impact.

Conflict of interest – A number of experts mention this as being very strictly observed, and while the evaluators have full understanding of the need for avoiding any kind of bias, some mentioned it as a problem. In some scientific areas the number of true specialists are limited, and many of them have some sort of previous engagement (e.g. as co-authors, participating together in projects etc.), thus making researchers avoid the role as expert or being rejected by the European Commission, shrinking the pool of eligible evaluators.

Diligence by experts and the European Commission in avoiding accusations of biased expert reviews can of course only be applauded. Nevertheless, it is necessary that the above-mentioned problem are addressed so competent evaluation are ensured at all times.

The crosscutting issues - In both survey and interviews, the experts was asked about crosscutting issues (Responsible Research and Innovation, gender and ethics), since they are promoted in all parts of H2O2O and therefore should be taken into account by proposers. It should be noted, that given the importance the EU have placed on the mainstreamed issues, it was not always touched upon in briefings (or at least left an impression) as indicated in the figure below.

Figure 17

Please indicate which of the following subjects you were briefed on before panel meetings or consensus meetings



It is recommended that the briefings on the crosscutting issues are done so they more directly reflect the importance, since it could otherwise be feared that experts and proposers alike will disregard it.

As a closing remark, it is a strong impression that the Danish H2O2O experts are dedicated and very much aware that they have an important task. They generally expressed confidence in the way proposals are reviewed, and in many of the interviews, the experts highlighted the merits of the H2O2O evaluations. Quite many of the complaints heard were not addressed at the processes or methods used as such, but against a *specific* situation, e.g. another expert who tried to ride a hobbyhorse, a heavy-handed official trying to impose a point of view etc.

Hopefully, the report will provide future proposers, support personnel and the European Commission with some useful insights that can ease work and increase interest in H2020 and future Framework Programmes.

Evaluations and evaluators in Horizon 2020

About the report

Project design and survey

Cecilie Ydemann Hansen, Office of Research Services, University of Copenhagen Louise Thers Nielsen, Office of Research Services, University of Copenhagen Torben Høøck Hansen, Office of Research Services, University of Copenhagen Christian Holstein, Danish Agency for Science and Higher Education

Interviewers and report

Cecilie Ydemann Hansen, Office of Research Services, University of Copenhagen Torben Høøck Hansen, Office of Research Services, University of Copenhagen Christian Holstein, Danish Agency for Science and Higher Education Svend Binnerup, Research Support Office, Aarhus University

Technical design of survey, data management

Rasmus Damkjær, Office of Research Services, University of Copenhagen Nikolaj Herning Gitz-Johansen, Office of Research Services, University of Copenhagen

If you have questions or comments about the survey or the report, please contact Senior Executive EU Liaison Officer Torben Høøck Hansen, Office of Research Services, Research & Innovation, University of Copenhagen (torben.hansen@adm.ku.dk / direct phone 35 32 07 74), or Special Advisor Christian Holstein, Danish Agency for Science and Higher Education (cho@ufm.dk / direct phone 72 31 82 66).

Evaluations and evaluators in Horizon 2020

Evaluations and evaluators in Horizon 2020



Danish Agency for Science and Higher Education