Proposal selection for objective-oriented research funding

A world wide review of good practices in proposal evaluation for research and innovation grants



Patries Boekholt
Innovation Policy Matters
August 2018

CONTENTS

EXECUTIVE SUMMARY iii

1 Introduction 1

2 Policy background for the review 2

3 Objective and approach of the review 3

3.1 The review's objective 3

3.2 The review's approach 6

3.3 The agencies and programmes under review 7

4 Literature on proposal evaluation approaches 11

5 Evaluation process designs 15

5.1 The basic designs of proposal evaluation 15

5.2 Proposal evaluators 19

5.3 Other features in assessment processes 21

5.4 Final decision making and portfolio management 22

6 Evaluation Criteria 24

7 Lessons for future Framework Programmes 27

8 Recommendations 31

List of Acronyms 33

Appendix I Detailed Case Descriptions

Appendix II Stakeholder Workshop Report

Appendix III Interviewees and other persons consulted

LIST OF FIGURES

Figure 1 Aspects of the detailed review of evaluation processes in R&I agencies 4
Figure 2 A hierarchy of mission-oriented / objective driven structures 9
Figure 3 NSF Proposal and Award Process and Timeline 16

LIST OF TABLES

Table 1 Examples of single submission proposal evaluation processes 17
Table 2 Examples of double submission proposal evaluation processes 18
Table 3 Models using internal and external evaluators 21
Table 4 Award criteria for Grand Solutions Programme IFD IV

Executive Summary

The purpose of the study

Major funding decisions are taken based on proposal evaluation processes for Research and Innovation (R&I) grants. Over time the European Commission has developed a state-of-the-art system of proposal evaluation for its consecutive Framework Programmes. Nevertheless, the preparation for Horizon Europe offers a suitable opportunity to consider whether improvements could be made to the existing evaluation processes and criteria. The purpose of this study is to synthesise good practices world wide for grant proposal evaluation, and possible evaluation processes under Horizon Europe. The focus is on the evaluation criteria, and procedural steps, as applied to proposals for 'top-down' objective-oriented research proposals, including 'missions'.

The approach

The first step of the review was to identify agencies and research funders that were likely to have top-down and objective-oriented research funding mechanisms and agencies that have launched 'missions'. This was done using expert knowledge and research and innovation policy literature. The emphasis was on those with 'societal challenge' type objectives, similar to today's Horizon 2020's Third Pillar. From a long list of relevant agencies world wide a choice was made to focus on those with potentially good practice features relevant for Horizon Europe. The case studies are from Europe and the US. In US the research funding Agencies selected are the Defense Advanced Research Projects Agency (DARPA), the Energy Department's Advanced Research Projects Agency Energy (ARPA-E), the National Institutes of Health (NIH), and the National Aeronautics and Space Agency (NASA), which are all by definition mission-oriented. In Europe cases were selected of dedicated "grand challenge" driven programmes: the Grand Solutions Programme of Innovation Fund Denmark and the Challenge Driven Innovation Programme of Vinnova in Sweden. In addition, examples were selected from agencies with societal challenge driven R&I programmes, i.e. the Environmental Protection Agency (EPA-EI) in Ireland and the Water Technology programme by Netherlands Organisation for Scientific Research (NWO).

The review of cases is mostly based on desk review, consulting, where possible, written guidelines for both applicants and peer reviewers, material for specific research calls and general reports and documents published by the agencies discussing their proposal evaluation

processes. Additional interviews were held to elaborate on how the written procedures are implemented in practice, using a structured checklist. Academic and policy-oriented literature on proposal evaluation and particular peer review was also consulted to set the context of the good practice review.

After the draft results were synthesised in an executive summary, a stakeholder workshop was held to discuss the findings and possible additional points of view and ideas from the participants. Using the discussions and feedback from the Stakeholder Workshop all these inputs have been synthesised in this Final Report which draws conclusions and provides some recommendations for Horizon Europe.

Findings

The review clearly brought to light that there is no one 'best-practice' in proposal evaluation. The Agencies reviewed adapt their evaluation processes and award criteria to the objectives of their missions, programmes and specific calls. The political and administrative context of the agencies in different parts of the world has a strong influence on how Agencies organise their grant evaluation processes. In US mission-oriented agencies, programme directors and officers have the discretion to decide on the specific features of calls and their evaluation processes.

In comparison with other agencies world wide the Commission uses what can be called state-of-the-art evaluation process designs. The range of designs found in other agencies (one-stage submissions, two-stage submissions, stage-gate approaches) are in Horizon's 2020 rules of participation. Some additional features could be considered that other agencies use such as face-to-face meetings between proposal reviewers and applicants or video-conferences with the same purpose. None of the objective or mission-oriented funding agencies and programmes reviewed use blind evaluation as a method to overcome unfairness in the selection procedures.

The set of three generic award criteria used in Horizon 2020 – excellence, impact and quality and efficiency of implementation - can in principle cover the key assessment criteria that the Agencies under review also use. With the sub-criteria that can be adjusted, the spectrum of judgments that proposals typically undergo in selection processes can be covered. In agencies reviewed scientific and technological merit are key award criteria, and in most cases knock-out criteria. However, additional criteria such as feasibility to deliver outputs, relevance to the call and mission, expected societal and economic impacts are on equal footing to decide on funding of proposals. In particular when funders ask proposers to address complex societal challenges, the award criteria related to the background and composition of consortia receive greater weight. In the European examples the inclusion of stakeholders involved in the

implementation of the project outcomes is a prerequisite for awarding proposals. In all reviewed cases the 'strategic fit' of a proposal with the mission or programme objectives formed a key consideration for funding.

Almost all reviewed agencies involved their internal staff at some stage in the proposal review. Particularly in the US agencies internal staff has a large discretion in the award process. There were only few examples where stakeholders were actively involved in the review process. In these examples they were typically from other policy sectors and public authorities essential for the implementation of research results. No example was found with citizen involvement in the proposal review process.

Portfolio-management of mission-oriented proposals is, if at all, done in the last stages of the decision-making when the ranked proposals are prioritised. In general, this is done by the Agencies' Management Board or by a dedicated Advisory Panel that considers the set of proposals that have passed all thresholds of the award criteria. No example was found where portfolio-management was done by the reviewers assessing individual proposals.

Recommendations

The findings of the review lead to the following four recommendations addressed to the Commission to consider for Horizon Europe:

1. Use the **current set of three generic award criteria**, and depending on the objectives of the mission, (work)programme and call, apply more **flexibility in their weighting** and in the **definition of sub-criteria.**

In comparison with the Agencies under review, the Commission could pay more attention to elaborating what is intended in terms of a strategic fit with the objectives of a call, particularly as independent reviewers chosen for the scientific and technological expertise might not be informed well. Contrary to bottom-up and curiosity driven research where the quality of research (excellence) is the key award criterion, in top down and objective driven research, proposals should not only be of excellent quality, they must have a fit with the policy intention of a call or programme. In order for this to be thoroughly assessed, not only the applicants but also the reviewers should be well informed how projects are intended to contribute to the mission.

2. If missions require a **portfolio** of projects addressing different aspects of the mission's objectives, **combine a sophisticated 'front-end' strategic process** (i.e. fine-tune the call in relation to already funded R&I activities and outcomes) that can identify

priorities and gaps allowing to focus calls on completing a portfolio **with a transparent 'back-end' process** where the set of ranked proposals that passed the necessary quality thresholds but still need to be prioritised for funding can be re-arranged to complete a portfolio.

In cases where a mission approach would ask for a set of R&I activities, spread across multiple (potential) projects, **portfolio management** of proposals could be an option. However, particularly as independent proposal reviewers do not have the overview of all proposals and ongoing projects, this is **not a task that should be integrated in the individual proposal quality review process**. In a similar manner as some of the reviewed agencies, those proposals that have passed a certain threshold of scores on the award criteria, could be assessed by a high-level Board or Panel to decide whether the prevailing ranking could be altered/adjusted in favour of proposals that address a significant gap in the required/desired research activities. Alternatively, if the thematic gap is significant and not time sensitive, additional calls could be launched specifically focussed on that gap.

3. As Horizon Europe is designed for greater impact, develop more sophisticated concepts and tools that applicants should use in their proposals to mark their contribution to the 'pathway of impact'.

Examples in the review such as mapping stakeholders interested in the results of the project, the visualisation through intervention logic maps, user-oriented communication plans and the description of societal readiness levels could be further explored. This should be done in a way that helps independent experts to assess the credibility and likelihood of impact claims, without burdening the applicants with extensive additional socio-economic analysis.

4. For missions implying transformational changes with a complex set of participants and stakeholders: experiment by **using a stage-gate approach** similar to the Swedish Challenge Driven Innovation programme. This allows for a greater number of competing and high-risk initiation studies to develop -if successful - into larger implementation and demonstration projects using different weightings of award criteria in different stages.

The information and views set out in this study are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

1 Introduction

Major funding decisions are taken based on proposal evaluation processes for Research and Innovation (R&I) grants. Over time the European Commission has developed a state-of-the-art system of proposal evaluation for its consecutive Framework Programmes. Nevertheless, the preparation of each new Framework Programme offers a suitable opportunity to consider whether improvements could be made to the existing evaluation processes and criteria, currently laid down in the rules of participation for Horizon 2020.

This study was commissioned to review and synthesise good evaluation practices worldwide. The focus of the study is on the evaluation criteria, and procedural steps, as applied to proposals for 'top-down' objective-oriented R&I proposals including "missions". Chapter 2 sketches the policy context for this review in light of the preparation for Horizon Europe. The approach of the review and the stakeholder workshop that was held on behalf of this study is elaborated in Chapter 3. This chapters also elaborates on the different Agencies that have been reviewed. Chapter 4 has a succinct discussion on the key topics related R&I proposal reviewing as discussed in the literature. Chapter 5 describes the Evaluation Process steps identified in the Agencies that have been reviewed, including their design, the types of evaluators involved and other issues related to the evaluation process. Chapter 6 discusses the Evaluation Criteria as used in the cases that have been reviewed in more detail. Chapter 7 summarises the findings and draws the key conclusions. The last Chapter 8 has a number of recommendations that the Commission could consider for Future Framework programmes and notably Horizon Europe. In Appendix I has detailed descriptions of the Agencies and programmes that have been reviewed in more detail for this study. Appendix II has the Stakeholder Workshop Report. Appendix III lists the people that have been interviewed and consulted for the study.

2 Policy background for the review

The European Commission's proposal for Horizon 2020's successor Horizon Europe, published on June 7 2018, introduces missions as a new provision under the Global Challenges and Industrial Competitiveness Pillar. This pillar will be designed to be more impact oriented. A small number of missions with specific goals will be launched from within the pillar, establishing a comprehensive portfolio of projects, but also drawing on relevant activities and outputs from other parts of the programme. Both principles, the impact orientation and mission-orientation have been for a large part inspired by the High-Level Expert Group that published the so called Lamy report.[1] The report suggested that to design for greater impact, the next Framework Programme should develop different inter-connected pillars, including one focused on societal challenges, possibly including large scale missions. The report also suggested to adapt the proposal evaluation to the different pillar's objectives.

In the preparation towards Horizon Europe a number of reports have been published on the potential of mission-oriented research with a global challenge driven focus. The 2018 JIIP report states that there is no one singular definition of mission orientation. [2] However, directionality and intentionality are core features differentiating them from other types of policy. The report distinguishes two ideal-types: initiatives oriented towards (faster) scientific and technological advancement and those targeting societal challenges implying transformational changes. Particular in the last type of ideal type it can be expected that funded R&I proposals have a clear focus on a particular (pathway to) impact and consist of a number of different partners that have a role to play in the transformation.

A 2018 report by Mazzucato[3] stated that mission-oriented research asks for a different way of evaluating proposals. As missions are complex, proposal evaluation in her view should look at a portfolio of projects as well as at the excellence of individual proposals. Both reports pointed towards the involvement of stakeholders and users in the policy process.

[1] High level Expert Group (chaired by Lamy), (2017), Lab -FAB-APP, Investing in the European future we want, European Commission, Luxembourg

[2] Joint Institute for Innovation Policy et. al, (2018), Mission-Oriented Research and Innovation: Inventory and characterisation of initiatives, Report for DG Research and Innovation

[3] Mazzucato, M.,(2018), Mission-Oriented Research and Innovation in the European Union, European Commission, Luxembourg.

These developments form the European research policy context for this review into proposal evaluation processes and criteria. The key research question is whether processes could be further updated inspired by good examples from research and innovation agencies across the world. While this mission-oriented and objective driven approach in Pillar II of Horizon Europe has been one reason to conduct this study, the report has also looked at certain other practices in proposal evaluation.

3 Objective and approach of the review

3.1 The review's objective

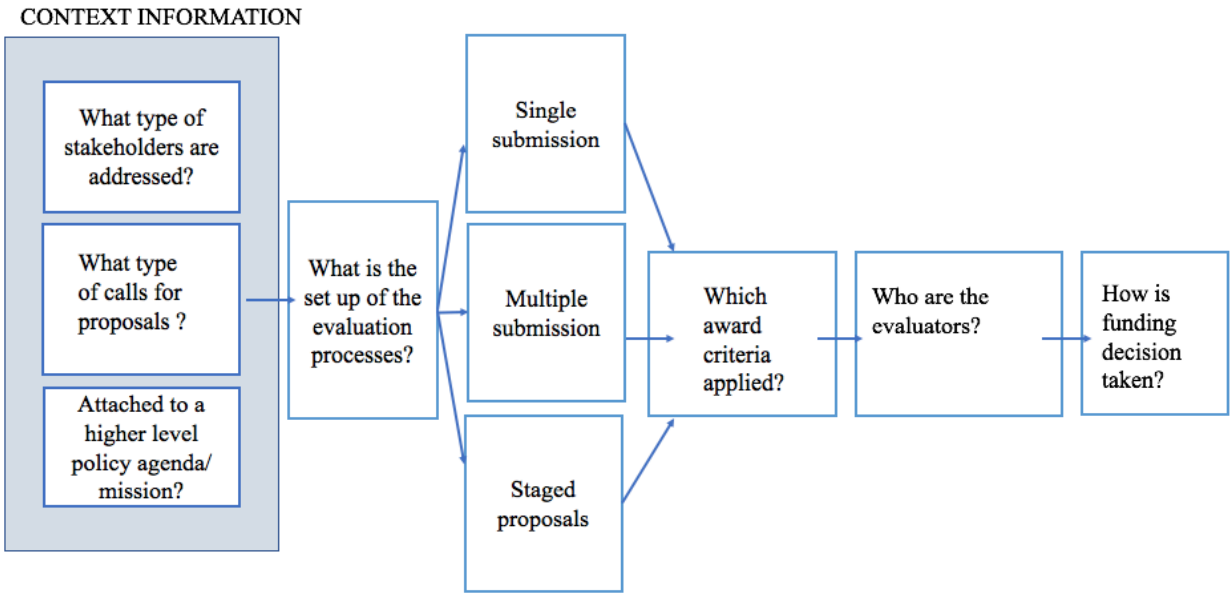
The objective of this review is to synthesise good practices worldwide for grant proposal evaluation, and possible evaluation processes under Horizon Europe and make recommendations.

The focus is on the evaluation criteria, and procedural steps, as applied to proposals for 'top-down' objective-oriented research proposals (including "missions"). A limitation was also made to focus on **research grants**, thus excluding selection procedures for prizes, single company innovation loans and subsidies and other types of policy instruments. The following Figure 1 provides a schematic view on the components of proposal evaluation that were part of the study. The very first step after receiving proposals, the eligibility check, that is mostly conducted by the administration of funding agencies has not been included in the detailed cases.

It is important to note that the review does not look into processes to select the missions themselves but rather, given a defined mission, the operational processes to select individual research and innovation projects that should contribute to the mission.

The scope of the review allowed to zoom in on a limited number of agencies and their practices. The first step of the review was to identify agencies and research funders that were likely to have top-down and objective oriented research funding mechanisms. The emphasis was on those with 'societal challenge' type objectives, similar to today's Horizon 2020's Third Pillar. This excluded funding approaches for curiosity driven basic research and non-thematic funding (e.g. European Research Council (ERC) type funding) as well as funding that is solely focused on competitiveness and (individual) firms. Appendix I has the description of the Agencies and/or programmes reviewed in more detail.

Figure 1 Aspects of the detailed review of evaluation processes in R&I agencies



Source: Patries Boekholt – Innovation Policy Matters

Although the study is also looking at a possible 'mission' approach in Horizon Europe, this was not the only avenue to look for good practices in proposal evaluation that could be considered for future Framework Programmes. The question for example how the objective oriented R&I funders deal with the 'fairness' principle by using blind evaluation was another consideration.

The shift of emphasis from funding bottom-up science (with 'excellence' as its dominant award criterion), towards 'mission-oriented' research leads to a number of hypotheses:

1. *Compared to bottom-up research programming, it can be expected that in objective-driven top-down approaches the balance shifts from 'scientific merit' as the key evaluation criterion to selection criteria related to relevance, impact and 'strategic fit' with policies or missions.*
2. *Given the complexity of societal challenges, the target groups 'carrying' the proposals become more complex and multi-disciplinary and include users, parts of value chains and a greater mix of actors.*
3. *In objective driven top-down research programmes one could expect that a substantial part of the reviewers are not chosen as scientific/academic specialists but in another capacity.*
4. *To ensure that projects are reviewed as components achieving a bigger mission more 'portfolio review' approaches can be expected, comparing proposals with a view how they fit as a set of complementary solutions to a bigger operational objective.*

There are a number of **boundary conditions** for proposal evaluation that affect all funding agencies. Evaluation practices need to take these boundary conditions into account and find trade-offs between different positions. For instance, a review that has multiple review steps using elaborate approaches to assess the merits of proposals is likely to cost more time and

resources which is in the general sense is not beneficial to the research community and the applicants.

From the perspective of **policy makers**:

- the evaluation process needs to ensure that the 'best' proposals are chosen, i.e. those that meet the objectives of the funding programme best. From this perspective a thorough review of the proposals and its proposers is necessary to ensure the best value for public investment in research and innovation.

From the perspective of the **funding agencies**:

- the time and resources for the entire review process should be manageable and in line with the funding initiatives. Choices need to be made on the costs and administrative burden on the organisation for selection processes;
- the evaluation processes should be attractive enough to appeal to the targeted proposers; and
- the limitation of available funds means that only part of the proposals can be funded requiring a ranking of high quality proposals.

From the perspective of the **beneficiaries:**

- the time resources needed to complete and submit proposals: this is the users side of the administrative burden. The required efforts need to be in balance with the chance of being rewarded for these efforts. The evaluation process can help to set that balance; and
- the time between proposal submission and signing of the contract (time-to-contract) should not be too long.

It was clear from the review that all R&I funding organisations make choices on what evaluation processes to use by finding a balance between these boundary conditions. The Danish Innovation Fund (DFI) for example wants to adhere to the principle that time-to-contract is no more than 100 days. The logistics of panel reviews and interviews with the best applicants need to be very tightly managed beforehand. The National Science Foundation (NSF) Biological Science Department in the US has experimented with two stage proposal submission in order to limit the burden on the organisation's programme officers. After an evaluation it found that this only increased the burden of its staff and the department went back

to one-stage submissions. The high administrative burden for applicants has led to more experimentations (e.g. two-stage proposals with very short first proposals).4

3.2 The review's approach

Using expert knowledge and literature, a first scan was made of R&I funders that have a clear policy mission and of Agencies that have 'challenge driven' top-down programmes. Section 3.3 gives a description and explanation of the R&I funders that have been screened and from which a number of more detailed cases were selected with potentially interesting elements in their proposal evaluation processes.

The review of cases is mostly based on desk review, consulting, where possible, written guidelines for both applicants and peer reviewers, material for specific research calls and general reports and documents published by the agencies discussing their proposal evaluation processes. Additional interviews were held to elaborate on how the written procedures are implemented in practice, using a structured checklist. Given the scope of the review access to published and transparent evaluation processes as well as availability of interview partners with an overview on the processes for particular agencies and programmes was also a consideration for selecting the cases.

Academic and policy-oriented literature on proposal evaluation and particular peer review was also consulted to set the context of the good practice review.

After the draft results were synthesised in an executive summary, a stakeholder workshop was held to discuss the findings and possible additional points of view and ideas from the participants.[5] A separate Stakeholder Workshop Summary Report has been published reflecting the findings and can be found in Appendix II.

Using the discussions and feedback from the Stakeholder Workshop all these inputs have been synthesised in this Final Report which draws conclusions and provides some recommendations for Horizon Europe.

4 Jeffrey Mervis, (2018), Science, American Association for the Advancement of Science, see: http://www.sciencemag.org/news/2017/10/nsf-drops-preproposals-deadlines-biologists-seeking-funding and Abt Associates, (2017), Evaluation of the NSF DEB and IOS Switch to Preliminary Proposal Review: Final Report, Prepared for NSF.

5 Held in Brussels on April 25th 2018 with representatives from various umbrella organisations from science, research and industry including some experts and practitioners on proposal evaluation for research and innovation.

3.3 The agencies and programmes under review

The review aimed to find examples of review approaches used by other R&I funders than the Commission. A selection was made of research and innovation agencies across the world and particularly those with a 'top-down' or policy challenge driven approach and funding programmes. The examples sought in the review can be presented as a hierarchy of mission-oriented/ objective driven structures. While there are quite some high-level national strategies that can be presented as a 'mission-approach', they are not always matched with a dedicated implementation plan. They act as 'Accelerators' adding additional 'earmarked' funding to existing research and innovation agencies, such as the Beau Biden Cancer Moonshot in the USA and the Brain Cancer Moonshot in Australia. There are government agencies that have a clear objective driven mission, often in the sectoral policy areas (e.g. Environment, Energy, Transport etc.) that amongst other activities have an R&D pillar such as the Irish Environmental Protection Agency and Defense Advanced Research Projects Agency (DARPA) in the US (see below). In some cases, these national strategies do lead to the launch of a dedicated R&I programme or initiative such as the Swedish government's 2030 Agenda for Sustainable Development leading to the Challenge Driven Innovation programme at Sweden's Innovation Agency Vinnova.

As an advanced R&I country the USA was a starting point, as challenge driven and mission-oriented R&I has been in operation there for some time across different policy domains. The review looked at several 'sectoral' policy agencies and funders with objective driven R&I as well as USA's key research funder the National Science Foundation (NSF). There are a number of significant differences in the role, the governance and mandates of the US agencies in comparison to most EU agencies and in particular the European Commission executive functions. These differences have an impact on the way proposal selection is organised. While this gives interesting perspectives in the search for good practices, it also calls for caution as it limits the transferability to the EU context.

In US research funding Agencies such as the Defense Advanced Research Projects Agency (DARPA), the Energy Department's Advanced Research Projects Agency Energy (ARPA-E), the National Institutes of Health (NIH), The National Aeronautics and Space Agency (NASA) are by definition mission-oriented. Each of these Agencies have their specific mission, often very broadly defined. For example, DARPA's mission is to make pivotal investments in breakthrough technologies for national security. NIH's formally stated mission is "*to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability*". The USA's officially launched Cancer Moonshot that Europeans often refer to as an example of a Big Mission, fits perfectly well in NIH's 'normal' research funding approach and as such does not

have different funding mechanisms or evaluation procedures compared to mainstream funding activities.

Thus, the term 'mission-oriented' research and Grand Challenges as used in Europe has quite a different meaning in the US context. Diane Hicks (2016) gives an interesting historical account of how the terminology of Grand Challenges emerged in US Science policy in the 1980s, mostly defined in scientific and technological terms, and taken over in Europe by shifting it towards societal challenges.[6]

Other characteristics that have a big influence on the evaluation processes in US agencies are:

- programme directors in US agencies can act quite independently and thus also decide on the characteristic of the calls, the design of the proposal evaluation process and the award criteria by which to choose between competitive proposals. Therefore, within one agency or even activity area there can be quite a variety of evaluation processes;
- programme managers and project officers at US agencies are mostly hired as scientific and/or technological experts in a specific domain and also have a much larger discretion to decide on the contents of the call, the design of the evaluation process and the precise definition of the selection criteria; and
- the calls for research proposals (called a Funding Opportunity Announcement (FOA) across all agencies) particularly of the more technology-oriented agencies (ARPA-E, DARPA, NASA) are often very narrowly defined in terms of the technological solutions that are called for. Much more so than in for instance EU calls, the specifications of the activities and outcomes expected of the research organisations and companies are prescribed by the scientific and technological experts (i.e. the project officers and managers of the agency). Thus, it makes more sense that they also perform the proposal selection internally.

In Europe agencies were selected with dedicated "grand challenge" driven programmes (the Grand Solutions Programme of Innovation Fund Denmark, Challenge Driven Innovation Programme of Vinnova in Sweden). In addition, examples were selected from agencies with societal challenge driven R&I programmes, i.e. the Environmental Protection Agency (EPA-EI) in Ireland and the Water Technology programme by Netherlands Organisation for

[6] Hicks, D. (2016), Grand Challenges in US Science Policy attempt policy innovation, International Journal Foresight and Innovation Policy, Vol. 11, Nos.1/2/3, 2016.

Scientific Research (NWO) in partnership with the water technology research centre Wetsus in The Netherlands.

The following Figure 2 illustrates this 'hierarchy' with some examples from the study.

Figure 2 A hierarchy of mission-oriented / objective driven structures

Hierarchy of mission-oriented/objective driven structures	Examples discussed in review
National Government Strategies and 'Missions'	Beau Biden Cancer Moonshot, Brain Cancer Moonshot (USA), Sweden's 2030 Agenda (SE), Industry Challenge Fund (UK)
Agencies with a Top-Down Mission Approach	EPA(IE), ARPA-E (USA), DARPA (USA), NASA(USA)
Agencies with some top-down R&I programmes amongst mostly bottom-up research programmes	NWO (NL), NSF(USA)
Objective Driven R&I Programmes	Challenge Driven Innovation – SE Grand Solutions Programme - DK

Source: Patries Boekholt - Innovation Policy Matters

Initially examples were considered such as the Industry Strategy Challenge Fund (ISCF) initiated by the UK Department for Business, Energy & Industrial Strategy (BEIS) and managed by UK Research and Innovation (UKRI). This fund is set up as part of a long-term strategy to raise productivity and earning power in the UK. The Fund invites consortia to submit expressions of interest to define Challenges that are in line with the objectives of the government's Industry Strategy. This Strategy defines four grand challenges: 1) artificial intelligence and the data economy, 2) clean growth, 3) the future of mobility and 4) meeting the needs of an ageing society. In the first two waves of invitations of proposals 15 Challenges have been launched.[7]

Basic requirements are that these proposal for challenges are business led and chose topics where the UK already has strong world leading research and business. From the long-list of Expressions of Interest for Challenges UKRI makes a short list of potential new challenges and sets up panels with representatives from business and academia from the domain to elaborate

[7] https://www.ukri.org/innovation/industrial-strategy-challenge-fund/

in detail what the challenge is exactly. This can be backed up by evidence such as forecasts and market data. The panels have to develop a plan how the money will be spent if their Challenge gets chosen. Subsequently at high level (involving BEIS, UKRI and ISCF Steering Board) a decision is taken which Challenge will be awarded Funding. Once a challenge is chosen, a director is appointed. For the management of the Challenge BEIS has opted for the DARPA model: the director decides the nature of the calls and how projects are selected. While this example is interesting as an example of how Challenges or Missions can be selected, for the topic of **project** selection the initiative is too recent to acquire a detailed overview of the R&I grant evaluation processes applied within each of the challenges.

In Australia an independent Advisory Board has recently proposed to launch a National Missions Programme in the Australia 2030: Prosperity Through Innovation report.[8] In May 2018 the Australian Government endorsed the recommendation to launch ambitious National Missions and to develop a framework to identify and implement National Missions. The launch of a Genomics Futures Mission was announced with additional government funding for the next ten years.[9] In October 2017 the Australian Minister for Health and Sport already launched a Brain Cancer Mission with additional funds from the Medical Research Future Fund, which has launched its first calls for grants in 2018.[10] Research grants for this mission will be implemented by a number of different funding partners, including foundations. The government's umbrella organisation Cancer Australia coordinates the mission. Thus, there is not one single proposal evaluation process for this specific mission.

In the initial phases of the review good comparators were sought in the Asian countries China, Japan and South Korea. However, in these three countries research and technology agencies either focused at investigator driven basic research or single technology-oriented programmes for specific strategic industrial sectors. There are only few open competition collaborative programmes, comparable with Horizon 2020 programmes and funding for applied research is mostly directed towards industry and technology oriented public research institutes.

The New Energy and Industrial Technology Organization (NEDO) in Japan has Strategic Innovation Programmes, which run open competitions with an emphasis on applicants from industry, for instance one related to energy conservation technology. Similar to the Vinnova

[8] Innovation and Science Australia, (2017), Australia 2030, Prosperity through Innovation, see https://www.industry.gov.au/data-and-publications/australia-2030-prosperity-through-innovation

[9] See https://www.industry.gov.au/data-and-publications/government-response-to-innovation-and-science-australias-australia-2030

[10] https://canceraustralia.gov.au/research-data/research/australian-brain-cancer-mission

Challenge Driven Innovation approach it calls for projects in three different stages: 1) an Incubation Research phase, 2) a Practical Application phase and 3) a Demonstration Phase. For each phase the applicants have to submit a new proposal and go-no-go decisions are taken whether projects can go to a next phase.

When drawing any lessons from examples from across the world one should be aware of the different research and innovation governance structures in which agencies operate, their specific mandate and general ways of operation. Particularly in the European context where the cooperation between Commission and Member States bring additional considerations, not all examples found elsewhere are applicable.

4 Literature on proposal evaluation approaches

The academic literature on proposal evaluation for research and innovation funding is mostly focused on the role of peer review in making funding decisions for academic research grants. The issues addressed in the literature are typically questions whether peer review is the most effective way of awarding grants (e.g. does it allow to select the 'best' science, or the most innovative science), whether its processes are fair and transparent and whether the burden on proposers, reviewers and funding agencies outweigh the benefits (efficiency). In general, while there are indeed critical points to be made on peer review as a key component of research funding decisions, there is consensus amongst research funders across the world, that it is still preferred method of prioritising and selecting research proposals.

Two comprehensive literature reviews made by RAND in 2009[11] and an update in 2018[12] for the synthesise the key issues regarding proposal evaluation. Although commissioned specifically by the health research domain, the two literature reviews cover general non-domain related literature and look particularly at effectiveness and the burden of peer review for grant funding. The question whether peer review selects the 'best' science is difficult to answer as

[11] Ismail, S,, A. Farrands and S. Wooding, (2009), Peer review in the Health Sciences, A review of the literature, RAND Corporation, Cambridge, UK.
[12] Guthrie S., I. Ghigha, S. Wooding, (forthcoming 2018), What do we know about grant peer review in the health sciences, RAND Corporation, Cambridge, UK.

the meaning of 'best' science is not fixed according to the literature reviews. Most of the cited studies in the RAND studies relate to bottom-up research grants for individual researchers, similar to ERC funding. In the rare cases where the performance of peer review is evaluated this is done in academic terms, looking at the bibliometric performance of the grantees in comparison with non-grantees. The RAND review has found no major alternative approach than using some type of peer review in agencies across the world, so there is no evidence that there is a better way of allocating grant funding. Another topic frequently discussed in the research funding debate is the fair chances of interdisciplinary proposals. This is addressed by broadening the composition of panels with experts from more disciplines and by analysing the scoring patterns of individual reviewers (seeking for extreme scoring) to spot controversy between reviewers (rather than looking at average scores). Mission-oriented research and particularly the transitional change type is likely to attract interdisciplinary solutions so this is clearly relevant for future framework programmes.

The 2018 RAND review suggests that there has been some experimentation using lotteries within certain sets of applicants (e.g. the group of applicants scoring just under a set percentage of the best applicants). But these experimentations are applied to bottom-up individual research grants such as fellowships and by Foundations who are more at liberty to experiment with different forms of grant allocations. Indeed in a short article in Nature, Thomas Sinkjear, former director of the Danish Research Foundation suggests that experimentation with blind evaluation at a number of Foundations led researchers to propose much bolder proposals than they would have dared to submit if the proposals were not anonymous. In addition, reviewers were allowed to give golden tickets to proposals to get funded even though their cumulative score would not position them to get funded. The experiments have been conducted recently so there is little evidence of their effects as yet. However, in these cases of bottom-up individual researcher grants blinded evaluation could help funding more unorthodox proposals.[13] The Stakeholder Workshop (see Appendix II) however clearly concluded that for objective-driven/ mission-oriented research blind evaluation is not suitable. And indeed, no examples were found of this being used in the Agencies reviewed.

The Wellcome Trust in the UK is another Foundation that in July 2018 launched a new Leap Fund to support high risk 'left-field' ideas in health and life sciences that need considerable funding, taking inspiration from venture capital sectors.

[13] Thomas Sinkjear, Fund ideas, not pedigree to find fresh insight, Nature, 8 March 2018, volume 555.

For the purpose of this study for the future European Framework Programmes it is interesting to note that in broad lines the evaluation process design as described in the literature is quite similar to what is used by the Commission and its agencies. There are single and multi-stage reviews as well as single and multi-stage proposal submissions in place. The differences are in the details of the design.

Compared to the attention and focus on prospective evaluation of academic research proposals, there is very little literature on the selection processes and methods in innovation agencies. The proposals for innovation programmes, which support projects that are closer to the market, add evaluation dimensions such as commercial potential and plans, financial capacity of the proposers, access to markets.

In 2016, TAFTIE, the European Network of Innovation Agencies made a review of the project selection processes of a large part of its members.[14] The review came to a number of conclusions:

- given the differences in regulatory, budgetary and governance between European agencies and their different functions, there is no sense in defining 'best practice', for selection processes. The review did define some 'good practices';
- quite a number of Agencies use two-stage proposal submission where the first proposal is short;
- Innovation Agencies mostly use a mix of internal and external experts. Single company grants are more likely to be assessed internally, while collaborative grants are more likely to include external evaluators;
- collaborative programmes are more likely to have a panel (e.g. selection committee, expert committee) that make a funding recommendation;
- there is a trade-off between the efficiency and the effectiveness of selection processes: selection processes with a higher reliability (e.g. more eyes looking at a proposal) generally cost more;
- another trade-off is between the evaluation of project excellence against considerations at a systemic level. The latter might lead to portfolio related choices for a better spread of proposals against policy objectives. The TAFTIE review found only one Agency (Research Council Norway) that has an explicit step of portfolio management in its

[14] TAFTIE, (2016), Task Force SELECT, Final, May 12, 2016, http://www.taftie.org/content/task-force-select

evaluation procedure, executed by its Programme Board, that makes the final funding decision; and

- two Agencies, Vinnova (SE) and the Spanish Centre for the Development of Industrial Technology (CDTI) experiment with the submission of short video pitches of their proposal, along with their written proposals. Evaluators also use these videos in their (remote) assessment.

The TAFTIE exercise also showed that the evaluation processes and criteria have strong similarities with what is seen as good practice in the leading agencies:

- either a single of two-stage submission of proposals;
- the use of independent external reviewers; and
- a scoring system with a wider spectrum of award criteria including quality, impact and implementation.

On the topic of fairness there is no clear consensus in the literature on the benefits versus the drawbacks of using blind peer review. The RAND literature review, reports on two studies that did find a clear difference in scores between blinded and non-blinded reviews, one based on the National Science Foundation (NSF) peer reviews in the USA (Bhattacharjee, 2012) and one in South-Korea (Lee et al. 2000).[15] The former article also states that even if proposals are anonymised, it is often still possible to identify the proposers. This point of view was also clearly stated in the Stakeholder Workshop where it was pointed out that previous Framework Programmes have used anonymised (part-)proposals, but the identity (of institutions and people) could still be detected. In addition, it costs time and resources of project officers who have to check each proposal in detail before they are sent to peer reviewers. The literature mostly describes situations were peer have to select individual researchers and could be biased regarding age, length of career and academic recognition, institution that the proposers belongs

[15] Bhattacharjee Y: Science funding. NSF's 'Big Pitch' tests anonymized grant reviews. *Science.* 2012; 336(6084): 969–970 and Lee M, Om K, Koh J: The bias of sighted reviewers in research proposal evaluation: a comparative analysis of blind and open review in Korea. *Scientometrics.* 2000; 48(1): 99–116.

to and gender. In the objective driven R&I funding, a key award criterion is the likeliness that a research project gets implemented. To assess this, knowing the identity and track record of the proposers is essential. In none of the cases of objective-driven research that have been reviewed do the agencies and funders use blind evaluation.

5 Evaluation process designs

5.1 The basic designs of proposal evaluation

In general, all grant proposal evaluation designs have a similar set of key phases:

a) Submissions of proposals
b) Eligibility check phase
c) Review phase
d) Grant decision phase

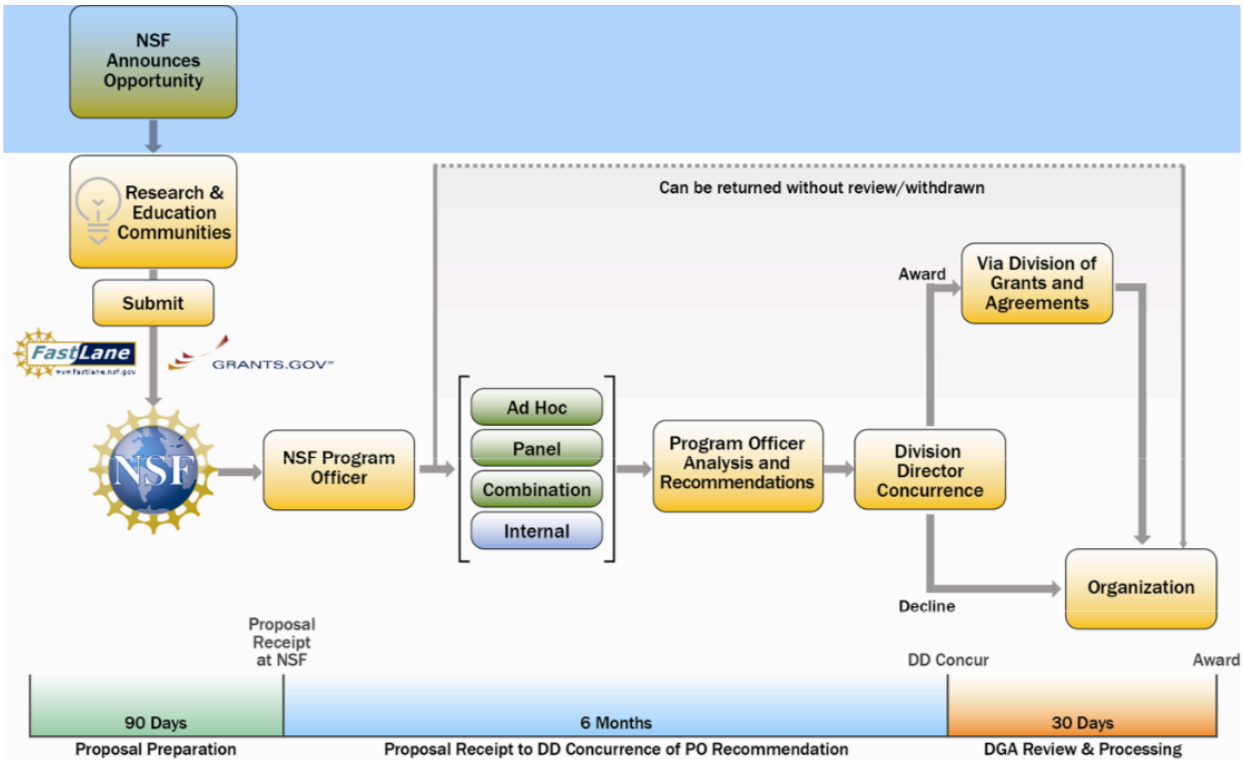
Given the scope of the study eligibility check phases have not been included in the review. The focus has been on the (quality) review aspects of the evaluation process and the grant decision phase. The first key characteristic that differentiates evaluation processes is the number of times participants are asked to submit proposals for one particular call. This can be only once (a one-stage submission), twice (two-stage submission) or more. This affects the design of the evaluation process as quality reviews are needed for all stages. A second key characteristic is the background and position of the evaluators involved in the evaluation process. They can be internal staff of the funding agency or independent external experts not formally attached to the agency, or a mix of both. In addition, the type of expertise that evaluators bring to the proposal assessment process also varies considerably. Figure 3 below illustrates how the entire review process can be organised, in this case by NSF in the USA.

5.1.1 One stage submission

One-stage submissions, where applicants submit only one full proposal is used by most of the European agencies in our review. The advantage of a one stage proposal is that it is easier to organise the quality review phase and thus helps to speed up the time-to-contract. The disadvantage is that the proposers have to spend considerable efforts to write and submit a full proposal. Particularly in cases where there is considerable oversubscription and low success rates, this could become a disincentive to apply for many. The literature reports efforts of Agencies to make proposal submission worthwhile even without being awarded a grant to

counteract this effect. This is usually done through an extensive feedback process. In the case of the Grand Solutions programme in Denmark, proposers who get through to the second round have a face-to-face meeting with the expert panels and reportedly getting direct feedback is considered as beneficial even for those who get a rejection. In the case of the Irish EPA (see below) even large proposals are restricted to 25 pages maximum so the written material has to be succinct.

Figure 3 NSF Proposal and Award Process and Timeline



Source: NSF, 2018

Single submission processes can vary in the number of stages by which all proposals are reviewed for their quality. The table below summarises the main review steps. Models were identified where proposals are first pre-selected internally and then sent to external peers (e.g. IFD) as well as models that are exactly the opposite: peers evaluate the proposals for technical merit and subsequently rated internally for relevance (programmatic fit) and financial feasibility (e.g. NASA). The different models each make sense for their specific programmes. IFD has completely open calls with a broad thematic focus. An internal evaluation first helps to filter out the wacky ideas that have little potential to be implemented. NASA on the other hand has mostly technically detailed calls where an external check is needed to decide on the scientific and technological merits.

Table 1 Examples of single submission proposal evaluation processes

Agency/Programme	Key review steps
Innovation Fund Denmark, Grand Solutions (SE)	• Internal Review selecting top 1/3rd proposals; • Panel review (recommending about 50% of remaining proposals); • Panel Review includes an interview with proposers.
Environmental Protection Agency - Ireland	• Remote scientific experts give scores to all proposals. Those with 60% score and at least 40% score on sub-criteria excellence and impact go to round two; • A national Panel including stakeholders comes to consensus score which proposals get funded.
DARPA -USA	• While proposal abstracts can be informally sent to project officers for comment before the final deadline there is only official submission; • Proposals are in general assessed internally by DARPA project officers and sent to Management Board for approval.
NASA- USA	• The common practice is that full proposals are submitted to a Technical and Programmatic review. The Technical Review is conducted by external scientific peers. They can also be asked to rate relevance; • Subsequently for those that meet the evaluation criteria: a NASA programme officer will evaluate the proposals in terms of programmatic relevance and financial limitations. The Programme Officers then presents a recommendation for selection to a NASA Selection Official identified in the call.
NSF-USA	• Proposals undergo a merit review by external peers and the programme officer and subsequently a programmatic review by the same programme officer.

5.1.2 Two stage submissions

The introduction of two-stage submissions aimed at reducing the burden on participants - the first submission is often a short version of a proposal – as well as on the funding agencies. It is thought that a short proposal would reduce the amount of effort that proposers have to invest in producing the proposal, thus limiting the administrative burden on the users. The 2018 RAND literature did not find conclusive evidence that by reducing the complexity of the application process (e.g. through short proposals) reduced the preparation time of the applicants. The number of comprehensive studies available on this matter is limited.

It seems that many US Agencies use a two stage sub-mission process where the first proposal is a so-called pre-proposal or concept paper. This concept paper is typically max 8-10 pages long. As described above the US Agencies programme managers have much more freedom to adapt the selection procedure for particular initiatives of calls. The NSF Directorate for Biological Sciences introduced a the two-stage procedure for a five-year pilot and simultaneously reduced the submission deadline to once a year, aiming to reduce the burden

on staff and researchers.[16] However, in 2017 the Directorate decided to abolish these changes as it increased the peaks of burden on staff with increasing numbers of pre-proposals. An evaluation showed that the changes did not have the desired positive effects of better full proposals nor lessened the burden on researchers.[17] Subsequently, this Directorate has gone back to Full Proposals only.

The following Table 2 gives an overview of Agencies predominantly using two stage submissions.

Table 2 Examples of double submission proposal evaluation processes

Agency/Programme	Key review steps
NWO-Wetsus – Water Technology Programme (NL)	• A preliminary proposal (2 pages) is submitted and Committee of three assesses both Quality of Science criteria and Added Value/Impact criteria. Wetsus can give non-binding advice on strategic fit of proposals with call themes. Committee decides which ones get to submit full proposal; • Full proposal remotely reviewed by at least additional 3 international reviewers who submit an assessment report. Proposers are allowed to comment on the assessment report; • Committee reviews proposals, assessment reports and comments from proposers on these reports and ranks proposals accordingly. NWO Board to take decision.
ARPA-E -USA	• A common model is that applicants are asked to submit a Concept Paper of about 5-7 pages which are reviewed internally and by external experts. A selection (generally 20%) of the best Concept papers are asked to submit a Full Proposal.
DARPA/NASA/NSF All in USA	• Programme directors can decide to ask for a Concept Paper and subsequently a full proposal. But this is not the dominant model.

5.1.3 Stage-gate proposals with multiple submissions

The case reviews featured two examples, the Strategic Innovation Programme on Energy Conversation Technology of NEDO in Japan and Vinnova's Challenge Driven Innovation Programme in Sweden (see Appendix I for more details on the latter programme), where applicants, if successful, can enter into multiple stages of projects, each with its own proposal evaluation.[18] This starts with a shorter and simpler feasibility or initiation phase. Successful

[16] Jeffrey Mervis in Science, 11 Oct 2017, http://www.sciencemag.org/news/2017/10/nsf-drops-preproposals-deadlines-biologists-seeking-funding
[17] Abt Associates, (2017), Evaluation of the NSF DEB and IOS Switch to Preliminary Proposal Review: Final Report, prepared for NSF, March 2017.
[18] See https://www.vinnova.se/en/publikationer/challenge-driven-innovation/

participants of this stage can - after completion of their initial project - apply to take part in a second stage. This limits the number of applications for the more complicated and extensive second phase and eventually third stage where projects get more complex and receive larger sums of public support.

The Swedish Agency Vinnova responded to its national government's Agenda 2030 for Sustainable Development using the United Nation's SDG as guideline for the direction of the government strategy. Any consortium with at least three project partners who have identified a societal challenge can take part. The research activity is built up in three stages:

- Initiation (up to €50K, max 80% state funding)
- Collaborative project (up to €1 million, max 50% state funding)
- Follow-up investment (up to €2 million, max 40% state funding)

The stage-gate approach allows Vinnova to invite more competitive and risky projects in the initial phase and monitoring whether these projects make progress in their initial stage. Applications are evaluated based on how well they meet the principles of the programme, using the same three criteria in all stages: 1) Potential/Impact, 2) Actors/Consortium and 3) Feasibility/Implementation. However, the weight of the three criteria shifts as successful projects progress towards a new stage.

5.2 Proposal evaluators

With research funding organisations across the world it is generally seen as good practice to invite external proposal reviewers who are considered as peers to the research community whose proposals they evaluate. They are usually experts in a particular science or technology domain or experts as practitioners from the application field of the research topic. The experts score the proposals, a delegation of experts (often a panel) ranks the proposals and recommends the management or Board of the funding agency which projects to fund.

Innovation and Sectoral Policy Agencies on the other hand are more likely to rely mostly on internal expertise, i.e. from policy officers and experts employed by the Agency, possibly alongside external experts. External experts used by Innovation Agencies could be knowledgeable about the technological domain and/or the commercial and financial potential of proposals. The TAFTIE review of Innovation Agencies across Europe concluded that for collaborative research programmes which include industry and university partners adding external reviewers to the process is more common, while programmes that only have business

beneficiaries are mostly reviewed by internal experts only. [19]The availability of high quality experts without a conflict of interest was reported as an issue for many Agencies. Biegelbauer and Palfinger (2016) concluded from a comparative study of ten European innovation agencies that six out of ten use a mix of internal and external evaluators, three used exclusively internal evaluators and only one used exclusively external evaluators.[20]

The Horizon 2020 Rules of Participation state that reviews should be conducted by independent experts, i.e. not formally attached to the Commission or Executive Agencies. So the evaluation tradition in the Commission is mostly based on Research Funding organisations.

One could expect that with a greater mission or societal challenge orientation of research programmes, users and stakeholder communities are invited as experts, assessing the potential impacts of the proposed research projects. In fact, the study found only one example, the Irish Environmental Protection Agency that has systematically engaged users into their research funding evaluation process. These users are typically organisations that need to be involved in the implementation the results of research projects such as the national Water Authority, regional and local governments or the Statistic Office who would have to supply or process data related to some research projects.

None of the example Agencies have an operational **citizen engagement** in the proposal assessment process. Indirectly some civil society representatives could have an influence on the outcome of proposal evaluations when Agencies involve Advisory Boards or Programme Committees to give an input to the portfolio of project proposals that have been ranked as the highest. This depends on the composition of these Boards. The NIH National Advisory Board on Cancer for example consists solely of professors from various health research domains.

The review across agencies world-wide found a variety of combinations of internal and external evaluators active during different phases of the evaluation process. The European Commission already uses a broad definition of experts to evaluate research proposals so it can already involve practitioners and representatives from user organisations in its current set up. However, in the European context as this moment it is not allowed to involve internal experts

[19] TAFTIE (2016), page 8-9.
[20] Biegelbauer, P., T. Palfinger (2016), Auswahlverfahren von Forschungsförderungsorganisationen im internationalen Vergleich, in: fteval Journal for Research and Technology Policy Evaluation, Issue 41, March 2016.

in the selection process. The following Table 3 shows the types of evaluators used in our Agency cases.

Table 3 Models using internal and external evaluators

Type of evaluators	Agencies
Predominantly Internal	DARPA (USA), Vinnova (SE)
Mixed model	NIH (USA), EPA (IRL), IFD (DK), NASA (USA), NSF (USA)
Predominantly External	NWO (NL)

What is apparent from the review is the significant role of internal experts in the selection process both in the US Sector Policy Agencies (e.g. DARPA, ARPA-E, and to a lesser degree NIH) and in the European Innovation Agencies in this review. According to interviewees this is also related to the selection of project officers in these agencies, who are selected on the basis of their scientific and technical backgrounds and thus can be considered as experts themselves. This is particularly the case for the US agencies.

5.3 Other features in assessment processes

The review found a few additional steps in the evaluation process that were used alongside the mainstream components of proposal reviews.

Response to peer review comments

In the evaluation process of the NWO-Wetsus Water Technology Programme, applicants of full proposals get to read the assessment reports by the independent remote reviewers and are asked to provide comments on these reports. The final proposal evaluation Committee takes these comments as one of the inputs for their final ranking, alongside the full proposal and the reviewers assessment reports.

The possibility for applicants to provide written comments on the assessment reports written by remote expert reviewers that serve as an additional input for the final assessment of a proposal.

Face-to-face meetings between applicants and expert reviewers

In the evaluation process of the Danish Grand Solutions Programme applicants of the best one third of all proposals are invited to present their proposal in a face-to-face meeting with the expert panel. The face-to-face meeting helps the reviewers to get to know the people behind the proposals and to seek answers to questions they cannot retrieve from the written proposals. For proposers it is helpful to get immediate feedback from the experts, which is even appreciated by those whose proposals are rejected. IFD managed to organise interviews with 67 proposal teams in a period of two weeks and maintain their 100-days-time-to-contract rule in their most recent call.

The stakeholder consultation meeting in April 2018 suggested that the introduction of face-to-face meetings between applicants and reviewers, to present and explain their proposals and have the possibility for an interactive questions and answers session, would be beneficial for both parties. There are not many examples of this approach in the review. In practical terms, this can only realistically be done if the number of proposals is not very high. In a European wide context this would lead to considerable additional travel (costs) for applicants. This could be unfair to those proposers who are located at greater distance from the review site (mostly Brussels). However, this could be introduced if there are a limited number of proposals in a last stage of the assessment. An example where this could be used is for instance if the launch of a 'Big Mission' is implemented through a small number of bigger and complex projects.

5.4 Final decision making and portfolio management

In almost all cases the final decision for project funding is made by the Agency management, whether this is a single person (e.g. NASA) or a Board of several Agency representatives (e.g. NWO, NSF). Some organisations have set up dedicated Panels/Advisory Boards with internal and external members such as the National Evaluation Panel by EPA-Ireland or the National Cancer Advisory Board connected to NIH, who formally only recommend to the management. These Panels could consist of a broader representation than only the external evaluators and include members of the Agency, stakeholders and so on.

The US Agencies more actively pursue an up-front portfolio approach. DARPA for instance states that its investment strategy begins with a portfolio approach. Before launching a call, Programme officers are expected to conduct a thorough analysis of already existing R&I projects that have received funding (finished and ongoing) to see what specific gap in the research activities their planned call can address. These agencies have extensive databases to

help them with this analysis. The analysis is part of the officer's justification to their management before they get permission to launch a call. One has to keep in mind that in many cases the US Agency calls are much more specifically defined (in terms of expected outputs and technologies to be addressed) and launched in a very decentralised manner. So indeed, it is important to know that there is no overlap with already existing research activities funded by the Agency. In the Commission case such a front-end analysis of research that has already been funded with European (co-)funding (including partnerships) could perhaps be conducted at a particular mission level. This probably requires more sophisticated project databases than the Community Research and Development Information Service (CORDIS) can offer today.

Portfolio management as part of the evaluation process is less visible. This could for instance be done in a wider call where the call's objective requires a certain spread of types of research (e.g. by topic, by risk-profile, by Technology Readiness Level (TRL) position). The review has not come across examples where this was a task of reviewers of individual proposers (e.g. by adding this as an additional question to reviewers). In practice, as some interviewees suggest, this sort of 'back-end' portfolio management is done in decision and recommendation meetings by groups of experts, programme officers. This is when there is a list of (ranked) proposals that have already passed the required qualitative thresholds, which need to be prioritised for funding decision. Overseeing all proposed projects at these high-level structures the match with the mission or programme objectives is assessed again. Negotiations and debates take place whether the ranking as recommended by peers should be maintained or not. These elements of the process are usually not well documented. One interviewee mentioned that if funding allows it the panel would add an additional project that would address a certain gap. Other have noted that they would seek a certain spread of proposals across the Technology Readiness Levels.

6 Evaluation Criteria

Evaluation criteria used across agencies in the world use a similar spectrum that has been in use the last decades for European research funding. They all include variations of criteria related to:

1. The scientific and technological merits of the proposed research project
2. The credibility and background of the proposer(s) of the project
3. The expected outcomes and impacts of the proposed projects
4. The strategic fit of the proposed project with the programme or call objectives
5. The soundness of the execution of the project (or implementation)
6. The financial and other resources foreseen for the proposed project

There is a wide variety in the way agencies define each of these categories in detail and what weighting is given to each of them. There is a general sense that the set of award criteria should be adapted to the objective of a specific call, an initiative or a programme. A majority of agencies that are reviewed have the flexibility to adapt the definition and weighting of the award criteria when necessary. The simplest set of standard award criteria are used by NSF who as a standard only used two criteria: 1) Intellectual Merit (which encompasses the potential to advance knowledge and 2) Broader Impacts which encompasses the potential to benefit society and contribute to the achievement of specific, desired societal outcomes.[21]

It is seen as good practice across all Agencies that the selection criteria are clearly communicated to the applicants as well as to external reviewers. Some agencies have extensive guidelines and videos for reviewers to train them how to interpret the selection criteria and more generally the principles of conducting peer review.

The European Commission is relatively unique in having a standard set of **three award criteria** across all programmes: excellence, impact and quality and efficiency of implementation. Each of these can have sub-criteria and the impact criterion refers to impacts that are mentioned in the work programme. Thus, the broad categories allow for some flexibility in their interpretation.

[21] NSF, Proposal and Award Policies and Procedures Guide, January 29, 2018.

In the literature on research proposal selection two key issues concerning criteria lead the debate:

- whether the criteria (and the evaluation process) used are sufficient to select the most innovative, high-risk projects that could lead to breakthrough science and technology; and
- whether the criteria (and the evaluation process) stimulate multi-disciplinary approaches.

The Commission's current sub-criterion under 'Excellence' explicitly addresses both issues. For Research and Innovation Actions sub-criteria under 'Excellence' are: a) 'extent that the proposed work is beyond the state of the art, and demonstrates innovation potential (e.g. ground-breaking objectives, novel concepts and approaches, new products, services or business and organisational models) and b) Appropriate consideration of interdisciplinary approaches and, where relevant, use of stakeholder knowledge and gender dimension in research and innovation content'.[22] Thus, in terms of award criteria, these considerations discussed in the literature on peer review are both addressed in Horizon 2020.

An emphasis on track record of the proposers is discussed as one reason that could explain a bias against highly innovative proposals. Wooding et al. (2018) found in the literature that innovative proposals may have less preceding work supporting them, and hence receive less praise from reviewers. In addition, innovative proposals from young researchers may suffer a 'double disadvantage': lacking previous work, both because of their novelty and the researcher's shorter track record.[23]

A number of observations can be made from the review of agencies:

a) In almost all cases under review, scientific and/or technological merit are knock out criteria. Obviously, no one has an interest to fund second rate R&I projects. Nevertheless, the detailed definition of merit varies. In the technologically very detailed calls from for instance DARPA and in the US Department of Energy's SunShot mission, technological feasibility is the key issue: can the proposers deliver the desired technical output? In the Swedish Challenge Driven Innovation programme scientific excellence

[22] Horizon 2020 – Work Programme 2018-2020, General Annexes, H. Evaluation Rules
[23] Guthrie et al. (2018).

in the academic sense is not an explicit selection criterion. Nevertheless, the proposed projects need to be lead to radical transformative changes. This could be based on science and technology but also on radical social innovations. The impact on the sustainable development goals is not necessarily dependent on the science involved.

b) Particularly for the sector-policy oriented agencies and the challenge driven programmes, the strategic fit of a project - with either the call, the organisation's mission or even national policies - is given a high prominence as selection criterion. The assessment of this criterion is made both by independent peer reviewers, but even more so by the Agency staff and the Management Boards that take a final decision on which set of proposals get funded.
c) More attention is given to assess proposals to societal impacts that go beyond the direct economic impact of the beneficiaries. The Danish Innovation Fund has developed a Societal Readiness Level (SRL) measurement system that proposers have to fill in as annex to their proposal. This is used to assess potential impacts although not suitable for all proposals. The Irish Environmental Protection Agency has used the Danish SRL model and adapted it for their programmes.
d) Almost all agencies under review can adapt the criteria for specific programmes and even for calls. This is particularly the case for US agencies where programme managers have a great deal of discretion deciding on the proposal evaluation approach. There is flexibility of adapting evaluation criteria to different policy objectives.

7 Lessons for future Framework Programmes

It is clear with the variation of models identified that there is not a clear 'best practice' design and implementation of proposal review processes. The design of the process is strongly influenced by the history and mission of each of the research organisations, the policy context of the research programmes that allocate funding, as well as the objectives of the specific programmes and calls.

For 'narrow' calls with very precisely defined expected outcomes, the evaluation process can be more straightforward than for 'broad' calls which ask for complex transformative solutions that require much more than only excellent science and technology. There is clearly no one-size-fits all. Good practice here is that funders can adapt the design of the evaluation process to fit with the nature of the programme and call.

The models identified range from traditional academic peer review models where proposals are reviewed and ranked by independent 'fellow scientists', to hybrid models where policy needs and commercialisation goals require a mix of internal decision and external expertise, to models where agency officers have the full authority to select proposals in a process they have designed for the purpose of a call. The TAFTIE review of proposal evaluation practices in Innovation Agencies came to a similar conclusion: "A major outcome of the [TAFTIE] Taskforce was the realisation that in hindsight of the differences between the agencies, their regulatory, budgetary and governance environment and the functions they have to fulfil in the respective innovation systems, it does not make sense to define a "best practice" for the selection processes.

Where some of the identified practices differ from how proposal evaluation is done in H2020 are the following characteristics:

- the significant **role of project officers** in the review of individual proposals and **discretion of programme directors** to decide on evaluation processes and criteria per call;
- multi-stage **review** process where all (full) proposals get reviewed in a first phase and only the best go through to a second, often more extensive round of review;
- a **split in the review process** where scientific merit is reviewed in the first instance by scientific peer reviewers and impact in a second stage by a wider set of evaluators;
- the involvement of **user organisations** in evaluation panels as for instance in the Irish EPA; and
- the explicit award criterion of the '**strategic-fit'** of a project with the mission, programme or call used to assess proposals.

Transparency and fairness are important principles for European research funding. Project officers in the Commission and its agencies are not necessarily hired as domain experts in the thematic field that they work. For these reasons copying the US Agency model with a large discretion of project offices and programme directors does not seem a sensible option in the EU context.

A matter that is important to address with objective-oriented research is how and by whom proposals are assessed in relation to their strategic fit with the policy objectives of a specific call. Today this assessed by the independent experts when they score proposals for Impact: *"The extent to which the outputs of the project would contribute to each of the expected impacts mentioned in the work programme under the relevant topic;"* The questions to raise here are whether independent experts are well positioned to answer this question (they need to be able to link foreseen outputs to high level impacts) and whether there isn't an additional way to assess if proposals have a strategic fit with the programme objectives. In the examples that are sketched in the review some Agencies deal with this by adding an additional Panel or Board with a more generalist and user oriented expertise than the individual reviewers. They can assess whether the proposed ranking of the proposals could be changed in favour of projects that have a better fit to the objectives.

With the debates on mission-oriented policies and societal challenge driven R&I in mind the review had a number of hypothesis.

1. *Compared to bottom-up research programming, it can be expected that in objective-driven top-down approaches the balance shifts from 'scientific merit' as the key evaluation criterion to selection criteria related to relevance, impact and 'strategic fit' with policies or missions.*

This hypothesis can be confirmed. As stated above scientific and technological merit are still important criteria and in most cases knock-out criteria. However, additional criteria as feasibility to deliver outputs, relevance to the call and mission, expected societal and economic impacts are on equal footing. It is impossible to quantify the weight of each of these sets of criteria as in many cases there is no quantitative weighting of different criteria and the weighting can vary in multi-stage reviews.

2. *Given the complexity of societal challenges, the target groups 'carrying' the proposals become more complex and multi-disciplinary and include users, parts of value chains and a greater mix of actors.*

This is partly confirmed. Only in the dedicated Challenge driven programmes (such as the Danish and Swedish cases and for a lesser degree in the Irish case) are proposers stimulated to form such complex constellations within the project. The American cases all seem to aim at quite traditional single R&D performers or collaboration of R&D institutions in their calls, with the exception of some health research areas. As the focus of the objective-oriented funding Agencies is mostly on outputs and impacts, the assessment of the people and organisations behind the proposals is very important. Are the proposers likely to deliver the requested technology, solution, cure or system change? is an essential question in the review process. Therefore, attention to the proposers' background, track records, institutions and collaboration agreements is key in the agencies in the review. None of the agencies in the review used blind evaluation to assess objective-driven research and innovation. Citizen involvement has not been implemented in the evaluation processes reviewed.

3. *In objective driven top-down research programmes one could expect that a substantial part of the reviewers are not chosen as scientific/academic specialists.*

This is partly confirmed. Scientific and technology experts still have an important role in the review in a majority of the cases. Programme officials have shown to be directly involved in the review process particularly in the US cases and also in DK, EI and SE. Innovation oriented programmes have reviewers from industry understanding the potential market impacts of proposals. Only one case has been found (EPA- Ireland) where users were actively involved in the entire review process. In this case users were mostly from different public bodies and institutions involved in the implementation of research results. The involvement of citizens or representation of groups of citizens were not found in the examples cases in this study.

4. *To ensure that projects are reviewed as components achieving a bigger mission more 'portfolio review' approaches can be expected, comparing proposals with a view how they fit as a set of complementary solutions to a bigger operational objective.*

This hypothesis has been partly confirmed. The study indeed found that Advisory Groups or Boards conduct the last stage of the review process, assessing whether projects fit with the programme objectives. However, this was done after single project reviews were made assessing proposals by their individual merits. In cases where programme officers have a large discretion in choosing proposals, the calls were usually very narrowly defined and not complex societal missions.

The Stakeholder Meeting[24], discussing the preliminary findings provided fruitful inputs and perspectives on the topic of this review (see Appendix II). The key take-away points for the Commission that came out of this workshop were as follows:

- the three main criteria (Excellence, Impact and Quality and Efficiency of the Implementation) currently used in Commission are suitable for the future. Flexibility in defining and adapting the sub-criteria would be of importance to cater for different types of research, including mission-oriented research;
- the design of the evaluation process (e.g. whether one-stage or two stage) should be in line with the incentives of the call and take into account important imperatives such as time-to-grant and the burden on both applicants and reviewers;
- blind evaluation is not suitable for mission-oriented and top down research where the potential for impact and feasibility of implementation are closely related to the background of the consortium partners;
- recruitment of evaluators should be reviewed to see what measures might enable more people in employment outside the research community to participate; and
- engaging a high level Advisory Board or similar group in a later stage of the evaluation process was considered an interesting option to be used for missions and particularly for portfolio management of research and innovation projects.

The review of proposal evaluation practices has generated a wealth of information. Nevertheless, the examples need to be understood in their specific institutional, policy and national context to judge whether they are indeed good practices. This also implies that the relevance of these examples for a future Framework Programme need to be assessed in the European research and innovation policy context and the specific institutional set up of the Commission. What is a good practice in one country is not necessarily applicable to the Framework Programme. Therefore, the Recommendations in the next Chapter are specifically focused on how European Commission proposal evaluation could be improved taking on board some of the relevant lessons from this review.

[24] Held in Brussels on April 25th 2018 with representatives from various umbrella organisations from science, research and industry including some experts and practitioners on proposal evaluation for research and innovation.

8 Recommendations

The findings of the review lead to the following four recommendations addressed to the Commission to consider for Horizon Europe:

5. Use the **current set of three generic award criteria**, and depending on the objectives of the mission, (work)programme and call, apply more **flexibility in their weighting** and in the **definition of sub-criteria.**

The set of three generic award criteria used in Horizon 2020 – excellence, impact and quality and efficiency of implementation - can in principle cover the key assessment criteria that the Agencies under review also use. With the sub-criteria that can be adjusted, the spectrum of judgments that proposals typically undergo in selection processes can be covered. In comparison with the Agencies under review, the Commission could pay more attention to elaborating what is intended in terms of a strategic fit with the objectives of a call, particularly as independent reviewers chosen for the scientific and technological expertise might not be informed well. Contrary to bottom-up and curiosity driven research where the quality of research (excellence) is the key award criterion, in top down and objective driven research, proposals should not only be of excellent quality, they must have a fit with the policy intention of a call or programme. In order for this to be thoroughly assessed, not only the applicants but also the reviewers should be well informed how projects are intended to contribute to the mission.

6. If missions require a **portfolio** of projects addressing different aspects of the mission's objectives, **combine a sophisticated 'front-end' strategic process** (i.e. fine-tune the call in relation to already funded R&I activities and outcomes) that can identify priorities and gaps allowing to focus calls on completing a portfolio **with a transparent 'back-end' process** where the set of ranked proposals that passed the necessary quality thresholds but still need to be prioritised for funding can be re-arranged to complete a portfolio.

In cases where a mission approach would ask for a set of R&I activities, spread across multiple (potential) projects, **portfolio management** of proposals could be an option. However, particularly as independent proposal reviewers do not have the overview of all proposals and ongoing projects, this is **not a task that should be integrated in the individual proposal quality review process**. In a similar manner as some of the reviewed agencies, those proposals that have passed a certain threshold of scores on the award criteria, could be assessed by a high-level Board or Panel to decide whether the prevailing ranking could be altered/adjusted in

favour of proposals that address a significant gap in the required/desired research activities. Alternatively, if the thematic gap is significant and not time sensitive, additional calls could be launched specifically focussed on that gap. Lessons could be learned from US agencies such as NIH that have the information, IT systems and analytical tools to conduct R&D project portfolio analyses prior to the launch of calls, in order to guide their thematic focus. This also allows the mission managers not only to compare proposals within one call with each other, but also assess the portfolio of ongoing and completed R&I projects.

7. As Horizon Europe is designed for greater impact, develop more sophisticated concepts and tools that applicants should use in their proposals to mark their contribution to the 'pathway of impact'.

Examples in the review such as mapping stakeholders interested in the results of the project, the visualisation through intervention logic maps, user-oriented communication plans and the description of societal readiness levels could be further explored. This should be done in a way that helps independent experts to assess the credibility and likelihood of impact claims, without burdening the applicants with extensive additional socio-economic analysis.

8. For missions implying transformational changes with a complex set of participants and stakeholders: experiment by **using a stage-gate approach** similar to the Swedish Challenge Driven Innovation programme. This allows for a greater number of competing and high-risk initiation studies to develop -if successful - into larger implementation and demonstration projects using different weightings of award criteria in different stages.

List of Acronyms

ARPA-E	Advanced Research Projects Agency Energy
BAAs	Broad Agency Announcements
BEIS	Department for Business, Energy & Industrial Strategy
CDI	Challenge Driven Innovation
CDTI	Centre for the Development of Industrial Technology
CORDIS	Community Research and Development Information Service
DARPA	Defense Advanced Research Projects Agency
DFI	Danish Innovation Fund
EPA-EI	Environmental Protection Agency Ireland
ERC	European Research Council
FOA	Funding Opportunity Announcement
ISCF	Industry Strategy Challenge Fund
NASA	National Aeronautics and Space Agency
NEDO	New Energy and Industrial Technology Organization
NIH	National Institutes of Health
NSF	National Science Foundation
NWO	Netherlands Organisation for Scientific Research
PM	Programme Manager
R&I	Research and Innovation
TAFTIE	European Network of Innovation Agencies
TRL	Technology Readiness Level
SRL	Societal Readiness Level
UKRI	United Kingdom Research and Innovation
Vinnova	Sweden's Innovation Agency

Appendix I: Detailed Case Descriptions

This appendix has the following detailed case descriptions

1 Innovation Fund Denmark – Grand Solutions Programme
2 Ireland - Environmental Protection Agency
3 The Netherlands NWO Wetsus Programme
4 Sweden – The Challenge Driven Innovation Programme – Vinnova
5 United States – NIH
6 United States – ARPA-E
7 United States - DARPA
8 United States - NASA

1) Innovation Fund Denmark's Grand Solutions Programme

Innovation Fund Denmark's (IFD) purpose is to advance research, development, testing and validation of innovative solutions for the benefit of growth and employment in Denmark as well as to solve societal challenges.[25] IFD therefore invests in projects across the entire value chain, from strategic research to implementation. The Grand Solutions Programme is focused on research based innovation and while the programme is business oriented in a majority of projects universities and research centres play a key role.

Grand Solutions projects are characterised by their high-risk profile and focus on ambitious results with high value creation whether in the form of new knowledge, improvement of/or new processes, systems, products or solutions to societal challenges. The objective is that an investment from IFD should ultimately result in significant value creation in the form of growth, employment or solutions to societal challenges. The projects benefits should clearly extend those of the project grant holders. In most cases a combination of public and private actors aims to solve a societal challenge. It is therefore expected that an entire value chain is involved in an early stage.

The guidelines state that the application form should be filled in as part of a combined effort by all the project partners. It is therefore seen as advantageous if beneficiaries of the results of the project and other key stakeholders are active participants in the development and duration

[25] Innovation Fund Denmark, Guidelines for Grand Solutions 2018

of the project. A beneficiary can be an end user of the innovation, or a public or private investor willing to invest in the project after IFD's involvement has ended. The breadth of Grand Solutions projects means that the number and type of project participants will vary considerably from project to project.

Five call themes are foreseen in 2018 and announced on the Innovation Fund's website:

- Green growth (bio-resources, energy, environment, water, climate, transport and construction)
- Better health and clinical research
- New technological opportunities
- The future society
- Open call

There is no detailed thematic or technological description of what type of project is expected from proposals. Within the boundaries of the four thematic areas the programme is quite 'bottom-up. An important boundary condition that IFD has set for the evaluation process is that time-to-contract is no more than 100 days.

Evaluation Process

IFD uses a one-step application process with a multi-review assessment process. The evaluation process, includes several steps of qualification before IFD's Board makes the final funding decision.[26]

- All applications are screened by the IFD's internal scientific officers. Approximately 1/3 of the most qualified projects are then selected to continue in the evaluation process. If needed external experts are also involved in this phase;
- The remaining proposals are sent to remote scientific and technological evaluators who provide a score to each proposal;
- An expert panel which consists of a combination of internal experts and external experts assesses the remaining proposals;
- A hearing procedure of the applicants that have been qualified for the peer review. This consist of a 15 minutes presentation by the proposers, a 30 minutes discussion with the peer reviewers;

[26] IFD Peer Review Guidelines 2017 and interview with Tore Duvold IFD

- After the hearing the panel then adjusts the scores for each proposal if needed and ranks the proposals which should be invested. This is done by consensus. Approximately half of the proposals that went through peer review have a good chance to get funded;
- The ranking is recommended to the IDF Board; and
- The IFD Board makes the final investment decision. IFD's Board has representatives of Industry, Academia, Trade Unions and Foundations.

IFD and EUREKA have made a partnership in view to optimize the evaluation of IFD calls. IFD use the EUREKA Experts database to find the most appropriate experts for the applications received. Therefore, experts who are part of the EUREKA Technical Experts data- base have been asked for their interest to become a Scientific and Technical Expert for the next IFD call for applications.

The emphasis on value creation requires additional effort from the proposers. They have to state where in the value chain a project is positioned. They have to estimate the time-to-market of the final product or solution and state how the value will be created over time. Proposers have to indicate what implementation, business or sales model they expect will be relevant and optimal. In addition to making a statement on TRLs they have to fill in a Societal Readiness Level statement. This is not relevant to all projects, but particularly for those projects where the risks are not mainly technical but also face societal and maybe ethical barriers. The status of the SRL form needs a lot of additional communication to the applicants. Proposers find it difficult to fill in, but it does raise their awareness that they need to involve the whole chain of actors, also those who are going to implement and or invest in the results of the research project. The SRL is not a formal award criterion.

The hearing stage of the process is considered as a valuable addition to the whole assessment. During the hearing the peer reviewers get a better view of the people behind the proposals, whether all partners are committed and how they address open questions that reviewers could not find in the written proposal. The logistics of these hearings are well planned in advance, so it hardly delays the time-to-contract. Another advantage is that proposers already get feedback from the panel during these sessions so even when they are rejected, the applicants receive good expert guidance.

Portfolio management is done by IDF staff either in the early stages and in the final decision round. This could be to ensure a good mix of TRLs. The IDF expert teams for the different thematic areas can also decide that in view of megatrends in a domain, certain projects are missing. So if IDF takes decisions that are not in line with the expert ranking it is usually to add an additional project.

Evaluation Criteria

The following Table 4 gives an overview of the three main award criteria and the questions that reviewers will address for each proposal that comes to the peer review process.

Table 4 Award criteria for Grand Solutions Programme IFD

Criteria	**Questions for the review**
Quality of research and innovation	• Clear objectives in accordance with the call; • Clear well-argued description of the needs and challenges addressed by the project; • State-of-the-art of research and innovation and quality of the question(s) investigated by the project as well as the quality of the solution and its method(s) in the project; • The research and innovation excellence in terms of originality, project use or combination of new methods as well as the ability to create synergy among the project participants; • Partner composition and inclusion of relevant competencies and qualifications that together contribute to the idea, solution and value creation of the project, including potential involvement of relevant international partners.
Value creation	• The relevance of the project for growth and employment in Denmark created through knowledge, innovation and technology; • How the project benefits Danish society and/or contributes to solving tangible societal challenges; • The value of the project for society as well as beneficiaries in relation to state-of-the-art in the short, medium and long term; • How the project seeks to accommodate an unmet or latent need or to explore an opportunity; • The value creation for the project partners in the short, medium and long term, including commercial potential if relevant; • The project has one (or more) beneficiary(ies) or particularly important stakeholders involved in the project; • The readiness of the market/society in terms of the solution(s) of the project, including the actions undertaken to ensure a match between project solution(s) and market/society; • Link of the project to national or international strategies and Danish positions of strength • The fit of the project with the general IFD strategy and/or the specific strategy in the field • The project results may lead to innovation in other fields;
Efficiency and Implementation	• A clear, shared vision of the idea of the project and value creation by project partners; • Description of actions undertaken to ensure the partners' ownership of the project; • Feasibility and organisation: o i) Detailed risk assessment and risk management o ii) Interdependencies between work packages and results, incl. milestones and stop/go-criteria o iii) Efficient decision-making processes in the project o iv) Any legal, ethical and regulatory aspects • Economic efficiency, including resource optimisation in the project and "financial gearing", e.g. other grants and/or investments which could be generated after the end of the project;

Criteria	Questions for the review
	• The commitment by public and private companies including co financing of the project.

2) Ireland - Environmental Protection Agency

The Irish Enterprise Protection Agency (EPA) is responsible for environmental protection and policing. The Agencies' tasks are issuing licensing and permitting, enforcement of policies, monitoring and assessment and research and education. For its mandate for research, the EPA manages both transnational research funding tasks (such as JPI Water) and national research funding tasks. EPA's three research pillars are climate, water and sustainability. National calls are announced on the EPA's website and opened once a year. The 2017 national research call was for fellowships, for desk studies, for medium-scale projects and large-scale projects. The calls are defined top-down. The call themes can be broadly defined as well as very specific (e.g. potential for the use of drones for water sampling).

A key feature of the EPA research funding is that users and stakeholders are involved at all stages: they have a role in defining the research topics and selecting the priorities for Ireland. They are part of the proposal selection process (see below) and (depending on the project size) take part in the project monitoring by participation in the Steering Committee. Across the three research pillars in EDA quarterly coordination meetings are held to discuss the topics for research calls. The EDA also receives suggestions for topics from the research community and stakeholder organisations.

While there are calls per pillar of research, there is one central Applicants Guide that explains the procedures and the evaluation criteria.

Evaluation Process

The EPA invites research proposals through its website by means of an annual call for all its research pillars.[27] Each pillar identifies a number of topics each with a different type of grant (i.e. small, medium sized and large studies, capability development projects, scholarships). There is only one single proposal submission required, with the largest project type allowed a maximum of 25 pages and scholarships 15 pages.

The evaluation process is divided in two stages. First proposals go to scientific experts (minimum of three of which two are not from Ireland, but for multi-disciplinary proposals this can be more than three) who fit with the proposal's thematic area. They assess the scientific merits (50%) as well as fit to Irish priorities and ability to identify emerging topics for Ireland (50%). Proposals with a total of 60% mark get through to round two, if they have at least a 40% score on each criterion.

In the second stage of assessment remaining proposals are assessed by a national panel which consists of experts from EPA and (policy) stakeholders (e.g. Irish Water Board, Statistics Office if proposers need access to data, relevant government departments, regional authorities, other Environmental Authorities, etc) and other users. This panel – typically with 10-20 people - comes to a consensus view which proposals get funded. Typically for each call topic only one project gets selected. A reserve list is created if there is still funding available.

Evaluation Criteria

As described in the paragraph above, in the first review round external scientific experts rate the proposals for scientific excellence (50%) and potential impact for Ireland (50%). The National Panel looks at a wider set of criteria in the second round:

- Relevance to national environmental priorities, including relevance to technical description
- Balance of support across research topics
- Value for money
- Research integrity
- Gender balance
- Capacity development in Irish research community
- Available level of funding
- Performance on previous EPA funded projects (where applicable) including reporting requirements

[27] https://www.epa.ie/researchandeducation/research/epafunding/

The results from the scientific review (round 1) accounts for approximately 35% of the weighting in the assessment, while relevance counts for 25% and value for money 20%. Separate scores are given for the communication plan (5%) of the proposal which needs to include a listing of key stakeholders and the key messages for users that the proposers are foreseeing at the end of their project. Research integrity, gender balance and capacity development each get another 5% of the score. There is currently a debate in EPA to decrease the weighting of the scientific review and increase the score for relevance in the second round.

As relevance is an important criterion, applicants have to:

- Explain how the projects will help policy or provide solutions to emerging issues;
- Make a clear impact statement in their proposal;
- Fill in their assessment of the Technology Readiness Level of the project;
- Fill in their assessment of the Societal Readiness Level of the project; and
- Develop a communication plan and provide a realistic budget for this.

The Societal Readiness Level format is copied from the Innovation Fund Denmark. Proposers are required to elaborate that in their proposal and remote evaluators have been explained how to assess this. It is considered as more useful than the traditional TRLs.

3) The NWO-Wetsus programme in The Netherlands

The Netherlands Organisation for Scientific Research (NWO) is the research funding agency mainly focused on public research organisations. NWO funding covers all scientific disciplines and fields of research. While a considerable part of the funding is bottom-up and aimed at talent development, a smaller part of the NWO programmes has a thematic focus and is aligned to various societal challenges or to one of the nine Top Sectors identified by the Dutch government. One of these thematic and more top-down funding schemes is the Partnership-Programme NWO-Wetsus, on the Impact of Water Research on Energy, Industry, Health, Economy and Environment.[28] Consortia consisting of one or more public research institutions and at least one private sector company from the water sector, and possibly other partners, can

[28] https://www.nwo.nl/en/funding/our-funding-instruments/alw/sustainable-water-technology/sustainable-water-technology.html last visited on 14 June 2018.

propose innovative research projects in the field of water technology. The partnership is with Wetsus, a European Centre of Excellence for Sustainable Water Technology which is funded by national, regional and local research and innovation support and one of the partners of the Dutch Top Sector Water and Maritime. The research that is proposed should be conducted at the laboratories of the Wetsus centre. The calls are once a year

Evaluation process

The process used for this programme is a double submission and multiple review rounds.

The first round of submission asks for a Preliminary Research Proposal of about 2 pages. A three person panel of independent reviewers looks at the scientific quality (set A criteria) and the potential value (set B criteria) of each short proposal (see below). The Chair of the panel is a Dutch professor with an overview of the field, while the other two are foreign professors. Wetsus, the Centre of Excellence on Water Technology that will host the proposed research activities in its laboratories has an observer role in the evaluation process. In an early stage of the assessment Wetsus can indicate whether the proposed research projects have a strategic fit with the thematic areas of the call and can indeed be executed at their laboratory. It can provide a non-binding advice to the panel to include or exclude proposals. The panel decides who to invite for a full proposal. In the very first NWO-Wetsus call in 2017 no proposal got excluded (for not having a strategic fit) and all preliminary proposals were asked to submit a full proposal.

In stage 2 with the full proposals, at least three additional independent and foreign reviewers provide a written (remote) assessment of the proposals. Each proposer has a chance to give a written reaction on the assessment by the remote reviewers. The Panel subsequently forms a final score on the basis of the proposals, the evaluator's written reviews and the reactions of the proposers. The proposals have to have a 'very good score on both sets A and B criteria. The proposals are ranked and subsequently provided to the NWO with a recommendation which ones to fund.

Evaluation criteria

The review uses two sets of criteria: set A on scientific quality and set B related to economic and societal value and strategic fit.

Set A. Originality and scientific quality

- Originality and innovative nature, potential for scientific progress;
- Scientific quality of proposal: objectives, approach and methods, fitness for purpose and feasibility;

- Scientific quality of the group: national and international embedding, publications, expertise.

Set B. Value in economic or societal terms and strategic fit

- Urgency of the proposed research in terms of the scientific reinforcement of the themes described in the call for proposals;
- The added value and potential for practical application of the envisaged research results in economic and/or social terms;
- Extent to which the application is within the aim and themes of this call for proposals

Each proposal must score a 'very good' on both sets.

4) The Challenge Driven Innovation Programme -Vinnova in Sweden

Vinnova is Sweden's Government Agency for Innovation with a vision to strengthen Sweden as a country of research and innovation. The Agency supports R&I collaboration between companies, universities and other higher education institutions, civil society and other actors. Vinnova's Challenge-Driven Innovation (CDI) Programme is launched in 2011 and has the vision to fund projects of international eminence and develop sustainable solutions to tackle key societal challenges.

In 2018 Vinnova changed direction with the programme and rather than a limited set of societal goals, that were identified as 'good for Sweden', the programme opened up to a list of societal challenges set in the Swedish Government's 2030 Agenda (following the United Nations Sustainable Development Goals).[29] The programme does not prescribe whether it is for universities or firms, but it does have an innovation objective and it is expected that problem owners and users take part in the project.

Any consortium with at least three project partners who have identified a societal challenge can take part. The research activity is built up in three stages:

- Initiation (up to €50K, max 80% state funding);
- Collaborative project (up to €1 million, max 50% state funding);
- Follow-up investment (up to €2 million, max 40% state funding).

[29] See: https://www.vinnova.se/en/publikationer/challenge-driven-innovation/

The principles to get project funding from CDI are[30]:

- Based on societal challenges that are defined bottom-up;
- Have a systemic approach, a holistic view of what is required to solve the societal challenge;
- Based on a profound cross-sector collaboration between all actors needed for successful implementation;
- Result in innovations that combine international business potential and societal benefits;
- Builds on active involvement of users and customers;
- Contribute to gender equality through the projects results and implementation; and
- The project's impact logic visualizes the desired results and how to reach them.

The evaluation processes are specifically defined for the CDI programme, so there is no central proposal evaluation process across Vinnova.

Evaluation Process

Proposals are invited by calls for proposals advertised on the Agencies website. For each stage one proposal can be submitted, but applicants are only allowed to submit proposals in Stage 2 if they have completed their Stage 1 project and likewise with Stage 3. For each call an evaluation panel is created. The proposals are primarily evaluated by a team of Vinnova administrators with backgrounds in industry, academia and the public sector, both experts and generalists. In some cases external experts are hired to contribute to evaluation panel. Applicants can be asked for an interview in the evaluation panel meeting.

Evaluations are all competitive, there are no predefined quota for certain topics. The results of the team's evaluations will be presented to Vinnova's Challenge-Driven Innovation Steering Group for discussion. The final decision on each proposal is taken by Vinnova.

Evaluation Criteria

[30] Andreas Netz, CDI Programme Director at Vinnova, presentation Stakeholder Meeting, 25 April, Brussels

Applications are evaluated based on how well they meet the principles of the programme, using the same three criteria in all stages: 1) Potential/Impact, 2) Actors/Consortium and 3) Feasibility/Implementation.

Thus, the first criterion is the potential to help create impactful and innovative solutions to a societal challenge, in keeping with the programme's principles and the global goals of the 2030 Agenda. In Stage 1 *Initiation* proposals this criterion has the highest weight. As the projects are small in size, Vinnova has room to experiment and also choose high-risk projects. The applicants are expected to deliver a visualisation of a logic framework with expected impacts, in the short, medium and long term. The credibility of the ideas on how a larger scale project could eventually create these impacts decides which proposals get chosen. Thus 'excellence in proposed impact' is the main decision criterion in this first stage. Those that have concluded their Initiation stage project can apply for a Stage 2 collaborative R&D project. It is expected that the consortium of a Stage 2 *Developing Solutions* proposal will have more partners, and particularly partners that are potential users and/or from civil society. Thus, in this stage the second criterion 'Actors' has the most weight in the selection procedure. The question whether the consortium has the right actors who can be held responsible for the project's performance and future implementation is leading for the evaluators. Those who have completed Stage 2 can apply for Stage 3 of follow-up investment. Proposals for this stage are again assessed according to the three criteria, but in this stage feasibility and implementation have the most weight. Again, it is expected that additional partners and investors have joined the project to ensure the implementation of the project's outputs. And along the project's development changes in the plans are allowed if it improves the foreseen outcome.

Thus along the three stages the projects are assessed in terms of how they progress and whether their potential for impact still holds until the final investment.

5) National Institutes of Health -USA

The National Institutes of Health (NIH) is not simply an agency that funds research. The umbrella organisation consists of 27 independent research institutes, of which 24 centres have research funding initiatives. NIH's formally stated mission is *to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to*

enhance health, lengthen life, and reduce illness and disability. Each NIH Institute has its own specific mission.[31]

The NIH Cancer Moonshot is often referred to as an example of a Big Mission that could be an example for a European approach. When the initiative was launched the National Cancer Institute (NCI) was charged with implementing the initiative. NCI convened the Blue Ribbon Panel (BRP) in 2016 to provide recommendations for achieving the Cancer Moonshot's ambitious goal of making a decade's worth of progress in cancer research in 5 years, now called the Beau Biden Cancer MoonshotSM Initiative. The BRP was charged with assessing the state-of-the-science in specific areas and identifying major research opportunities that could uniquely benefit from the support of the Cancer Moonshot and could lead to significant advances in our understanding of cancer and in how to intervene in its initiation and progression. The recommendations focused on areas in which a coordinated effort could profoundly accelerate the pace of progress in the fight against cancer and were not intended to replace existing cancer programs, initiatives, and policies already underway. The BRP final report was approved by the National Cancer Advisory Board and included a recommendation for a strategic investment in the development and testing of implementation strategies to significantly improve CRC screening, follow-up, and referral-to-care, particularly among underserved populations. The 21st Century Cures Act was signed into law in December 2016 dedicating new funds to support efforts associated with the Beau Biden Cancer MoonshotSM Initiative.[32]

With this initiative NIH received an additional $670 million to spend which did not come out of the normal NIH budget. As it was supported by Congress and the White House the research money is earmarked and it was up to the National Cancer institute to implement. In terms of day-today management of this additional funding, the Cancer Moonshot is 'business-as-usual'. The research funding is spent through the mainstream mechanisms used for any other NCI research funding. Thus, proposal selection is also not specially designed for the Cancer Moonshot but follows normal NCI procedures.

Within a certain thematic domain project officers have the mandate to design a call or Funding Opportunity Announcement (FOA) in their thematic area. They do so on the basis of extensive

[31] https://grants.nih.gov/grants/oer.htm
[32] From: https://grants.nih.gov/grants/guide/rfa-files/RFA-CA-17-039.html#_Section_V._Application

research on scientific trends and issues and on extensive database research on already existing and ongoing research projects under the NIH umbrella.

Project officers have a choice of launching between over 20 different types of research grants varying from smaller individual grants to large projects and institutional funding.[33] The project officer's proposal has to be approved by the Institute's Board. The proposal evaluation process is quite similar for each call and the main set of selection criteria as well. However, for each FOA the project officer can decide to add specific selection criteria or steps in the procedure.

Evaluation Process

The NIH is a complex organisation with many independent centres and over 20 different instruments that can be applied for each call as aforementioned. Thus, there is no fixed NIH evaluation process, there are a number of variations, possible dependent on the decision of the programme managers. Applicants have to download the Funding Opportunity Announcement to find out what is requested from them and to see if a call has a number of specific criteria and requirements on top of the standard procedures.[34]

The **first level of review** is carried out by a Scientific Review Group (SRG) composed primarily of non-federal scientists who have expertise in relevant scientific disciplines and current research areas. This is done on the basis of the standard procedures unless a call uses different procedures.[35]

The **second level of review** is performed by Institute and Center (IC) National Advisory Councils or Boards. Councils are composed of both scientific and public representatives chosen for their expertise, interest, or activity in matters related to health and disease.

The Advisory Council/Board of the potential awarding Institute/Center performs the second level of review. Advisory Councils/Boards are composed of scientists from the extramural research community and public representatives. Members are chosen by the respective IC and are approved by the Department of Health and Human Services. For certain committees, members are appointed by the President of the United States.

[33] See: https://grants.nih.gov/grants/funding/funding_program.htm
[34] See: https://grants.nih.gov/grants/peer-review.htm
[35] See https://grants.nih.gov/grants/guide/notice-files/not-od-09-024.html

The recommendation process is as follows:

- NIH program staff members examine applications and consider the overall impact scores given during the peer review process, percentile rankings (if applicable) and the summary statements in light of the Institute/Centre's priorities;
- Program staff provide a grant-funding plan to the Advisory Board/Council. Council members have access to applications and summary statements pending funding for that IC in that council round;
- Council members conduct a Special Council Review of grant applications from investigators who currently receive $1 million or more in direct costs of NIH funding to support Research Project Grants. This additional review is to determine if additional funds should be provided to already well-supported investigators and does not represent a cap on NIH funding;
- The Advisory Council/Board also considers the Institute/Centre's goals and needs and advises the Institute/Centre director concerning funding decisions; and
- The Institute/Centre director makes final funding decisions based on staff and Advisory Council/Board advice.

Only applications that are recommended for approval by both the SRG and the Advisory Council may be recommended for funding. Final funding decisions are made by the IC Directors.

The NIH utilizes a 9-point rating scale (1 = exceptional; 9 = poor) for all applications; the same scale is used for overall impact scores and for criterion scores (see Scoring Guidance). Before the SRG meeting, each reviewer assigned to an application gives a separate score for each of (at least) five review criteria (i.e., Significance, Investigator(s), Innovation, Approach, and Environment for research grants and cooperative agreements; (see Review Criteria at a Glance). For all applications the individual scores of the assigned reviewers and discussant(s) for these criteria are reported to the applicant.

In addition, each reviewer assigned to an application gives a preliminary overall impact score for that application. This impact score is not necessarily the average of the scores for each specific criterion. In many review meetings, the preliminary scores are used to determine which applications will be discussed in full at the meeting. For each application that is discussed at the meeting, a final impact score is given by each eligible committee member (without conflicts of interest) including the assigned reviewers. Each member's score reflects his/her evaluation of the overall impact that the project is likely to have on the research field(s) involved.

The final overall impact score for each discussed application is determined by calculating the mean score from all the eligible members' final impact scores and multiplying the average by

10; the final overall impact score is reported on the summary statement. Thus, the final overall impact scores range from 10 (high impact) through 90 (low impact). Numerical impact scores are not reported for applications that are not discussed (ND), which may be reported as ++ on the face page of the summary statement and typically rank in the bottom half of the applications.

Evaluation Criteria

The standard award criteria used for NIH research proposals are the following[36]:

- Significance
- Investigator(s)
- Innovation
- Approach
- Environment

However, for each call project officers can decide to add specific criteria. As applicable for the project proposed, reviewers will evaluate the following additional items while determining scientific and technical merit and in providing an overall impact score, but will not give separate scores for these items.

- Protections for Human Subjects
- Inclusion of Women, Minorities, and Children
- Vertebrate Animals
- Biohazards
- Resubmission/ Renewal/Revision

6) ARPA-E (USA)

The Advanced Research Projects Agency-Energy (ARPA-E) is the US agency tasked with promoting and funding research and development of advanced energy technologies. It advances high-potential, high-impact energy technologies that are too early for private-sector investment. 37It's budget in 2017 was $306 million.

[36] See NIH Review Criteria at a Glance last updated March 9 2018.
[37] See https://arpa-e.energy.gov/

The proposal evaluation processes have been recently reviewed during an overall Assessment of ARPA-E.[38] For the report on world wide agencies this accurate account has been very useful. The National Academies report has described the entire process with its criteria as follows:

Applicants submit a five- to seven-page concept paper. External experts review an abstract of the proposal. This initial stage is intended to save applicants the time and cost of preparing applications with little likelihood of success. Concept papers were not widely used within the Department of Energy at the inception of ARPA-E. This step is widely viewed as a positive addition to the process and is critical in allowing ARPA-E program directors to shape projects. The reviewers assign each concept paper numerical scores according to two broad sets of criteria:

- **Impact of the proposed technology (50 percent)**—Reviewers consider (1)the extent to which the idea has potential for transformative and disruptive, rather than incremental, advancement over current technologies; (2)the extent to which the idea will positively impact at least one of the mission areas set forth by ARPA-E in the relevant FOA; and (3)the extent to which the applicant demonstrates awareness of competing commercial and emerging technologies and how the idea could result in significant improvements;
- **Overall scientific and technical merit (50 percent)**—Reviewers consider (1) the feasibility of the proposed work, (2) the extent to which the applicant proposes a "sound technical approach" superior to alternatives, (3) the extent to which outcomes and deliverables are clearly defined, (4) the extent to which the applicant identifies relevant techno-economic challenges, and (5) the demonstrated capabilities of the performers.

Following this review, the program director considers the evaluations and compiles a tentative list of concept papers to encourage to submit full applications based on a number of factors, including the numerical scores and comments received. Each FOA is accompanied by a Merit Review Plan, which is executed by an ARPA-E Merit Review Board. In carrying out the Merit Review Plan, a Merit Review Board, usually chaired by the program director who proposed the project and will oversee it, reviews and discusses the lists of papers with the program director, and the list is finalized through consensus.

[38] Committee on Evaluation of the Advanced Research projects Agency-Energy, (2017), National Academy of Sciences, The National Academies Press, Washington DC, pages 37-39.

The ARPA-E director reviews the Board's recommendations and makes the final decision as to which applicants should be encouraged to submit a full application and which should be discouraged from doing so.

Following review of concept papers, ARPA-E encourages some applicants to submit *full applications* providing information about prior, current, and pending sources of funding (both private and public) and an explanation of why other funding sources will not support the project. Full applications are reviewed by ARPA-E program directors and other personnel, including the director, along with leading external experts in the field.

Reviewers are asked to evaluate the applications and assign numerical scores, and provide comments according to four broad sets of criteria listed below, two of which—impact of the proposed technology and scientific/technical merit— are carried forward from the concept paper stage:

- **Impact of the proposed technology relative to the current state of the art (30 percent)**—Applicants are asked to demonstrate (1) the extent to which the idea is potentially transformative and disruptive (not incremental); (2) "profound understanding of the current state of the art" and an innovative approach to significantly improving it; (3) awareness of competing commercial and emerging technologies and how the idea would yield significant improvements over these technologies; and (4)a "reasonable and effective strategy" for transitioning the new technology from the laboratory to commercial deployment;
- **Overall scientific and technical merit(30 percent)**—Applicants must demonstrate (1) the extent to which the idea is "unique and innovative"; (2) clearly defined outcomes and deliverables; (3) the feasibility of the proposed work based on preliminary data, background information, and/or sound scientific and engineering practices and principles; (4) a sound technical approach, including "appropriately defined technical tasks"; and (5) identification of major technical R&D risks and feasible, effective risk mitigation strategies;
- **Qualifications, experience, and capability (30 percent)**— Applicants must demonstrate (1) the skill and expertise necessary to execute their project plan, based on prior experience with R&D of similar risk and complexity; and (2) access to the equipment and facilities necessary to execute the plan;
- **Management plan (10 percent)**—Applicants must demonstrate (1) a plausible plan for managing people and resources; (2) allocation of appropriate levels of effort and resources to tasks; (3) a reasonable project schedule, including major milestones; and (4)the appropriateness of the budget relative to the task.

Once external evaluations have been completed, the Merit Review Board, usually chaired by the program director who initiated the relevant FOA, reviews the evaluations, generates "discuss" lists, and meets in person with the review panel to discuss applications. Applicants are allowed to read reviewers' comments and submit rebuttal comments before the Merit Review Board makes funding recommendations. The Merit Review Board then considers the evaluations, any rebuttal comments from applicants, and the technical merit review criteria and program policy considerations and makes its recommendation to the ARPA-E "selection official" (the ARPA-E director). The selection official decides which full applications will be selected to negotiate terms of an award. Again, there is no obligation to use the scoring by the reviewers as the sole factor in determining which projects are selected, or to rank applications according their reviewer scores.[39]

7) DARPA (USA)

The Defense Advanced Research Projects Agency (DARPA) has for sixty years held to a singular mission: to make pivotal investments in breakthrough technologies for national security. It repeatedly states that the organisation is seeking for high risk projects with a potential for great impact.

DARPA comprises approximately 220 government employees in six technical offices, including nearly 100 programme managers, who together oversee about 250 research and development programmes.[40] At the Agency level, the DARPA Director and Deputy Director approve each new program and review ongoing programs, while setting Agency-wide priorities and ensuring a balanced investment portfolio. Programme managers are hired as top level experts in their scientific and technological field. Once they are recruited they have a considerable freedom to run their programmes as they see fit.

DARPA publishes funding opportunities primarily by posting Broad Agency Announcements (BAAs) that formally solicit proposals tied to programme-specific areas of research and development. In addition to those opportunities, each DARPA technical office maintains an "office-wide" BAA that offers a mechanism for researchers to reach DARPA with an idea that

[39] Ibid, pages 37-39.
[40] See https://www.darpa.mil/about-us/about-darpa, page version June 2018

may fall outside of DARPA's current priorities, but that the proposer feels could be valuable to national security.

As the Programme Managers (PM) have a large discretion to decide on the design and contents of the calls within agreed programmes, there is no singular DARPA method of evaluating proposals. DARPA has issued a Guide to Broad Agency Announcement and Research Announcements.[41]

Inspecting a number of open calls in early 2018 it was clear that these calls were very technically specified, outlining in detail the required outputs, the timelines some of the technical challenges that had to be addressed. As noted DARPA does allow open sollicitations of proposals. But the programmes that launch the BAAs are very focussed.

Evaluation Process

The common model for proposal evaluation is as follows:

After a call Programme Managers (PM) may allow proposers to submit preproposals before a full proposal. This is also called an Abstract. The PM can decide how this abstract is used, e.g. whether it is evaluated using all criteria or only some criteria. The format and status of the Abstract has to be communicated to applicants beforehand. From reviewing some of the examples of open DARPA calls in early 2018, we could see that abstracts could be sent voluntarily if applicants wished to do so (and receive comments) but it was not a formal first evaluation step. There is no prescribed way to review an abstract, but all conforming abstracts must receive consideration and a written response either encouraging or discouraging submission of a fully proposal.

Typically, prior or just after the call is published a Proposers Day is organised. Here proposers can discuss the requirements with the PM. Briefing materials are made available online (e.g. power points) so that those who have not attended the Proposers Days still can see the information material. The PM can maintain an open dialogue with the applicants before the proposal deadline, although there are guidelines what he or she is allowed to share with the applicants (e.g. no information from a competitor). Obviously for classified calls in the Defense area these guidelines are more strict.

[41] DARPA Guide to Broad Agency Announcements and Research Announcements, November 2016.

For each call the PM brings together a Review Team. As the Guidelines state, the Programme Manager is the lynchpin in the Scientific Review process. Additional members of this panel are a Scientific Review Official and reviewers and subject matter specialists. The PM can choose whether he or she takes a role as reviewer (filling in an individual evaluation form) as well as PM. Reviewers must be government personnel. Subject matter experts could be external, but they only get to see part of the proposal that their technical input is required for. Finally the PM decides which proposals to recommend for funding based on reviewers' evaluations, technical information provided by subject matter experts, and the PM's own independent judgment. The Technical Office Direct keeps oversight of the integrity of the process and the PM's conduct.

Evaluation Criteria

While again the PM has quite some discretion to decide on the details the three main criteria used are:

1. Overall Scientific and Technical Merit
2. Potential Contribution and Relevance to the DARPA Mission
3. Cost Realism

8) NASA (USA)

The National Aeronautical and Space Administration (NASA) is a US agency responsible for the civilian space programme as well as aeronautics and aerospace research. Its stated vision is: *We reach for new heights and reveal the unknown for the benefit of humankind.*[42] In the terminology of NASA a mission is clearly related to their large undertakings in outer space such as the Mission at Jupiter and the International Space Station that takes up a large share of their activities. Nevertheless, NASA also has research funding responsibilities and has calls for research called NASA Funding Announcements. They do not award funding to individuals, only to organisations.

[42] See: https://www.nasa.gov/about/whats_next.html

The proposal evaluation processes at NASA are described in The Guidebook for Proposers Responding to a NASA Funding Announcement (FA) (2017). Each of the four Mission Directorates (Science, Human Exploration and Operations, Aeronautics Research, Space Technology) implement the overall mission of NASA with a wide variety of programmes. The Guidebook is used as a reference, but as with all US Agencies, each Directorate has the freedom to deviate from the standard processes.

Evaluation process

Some FAs require that proposals be submitted using a two-step process. The Step-1 proposal is an abbreviated presentation of the intended research and is submitted by the Authorised Organisational Representative (AOR) by the Step-1 proposal due date. The required contents for the Step-1 proposal will be specified in the FA, including whether or not the Step-1 proposal is reviewed to determine if a proposer will be invited to submit a Step-2 proposal. Though the Step- 1 proposal is a prerequisite for submission of a full Step-2 proposal, it does not obligate the proposer to submit a Step-2 (full) proposal.

All proposals submitted in response to an FA are evaluated in accordance with the same peer review process regardless of the submitting organization, including NASA Centers. All proposals will have an administrative (eligibility), technical/programmatic and a financial review.

NASA always seeks the best possible evaluations by appropriately qualified peers of the proposer who are knowledgeable, though not necessarily specialists, in the objective(s) solicited by the FA. Only those proposals that the reviewers identify as fully meeting the evaluation criteria will be considered for funding. NASA peer review members may also participate in determining the relevance of a proposal to NASA program objectives and the realism and reasonableness of proposed costs. Evaluation is against the state-of-the-art. Review panels are instructed not to compare proposals to each other; any comparative evaluations are conducted by NASA programme personnel.

Following peer evaluation, the Programme Officer will evaluate the competitively rated proposals in the context of the programmatic objectives and financial limitations stated in the FA. The Program Officer will present a recommendation for selection based on the entirety of these factors to the NASA Selection Official identified in the FA. The Selection Official will select proposals as judged against the evaluation criteria, the objectives of the FA, programmatic considerations, and the available financial resources.

Selections are typically announced between 150 days and 220 days after the proposal due date.

Following the technical review and selection process, documents are submitted to the Grant/Award Officer for a review to determine if proposed costs are allowable, allocable and reasonable for the proposed work.

Evaluation Criteria

Unless otherwise specified in the FA, the evaluation criteria considered in evaluating a proposal are 1) its relevance to NASA's objectives, 2) intrinsic merit and 3) its cost. The failure of a proposal to be rated highly in any one of these elements is sufficient cause for the proposal to not be selected. These may be supplemented by specific criteria given in the FA itself.

Evaluation of a proposal's relevance includes the consideration of the potential contribution to NASA's mission as expressed in its most recent NASA strategic plans and the specific objectives and goals given in the FA. If an FA describes the program's relevance to the NASA strategic plans, it is not necessary for proposals to show relevance to NASA's broader goals and objectives but, rather only to demonstrate relevance to the goals and objectives of the specific goals and objectives of the FA. Evaluation of Intrinsic Merit includes the consideration of the following factors, as applicable to the particular proposal:

- The scientific quality of the proposed project, including, but not limited to, the scientific rationale and the expected significance and/or impact of the proposed work;
- Overall technical quality of the proposed work, including, but not limited to, the quality of the management plan and project timeline for carrying out the work and the effectiveness and resilience of the proposed experimental designs, methods, techniques, and approaches for achieving the proposed goals and/or objectives;
- The qualifications, capabilities, and related experience of personnel demonstrated by the proposal (e.g., publications, delivered products, and other measures of productivity and/or expertise) that would affect the likelihood of achieving the objectives;
- Facilities, instruments, equipment and other resources or support systems presented in the proposal that would affect the likelihood of achieving the proposed objectives.

Appendix II: Stakeholder Workshop Report

Stakeholder Workshop on Evaluation Practices

Brussels, 23 April 2018

Purpose of the workshop

The stakeholder workshop commenced with the presentation by Mr Alan Cross of the European Commission welcoming all participants. The participants of the workshop formed a mix of representatives from umbrella organisations from the research and innovation community as well as experts and practitioners dealing with research proposal evaluations.

As it is likely that the next Framework Programme will include a pillar on global challenges, it is important to collect ideas and experiences on how to select the best proposals in the objective-driven, 'top-down' domains. Furthermore, it is envisaged that a number of 'missions' are going to be identified within these domains. A working definition of a mission is an assignment given for a portfolio of R&I actions to achieve measurable results for a given impact within a set timeframe. While the details of the implementation of these mission-type initiatives are still to be defined, it is helpful to consider what proposal evaluation practices could be applied once these missions have been selected and launched.

Alan Cross welcomed input from the participants of the workshop. In particular he asked for comments on the study of worldwide practices, lessons learned in other European contexts and any additional models that could be of interest for future framework programmes. The review together with the stakeholder workshop will help the Commission to consider what proposal evaluation processes and criteria would be most suitable for future Framework Programmes.

Presentation of the review

Ms Patries Boekholt presented the draft results of her review of evaluation practices used by Agencies and other research funders that have 'top-down' and objective-driven programmes and initiatives. Agencies under review include Innovation Denmark (Grand Solutions programme), the Environmental Protection Agency Ireland, NWO-Wetsus (Water programme),

Vinnova (Challenge Driven Innovation programme), UK Government Agencies (Industry Strategy Challenge Fund), several US Agencies (ARPA-E, DARPA, DoE, NIH, NASA, NSF) and some examples from the rest of the world.

She pointed out that the mission-oriented and global challenge driven policy concepts use a vocabulary that is very European. In countries outside Europe the comparators are mostly found in Agencies with sectoral policy objectives such as health, environment or energy. She stressed the importance of understanding the programme or initiative context, to interpret the uses of certain evaluation processes and criteria. For instance, there is a wide spectrum of the granularity of calls to which proposals have to respond. They can vary from very narrowly defined and technically specified calls to very broad and open calls that ask for solutions to societal problems without prescribing how proposers should tackle that. Many US agencies such as DARPA and ARPA-E tend to use the first type of calls while the example of the Challenge Driven Innovation programme in Sweden (see below) is an example that uses very open calls without any prescription of the scientific or technical content of proposals.

The evaluation processes reviewed apply either single submissions (a full proposal is submitted only once) or multiple submissions. In the latter case proposers are asked to submit a short proposal such as a concept paper first. These are reviewed and those with the highest scores are asked to submit a full proposal. This two-stage proposal submission can currently be used in Commission evaluations as well.

The 'expert review' stage in the reviewed cases also vary between a single review process (proposals are reviewed once by a single set of evaluators) to multiple review processes where different set of evaluators look at all or a selection of proposals. An interesting feature that was discussed in the workshop, seen in a number of the Agencies surveyed (e.g. NIH and the Environmental protection Agency Ireland), was the addition of a Board or Advisory Council that reviewed the remaining set of proposals as a last step in a review process, after the scientific and technical peer review (see also below under portfolio management).

Across all agencies, a wide set of selection criteria is used, consisting of variations of scientific and technological merit, potential impact, strategic fit with the policy mission and implementation. While scientific/technological merit is still an important first knock-out criterion, potential impact and strategic fit with the policy objective or mission has considerable weight in the selection process.

There is no clear pattern that mission-oriented research and innovation selection processes include evaluators with a broader background than scientific or technological experts. The review found only a limited number of examples where representatives of users played a significant role in the review process.

In the discussion that followed the finding that programme officers US Agencies have far more discretion of being involved in the whole process was commented on by a number of workshop participants. In the European context of transparency and consistency it was considered that is would be problematic to have one-to-one discussions with programme officers before handing in proposals, as would be the sole responsibility for selecting proposals as is customary in a number of the examples reviewed.

On the question whether alternative models could be included in the review, the suggestion was made to also consider the use of interviews with the proposers in some stage of the review process.

Vinnova's Challenge Driven Innovation

Mr Andreas Netz, Programme Director of Vinnova's Challenge Driven Innovation (CDI) programme, described the set-up and background of the programme and how its projects are selected. CDI's thematic focus is on the complete set of Sustainable Development Goals (SDGs) as defined by the United Nations. Calls are open in thematic terms as long as projects address one of these SDGs and have a potential for systemic transformations. The expectations in terms of the consortium and the expected impacts of the proposed projects are high.

CDI is divided in three consecutive stages and once a project is successful in stage 1, a new proposal can be made for the second and if showing good progress, the third stage. The whole cycle could take a number of years to complete. No project can skip a stage.

In each stage Vinnova's financial contribution increases in absolute terms but decreases as percentage of the total project costs. The same three selection criteria are used for selecting projects in all stages: 1) Potential for Impact, 2) Consortium and Actors and 3) Feasibility and Implementation. In the first stage (with projects receiving less than €50K from Vinnova) the emphasis is on whether the idea is convincing and could potentially have a high impact. In the second stage the projects are selected if they have a convincing set of partners (including users) that can ensure the project can be implemented and in the third stage the emphasis of the project selection is on the feasibility of actually implementing the innovation. Therefore, the whole evaluation process is set up to ensure that the proposed projects have a good chance for impact on addressing a particular challenge.

CDI's peer review is done by a mix of internal (Vinnova) and external experts. With the broadening up of the programme to include all SDGs Vinnova needs a much broader set of evaluators and panels have become bigger. Proposers are interviewed as part of the peer review meetings. If needed, support from other national agencies is brought into the review process.

The CDI example is one where the selection process is geared towards rewarding impact. Scientific excellence is not an objective as such and therefore not a separate evaluation criterion. Workshop participants found the shift in weight of the three criteria across the stages and the involvement of peers from other agencies in the review process very interesting.

Afternoon discussion session

Following the presentations, the afternoon's discussion focused on how the examples presented and the lessons learned could point to improved evaluation practices for future framework programmes and mission-oriented research and innovation in particular.

Evaluation process design

There was consensus that the current design of the evaluation process overall works well despite that having the same criteria for all kinds of interventions hampers in some cases the optimal presentation of the case for funding. However, despite that retention of the basic criteria was considered appropriate, some argued that further adaptation could be appropriate according to the objectives of the particular calls, as well as the type of stakeholders that can be expected to respond.

As for the design of mission calls themselves, it was strongly suggested to take into account the incentives for proposers to apply for this form of grant, which may in turn be reflected in the evaluation system.

During the discussion, both benefits and drawbacks were pointed out of a two-stage proposal submission. This option can already be used for Commission evaluations.

Whereas a short first stage proposal could reduce the efforts of the proposers initially, the entire length of the proposal process and time-to-grant becomes longer.

The opportunity for the evaluators to have an interview with proposers, as has been used for instance in ERC, should be considered as an optional part of the review process for the future, particularly if the number of applications allow for that to be organised without a too heavy impact on the logistics and thus time-to-grant.

There was a clear consensus that in the objective-driven and top-down type of research funding, blind evaluations are not a suitable approach. The arguments used were both operational and conceptual. Experience of using blind evaluation in a previous framework programme (FP5)

showed that even without explicit reference to the proposer's identities, evaluators could in many cases infer who the proposers were from the proposal text. In addition, it is a greater burden on both the proposers (having to ensure that no references are made) and the project officers (checking each proposal in detail before review to delete references to people and institutions). Any accidental reference to an applicant body can make a proposal ineligible. The liability of partners in consortia would also suggest that evaluators need to know whether proposers can be held accountable.

The main conceptual reason that blind evaluation is considered not suitable for top down research is the greater weight on selecting on impact and implementation. Evaluators can only make a credible assessment of these criteria by knowing the proposers, their track record and their ability to deliver. If these aspects are reviewed in a separate part of the proposal, where applicants' identities are revealed, the procedure becomes tortuous. Significantly, in the worldwide review no example was found of Agencies using blind evaluation for objective driven research programmes.

Some participants were not convinced that splitting the proposal evaluation process in two separate parts, one for the scientific and technological merit (or excellence) and one part for impact and implementation would bring any benefits. In FP4 the experience was that proposals with high Science and Technology scores might score lower on Exploitation as they were further away from market and vice vers proposals scoring high in exploitation were not considered visionary enough so that the average of these two scores came out the same for most proposals leaving final selection a challenge. It could also lengthen the process.

There was appreciation of the multiple stage evaluation process, in which a Governing Board, Advisory Council or similar body is responsible for the last step in the proposal selection. They come in action after evaluators have assessed and ranked the individual proposals based on the criteria laid out for a particular call. Examples presented from the review are for instance the National Cancer Advisory Board in the USA or the national review panels used by the Environmental Protection Agency with a broad (multidisciplinary) representation of experts and users. These bodies are particularly involved in assessing whether proposals fit the policy objectives or mission, not so much with assessing scientific merit or the individual ranking.

Evaluation criteria

There was consensus that the set of selection criteria used currently in in H2020 (Excellence, Impact and Quality and Efficiency of Implementation) in principle covers all important aspects for proposal evaluation. It was seen as important to have more detailed definitions of the sub-criteria used. For instance, the precise definition of 'excellence' depends on the specific call

objective and the domain of the call. Flexibility is important, it was argued, either by adjusting the sub-criteria, or by clear reference to the specific call texts (e.g. excellence with respect to the objectives of the call) and related guidance for the experts. The weighting of the criteria should also have flexibility and adapted to the objectives of a call. As missions are more focused on developing solutions, more weight could possibly be given to the implementation and impact criteria.

"Ingenuity" was suggested as a characteristic that we would expect in a solution focused and global challenge oriented proposed project.

In relation to impact, it was suggested rather than implying that the onus is on proposers to provide evidence of impacts that could possibly occur years after the project or given the risky nature of a project, may not predictable at all, instead to ask them to give a convincing argumentation of the 'pathway to impact'. This would mean specifying activities that the project sets out to perform (as budgeted items) to increase the chances of impact. This could be related to stakeholder engagement or interaction with policy makers, all depending on the topic. This would make the assessment of potential impact more coherent and transparent as well as ensuring a commitment from the proposers to devote resources to measures that support impact.

A potential danger signalled by some workshop participants is a cognitive distance between the policy and governance levels that have set the goals and targets of a mission or programme and those who evaluate particular calls. A thorough briefing of evaluators what was considered essential for the assessment of the relevance of the proposal to the policy objectives set out in the call or the goals and targets of a particular mission.

The profile of the evaluators

It was suggested that we should keep in mind that acting as a evaluator should remain interesting and worthwhile to ensure that the Commission can keep attracting high level and good quality evaluators with a broad range of relevant experience. There was an interest in the examples that had a broader definition of 'expert' and involved representatives of user groups or organisations in the proposal review process as well as representatives from other agencies or sectoral policy domains. This could be worth exploring for the future.

The Commission needs to consider practical measures to recruit such people. The current rate of remuneration and commitment required in working time means that anyone working for an organisation that is unable or unwilling to release them for this task is precluded from making themselves available. This includes many people from industry or user organisations who

would otherwise have an interest in a personal capacity b in undertaking the task and whose experience would greatly improve the quality of assessment on innovation potential. To this end shortening the length of proposals so that they can be read and assessed outside normal working hours may be helpful.

Portfolio management

A last topic discussed in the workshop is how to tackle portfolio management that is currently part of the working definition of missions. The review found that portfolio management (or comparing proposals with each other in terms of their synergy and fit with the policy objectives), is not done in the early stage of the review process, where experts score and rank the proposals on the set criteria. There seemed to be consensus that this would not be advisable to introduce in Commission evaluation processes as it could lead to non-transparent decisions. In addition, specialist experts do not have the overview of the entire portfolio of research projects.

However, looking at the complete set of proposals and potential projects after the qualitative ranking is performed is implicitly performed by high level Boards such as for instance the aforementioned National Cancer Advisory Board in the USA. There was consent amongst the workshop participants that this could possibly be an approach for the future missions. An appointed Board, or group of top level experts (suitably briefed), could, akin to deciding on big investments, identify which of the top-quality proposals would bring the best added value in relation to the objectives of the mission. This group would not assess and rank the individual projects separately, but look at the combination of potential projects. This high level Board activity could possibly be connected to high level Science Advisory Boards.

In any case, all this should obviously be done in greatest possible transparency.

Another interesting approach is to have an initial simple and short application process, after which there is a stage where applicants can interact with each other, guided by the funding body, enabling regrouping, merging, gap-filling etc. Full proposals, the scopes of which adapted accordingly, would then be submitted and evaluated at a second step.

Otherwise, a strategy of multiple calls can be used to address possible gaps in the portfolio of topics, to assure multidisciplinary approaches to a particular global challenge or to spread the high risk and lower risk projects.

It was put forward that good strategic intelligence and knowledge management on what is already in the research and innovation projects portfolio would support the decision making. This could be an overall Commission task to generate the appropriate monitoring information.

Summary

Summarising the key take-away points for the Commission that came out of this workshop:

- the three main criteria (Excellence, Impact and Quality and Efficiency of the Implementation) currently used in Commission are suitable for the future. Flexibility in defining and adapting the sub-criteria would be of importance to cater for different types of research, including mission-oriented research;
- the design of the evaluation process (e.g. whether one-stage or two stage) should be in line with the incentives of the call and take into account important imperatives such as time-to-grant and the burden on both applicants and reviewers;
- blind evaluation is not suitable for mission-oriented and top down research where the potential for impact and feasibility of implementation are closely related to the background of the consortium partners;
- recruitment of evaluators should be reviewed to see what measures might enable more people in employment outside the research community to participate; and
- engaging a high level Advisory Board or similar group in a later stage of the evaluation process was considered an interesting option to be used for missions and particularly for portfolio management of research and innovation projects.

Appendix III: Persons contributing to data gathering

Interviewees

Tore Duvold	Danish Innovation Fund
Barbara Gerratana	REA (previously at NIH)
Eleanor Harris	UK Research and Innovation
Andreas Netz	Vinnova
Theo Saat	NWO
Alice Wemaere	Environmental Protection Agency Ireland

Other persons consulted

Laurent Bochereau	EEAS- Beijing
Alex Cooke	European Counsellor Industry, Innovation and Science Australia
Yumi Chung	EEAS - Seoul
Leonidas Karapiperis	EEAS - Tokyo
Mary Kavanagh	EEAS - Washington
Takahiko Miyao	NEDO