



## OBSERVER REPORT

**Programme:** Horizon 2020 Framework Programme  
**Call for proposals:** H2020-S2RJU-2020  
**Call deadline:** 27 May 2020

**Date of evaluation:** between 29 May 2020 and 02 July 2020

### Observer

Name	Date	Signature
Joseph Prieur	July 6, 2020	

**Number of pages in this report (title page included):** 16 pages

## 1. Background

This report describes the observer's assessment of the evaluation of the following call:

**Calls for proposals:** H2020-S2RJU-2020

**Published:** 07 January 2020

**Deadline:** 27 May 2020 (original deadline of 21 April 2020 was extended to 27 May 2020 due to restrictions related to the COVID-19 outbreak)

**Budget :** EUR 74.45 millions (EU contribution)

This call covers the following topic(s)/type(s) of action:

Topic code	PANEL	Type(s) of action	Number of eligible/ received proposals	Budget (estimated EU contribution)
S2R-CFM-IP1-01-2020	1-IA	IA	1/1	€ 8 971 000
S2R-CFM-IP1-02-2020	1-IA	IA	1/1	€ 3 983 000
S2R-CFM-IP2-01-2020	1-IA	IA	1/1	€ 14 970 000
S2R-CFM-IP3-01-2020	1-IA	IA	1/1	€ 11 408 000
S2R-CFM-IP4-01-2020	1-IA	IA	1/1	€ 5 207 000
S2R-CFM-IP5-01-2020	1-IA	IA	1/1	€ 7 879 000
S2R-CFM-IPX-01-2020	1-RIA	RIA	1/1	€ 2 021 000
S2R-CFM-IPX- CCA-02-2020	1-RIA	RIA	1/1	€1 041 000
S2R-OC -IP1-01-2020	2-RIA	RIA	3/3	€ 2 300 000
S2R-OC-IP1-02-2020	3-IA	IA	1/1	€ 4 600 000
S2R-OC-IP1-03-2020	4	RIA	3/3	€2 420 000
S2R-OC-IP2-01-2020	3-RIA	RIA	3/3	€ 1 340 000
S2R-OC-IP2-02-2020	3-RIA	RIA	4/4	€350 000
S2R-OC-IP3-01-2020	5	RIA	5/5	€ 1 350 000
S2R-OC-IP3-02-2020	6	RIA	4/5	€ 1 710 000
S2R-OC-IP3-03-2020	5	RIA	3/3	€ 2 700 000
S2R-OC-IP4-01-2020	7	IA	1/1	€ 2 000 000
S2R-OC-CCA-01-2020	6	RIA	4/4	€ 950 000
S2R-OC-IPX-01-2020	2-CSA	CSA	3/3	€ 250 000
TOTAL			42/43	€ 75 450 000

The topic code in the above table refers to the IP (Innovation Programme) to which it belongs:

IP 1: Cost-efficient and reliable trains

IP 2: Advanced traffic management and control systems

IP 3: Cost efficient and Reliable High Capacity Infrastructure

IP 4: IT Solutions for Attractive Railway Services

IP5: Technologies for sustainable and attractive European rail freight

CCA: Cross Cutting Themes and Activities

IPX: System Architecture and Conceptual Data Model (CDM)

It also refers to one of the 2 the topic categories, namely:

- The CFM (Call for Members) is restricted to S2R JU members only. Only one proposal per topic will be funded, provided that it passes all thresholds.
- The OC (Open Call) is open to non S2R JU members only. Depending on budget availability, more than one project may be funded.

Each of the 19 topics has its own dedicated EU budget, and therefore topics are not competing against each other for the same budget. However, proposals in response to the OC call for any specific topic may compete against each other within that topic budget line.

The total estimated EU budget for the call is 75.45 million euros. In addition, the S2R JU members (other than the Commission) are expected to provide in kind contributions up to 68.35 million euros (for proposals in response to CFM topics), while 2.8 million euros for non S2R JU members would not be funded from the EU budget, thus resulting in an overall estimated budget for the call of 146.6 million euros

All proposals were of the single stage submission type. A total of 43 proposals were received with a total value amounting to 185.8 million euros, of which 113.3 million euros were requested from EU, i.e. about 50% above the available EU budget (i.e. an oversubscription ratio of about 1.5). The 43 proposals were distributed over 7 evaluation panels:

- Panel 1-IA (6 IA proposals, in response to 6 CFM topics, i.e. 1 proposal for each topic),
- Panel 1-RIA (2 RIA proposals, in response to 2 CFM topics, i.e. 1 proposal for each topic),
- Panel 2-CSA (3 CSA proposals, in response to 1 OC topic),
- Panel 2-RIA (3 RIA proposals, in response to 1 OC topic),
- Panel 3-IA (1 IA proposal, in response to 1 OC topic)
- Panel 3-RIA (7 RIA proposals, in response to 2 OC topics)
- Panel 4 (3 RIA proposals, in response to 1 OC topic)
- Panel 5 (8 RIA proposals, in response to 2 OC topics)
- Panel 6 (9 RIA proposals, in response to 2 OC topics)
- Panel 7 (1 IA proposal in response to 1 OC topic)

As shown on the above table, 8 CFM proposals (2 RIAs and 6 IAs) were received, i.e. one per topic, whereas 35 OC proposals (30 RIAs, 2 IAs and 3 CSAs) were received, ranging from 1 to 5 per topic. The number of proposals to be evaluated per panel ranges from 1 (for Panel 7) to 9 (for Panel 6)

All 43 proposals were considered eligible at the start of the evaluation process. However, one proposal in panel 6, was later declared ineligible as explained in section 3 of this report

The present report analyses the efficiency of the procedures, usability of the instruments (including IT tools), conduct and fairness of the evaluation sessions, and compliance with the applicable rules. The objective is to give independent advice for improving the evaluation processes for the S2R JU calls.

It should be noted that the consensus phase of the evaluation was a “virtual local” exercise, i.e. performed 100% remotely (due to the COVID-19 outbreak) in contrast with standard local consensus phases normally held in Brussels

The independent observer Joseph Prieur has been retired from ONERA, the French national aerospace lab, since June 2011. He was not involved at all with the call under evaluation. He has a wide experience of European Commission research framework programmes and proposal evaluations, first as END/SNE (Expert National Détaché /Seconded National Expert) in DG RTD and DG ENTR (now DG GROW) from 2001 to 2005, and, from 2006, as an independent expert (evaluator, rapporteur/recorder and observer) on various occasions and on several themes (Aeronautics, Surface Transport, Security, Space, Clean Sky JU, S2R JU, BBI JU, NMP, EIT) for FP6, FP7 and Horizon 2020 programmes

## 2. Methodology

The approach followed by the observer was in line with the requirements of the S2R JU evaluation team and with the subsequent observer contract. All relevant information was made available to the observer in a timely manner.

The observer was involved in the evaluation process after the entry into force of his contract CT-EX2002B070194-147 on 19 May 2020 and was given access to all documents and to the SEP system from 18 June 2020. The observer work consisted in the following activities:

- Review general and specific Horizon 2020 documents and relevant evaluation rules ([https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2018-2020/annexes/h2020-wp1820-annex-ga\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2018-2020/annexes/h2020-wp1820-annex-ga_en.pdf)) and H2020 Grants Manual - Section on: Proposal submission and evaluation [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/pse/h2020-guide-pse\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/pse/h2020-guide-pse_en.pdf)
- Review and become familiar with strategic S2R JU documents, namely:
  - The S2R Strategic Master Plan: providing a high-level strategic vision to achieve the S2R objectives and identifying key priority research areas ([https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/jtis/h2020-masterplan-shift2rail\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/jtis/h2020-masterplan-shift2rail_en.pdf))
  - The S2R Multiannual Action Plan (MAAP) part B, providing a long-term investment plan, identifying projects, milestones and deliverables to achieve the Strategic Master Plan objectives [https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/jtis/h2020-maap-part-b-shift2rail\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/jtis/h2020-maap-part-b-shift2rail_en.pdf).
  - The annual work plan and budget 2020 document <https://shift2rail.org/wp-content/uploads/2019/11/ANNEX-GB-Decision-8-2019-Annual-Work-Plan-and-budget-for-2020-REV-20191128.pdf> containing the 19 relevant topic descriptions for this call (specific challenge, scope, expected impact, complementarity with other topics and/or on-going projects, and type of action)
- Review briefings and other relevant information provided to the experts/evaluators and recorders by the S2R JU evaluation team.
- Attend a specific observer briefing via teleconference on 18 June 2020.

- Review the specific document relevant for the evaluation of CFM proposals where the lump sum reimbursement is foreseen [https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/lump\\_sum/lumps\\_umdecision\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/lump_sum/lumps_umdecision_en.pdf)
- During the remote consensus phase, attend the overall general briefing to the experts, and then the consensus and ranking & panel report meetings via the Webex platform in such a way as to cover as many different evaluation panels and moderators and as many topics as possible.
- Throughout the consensus phase (25 June to 01 July 2020) have some exchanges by mail with the call coordinators, and via Webex with some panel moderators, to clarify some details of the consensus process.
- Prepare the present report

The observer did not participate to the individual evaluation phase and therefore this phase will not be addressed in the present report.

### **3. Assessment**

#### Scale of complexity of the evaluation task:

In terms of number of proposals to be evaluated the task was not extremely complex, as only 43 proposals were received.

More experts than the minimum required per type of action (3 for RIA, 5 for IA) were selected in order to get a high level of skills, experience and knowledge in particular in the case of CFM proposals. Each proposal was evaluated by 3 to 6 experts and an additional expert was selected to act as a recorder for each panel to prepare the consensus reports.

Observers from relevant European institutions such as EC, ERA (European Railway Agency) and GSA (European GNSS Agency), IP steering committees representatives (for OC proposals only), were present in the various panels to ensure coordination of all rail activities at EU level and Programme level. These observers are not evaluating proposals but, in their coordination role, can assist evaluators.

Further, a financial expert was present in Panel 1 to contribute to the evaluation (primarily in relation with the criterion “Implementation”), of the CFM proposals for which the grant agreement will be based on the EC financial contribution provided as a lump sum (as opposed to reimbursement of actual costs).

The recorders were dedicated recorders (i.e. not involved as evaluators). They had prepared and submitted to S2R JU moderators drafts of their consensus reports before the start of the consensus phase. The drafts were prepared in line with recommendations provided in a specific recorder briefing: they displayed common views as well as diverging opinions between experts, thus providing a good basis to start an effective and efficient discussion on the proposal merits.

Finally, 2 quality checkers (QC) from the S2R JU team were involved in checking the CRs established at the end of consensus meetings and providing their feedback to the moderators. Most of the CRs were finalized after one single iteration with QC, exceptionally only after a second iteration.

Overall, the relative complexity of the evaluation exercise was not so much related to the number of proposals to be evaluated, but rather to the number and variety of participants (moderators, evaluators, recorders, observers of different types and functions, financial experts, quality checkers) and the novelty of having an entirely remote consensus phase via Webex.

The observer is of the opinion that the S2R evaluation team was well prepared to meet the scale and complexity of the tasks, and the challenge of such a “virtual local”, totally remote, exercise. All phases of the evaluation process attended by the independent observer (briefing, consensus meetings, quality checks, proposal ranking and panel report meetings) and the roles of each of the participants were clearly defined. The planning of the evaluation phase (panel composition, and schedule of meetings for each panel) was also clearly communicated to all participants. The planning allowed for 2-hour meetings, with a time interval of half an hour as buffer time between consecutive meetings, which is a wise precaution, although the actual duration of consensus meetings is largely unpredictable. Some minor technical problems with the use of Webex created small delays and disturbances, but overall, the planning of the consensus phase was rather well kept.

#### Transparency of the procedures:

The procedures for the evaluation and selection of proposals were transparent to all participants (evaluators and observers) including the sequencing of the process into individual and consensus phases (with CR drafting in between), quality checks, review and approval of CRs, and finally ranking of proposals and panel report preparation and approval.

The modalities of the evaluation were clearly explained in briefings delivered to the experts at the start of the individual evaluation phase and of the consensus phase. The comprehensiveness of these briefings illustrates the willingness of the S2R JU staff to have a fully transparent evaluation process. During the consensus period, the independent observer was invited by the S2R JU call coordinators, to raise any question at any time, and indeed all questions received prompt and adequate answers, which reflects also the full transparency of the process.

The independent observer was also provided with a copy of the early briefing delivered on 29 May 2020 at the start of the individual evaluation phase, which contains a vast amount of recommendations, tips, and suggestions for evaluators on how to write high quality reports... a very useful tool, especially for first-time evaluators.

The procedures are in line with the rules described in Annex H of the Work Programme and they comply with the principles established by the Commission and reminded to experts during briefings (independence, impartiality, absence of conflicts of interest, objectivity, consistency, accuracy). Each aspect of the evaluation process is described in detail in the “Grants Manual: Section on proposal submission and evaluation” referred to in the S2R Annual Work Plans.

[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/pse/h2020-guide-pse\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/pse/h2020-guide-pse_en.pdf) available to all experts and to the independent observer

The evaluation criteria, scoring system and thresholds were well understood by all experts. The requirement of “commenting before scoring” was also clearly understood and applied by all participants. Moderators often invited experts to refer to the score interpretation table, after the consensus comments had been agreed upon, in order to ensure that the given scores indeed reflected the agreed comments.

The briefings covered also adequately specific aspects of the evaluation related to the Lump Sum Pilot application to the CFM proposals  
[https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/lump\\_sum/lumpsumdecision\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/lump_sum/lumpsumdecision_en.pdf) including the role of the financial experts in the evaluation process and their participation to the consensus meetings

Specific arrangements due to the unusual remote character of the consensus phase such as the lack of physical signature of the panel reports were explained clearly to the evaluators by each moderator in each panel: after approval of the CRs in SEP ( i.e. the usual procedure) and review of the complete panel report, including the ranked list, all evaluators expressed verbally their agreement (via the Webex platform) on the report and were then sent an e-mail requiring them to confirm in writing their agreement. This is to avoid sending around by mail the report itself which is a confidential document.

The observer is of the opinion that the remote character of the consensus phase did not adversely impact the transparency of the evaluation procedures He is fully convinced that the evaluation was conducted with high standards of transparency, fairness and diligence.

#### Throughput time of the evaluation and the efficiency of the procedures:

On average, only 6 proposals per panel were to be evaluated, ranging from 1 (panel 7) to 9 (panel 6). This rather wide scatter comes from the fact that panels must be organised, and evaluators recruited well ahead of the proposal submission deadline when it is still unknown how many proposals will fall under each panel.

The 3 financial experts involved in the financial evaluation of the 8 CFM proposals had about 1.5 week to complete their individual assessments (about 1/ working day), then 1 week to arrive at a consolidated financial report per proposal. The financial reports were made available to the experts and moderator of Panel 1 ahead of the consensus phase, and one of the 3 financial experts attended all 8 consensus meetings of this panel.

Two S2R JU staff members acted as quality checkers (QC) during the consensus phase. The observer noted that on 1 or 2 occasions a panel had to wait for the feedback from QC, but this waiting time never exceeded 10 minutes and therefore the evaluation schedule was never at risk for reasons related to the QC process.

Not only does the observer believe that having dedicated recorders is a good practice in terms of efficiency of the evaluation process, but also it allows all evaluators to take fully part into the consensus debate without one of them having to concentrate on collating the opinions of

his co-evaluators with the risk of not being in a good position to express/defend his own view, or the opposite risk of imposing his own views onto his fellow experts.

Recorders appeared to have good writing skills and be able to capture the substance of the individual comments of the evaluators in order to formulate common views and diverging opinions to be addressed and resolved during the consensus meetings. Some of them appeared to have also some general knowledge about the topics/subjects under discussion.

#### Efficiency, reliability and usability of the procedures, including the IT-tools:

The IT evaluation tool SEP was employed in the evaluation process. The tool is fully operational, and all participants seem familiar with the system.

The Webex platform used in this remote consensus phase appeared to be familiar to most, if not all, the participants. Minor technical issues appearing throughout the consensus phase were quickly resolved and did not generate any real disruption of the process.

The observer believes the procedures used are reliable and robust and the implementation of these procedures was overall very good.

#### Impartiality, fairness and confidentiality of the evaluation:

From his participation to this remote evaluation process, the observer is of the opinion that the evaluation was impartial, fair and conducted in full compliance with the rules set by the Commission, in accordance with specific S2R rules, and in line with the steps indicated in the “Grants Manual: Section on proposal submission and evaluation”. These rules and procedures were reminded to the experts as part of their briefings, and as needed during all steps of the evaluation.

The overall evaluation process is compliant with the principles established by the Commission: independence, impartiality, objectivity, accuracy and consistency.

The confidentiality requirements, which are annexed to the contracts of the experts and the observer, were reminded to the experts by the S2R JU staff involved in the evaluation

The following factors are believed to contribute significantly to achieve impartial and fair evaluations

- the rather large number participants: 3 to 6 independent experts, plus the moderator, the dedicated recorder, the independent observer, other observers from EC or Agencies, sometimes financial experts, and quality controllers
- The variety of the expert backgrounds, ensuring that all aspects of each proposal are considered, and that each proposal is looked at under different angles, with also a specific emphasis on the financial aspects for the CFM proposals due to the Lump Sum type of funding.
- the early detection and treatment of potential conflicts of interests. All experts, and the observer, sign a declaration of absence of conflict of interest as part of their contract. The importance of avoiding conflict of interest is also reminded to experts in each briefing. Pre-selected potential experts with one or more suspected conflict(s) of interest may be excluded from the evaluation process once the proposals are received. During the consensus phase, as



soon as a conflict of interest is detected with an evaluator in relation to a proposal, (s)he is immediately excluded from participating to the discussion of this proposal.

#### Conformity of the evaluation with the applicable rules (including guidance documents)

The “H2020 Vademecum section on proposal submission and evaluation” (<http://www.h2020.cz/files/pracna/H2020-Vademecum-Section-on-Proposal-Submission-and-Evaluation>) is a document meant to assist all staff in implementing proposal evaluation (from the Commission, the Executive Agencies or JUs). It provides a common basis for all personnel involved in the evaluation to have a similar approach in conducting the evaluation. The document is consistent with the rules and procedures given in the Grants Manual - Section on: Proposal submission and evaluation ([http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/pse/h2020-guide-pse\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/pse/h2020-guide-pse_en.pdf)), and with briefings provided to the experts.

In addition, experts have been provided with a SEP User guide [https://ec.europa.eu/research/participants/data/support/expert/expert\\_evaluation\\_user\\_manual.pdf](https://ec.europa.eu/research/participants/data/support/expert/expert_evaluation_user_manual.pdf) , as well as with very comprehensive and detailed briefings.

All the information and documents were in full conformity with existing applicable rules and the general annexes to the Work Programme [https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2018-2020/annexes/h2020-wp1820-annex-ga\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/wp/2018-2020/annexes/h2020-wp1820-annex-ga_en.pdf), especially Annex H which describes the evaluation rules

The observer is convinced that all phases of the S2R JU-2020 call evaluation were conducted in full compliance with the rules and guides contained in the above-mentioned documents.

#### Quality of the evaluation process in comparison with similar national/international evaluation procedures:

The observer has not been involved in evaluation processes other than EU research ones. However, based on his experience with the observation of many Horizon 2020 calls, the observer believes that the EU evaluation process is robust and of high quality and, from earlier discussions with some experienced evaluators over different previous calls, can be considered better than or as good as other national and/or other international research funding schemes. And it is generally acknowledged that several countries are looking to the EU process, or even transposing it, to improve their national process.

At international level, the main evaluation principles are objectivity and impartiality, independence of evaluators, participation of all parties concerned in the entire process, transparency and focus, reliability, completeness and clarity of reports, fairness and protection of the interests of the parties involved. The observer is convinced that all the above-mentioned international principles are well embedded and efficiently implemented in the EU evaluation process.

#### Overall quality of the evaluation:

The overall quality of the evaluation process was excellent, thanks to the contribution of all involved parties (experts, recorders, various observers, and S2R JU staff)

The evaluators were selected primarily on the basis of their expertise. They all are familiar with the Horizon 2020 programme and were well informed about the specificities of the S2R JU calls.

The independent observer participated in the remote consensus phase of the evaluation process, was welcomed to attend any consensus meetings he wished to attend, and free to raise any question with the S2R JU evaluation staff (call coordinators and panel moderators) at any time during the evaluation.

The S2R JU staff involved in the evaluation, showed full compliance with the established rules and procedures.

In conclusion, the whole evaluation process was conducted by expert evaluators, and S2R JU staff with the highest professional and quality standards and in accordance with the applicable rules and guidelines. The observer is of the opinion that the evaluation was conducted in a fair, transparent and open way and that each proposal received adequate and equal treatment.

#### Other remarks

- quality of the documentation provided to experts beforehand

Although the independent observer was not involved in the evaluation process until 18 June, he was given access to the briefing material delivered earlier to the experts (on 29 May) at the start of the individual evaluation phase. This briefing material appears to be of high quality and very comprehensive covering general information about S2R JU, an overview of the S2R JU call 2020, and extensive details about the evaluation process and procedures. All relevant aspects of the evaluation process were presented: the timing, the different panels, the lump sum pilot, the evaluation criteria & sub criteria, the scoring system, the role of the different participants, the different phases (individual evaluation, consensus, panel), the confidentiality requirements, the Conflict of Interest situations, and all major principles of the process ( impartiality, objectivity, accuracy, independence, consistency)

Also, very useful in this initial briefing, especially for first-time evaluators, were a lot of guidance and recommendations concerning the write up of IER for each criterion and sub criterion

A specific briefing was also delivered to the recorders on 17 June 2020 with very detailed information and guidance about what was expected from them and tips to prepare high quality CRs

The independent expert is convinced that this information was comprehensive and of high quality

- quality of the remote consensus briefing

A briefing was delivered to the experts on the morning of the first day of the consensus phase (25 June 2020), explaining in detail the purpose, sequence, rules and content of the consensus and panel meetings, with the actors involved and the roles of each participant. This briefing

repeats and expands some of the information provided to the experts at the start of the individual evaluation phase. This repeat often proves to be very useful. Further, it does provide more focused and detailed practical information relevant for the consensus and panel stages, as well as practical information to the experts about the payment. This general briefing provides also an opportunity for experts to raise questions and seek clarifications.

This briefing session was well organized, and the information provided was clear and comprehensive.

The independent observer noted also that brief and very practical specific mini-briefings were delivered by some moderators to their experts of their panel at the start of the consensus discussions, and at the start of the final steps of the evaluation (CR final approval, proposal ranking, and panel reporting). This is believed to be a useful initiative, specially to remind the experts about the purpose and content of the panel report

- the understanding by experts of the call (context, topics), of the evaluation process and their role and of the award criteria and scoring scheme

The experts benefited from 2 comprehensive and detailed briefings at the start of the individual and consensus evaluation phases. In addition, recorders were specifically briefed to produce high quality CRs.

The observer did not have the opportunity to talk directly with the experts due to the remote character of the evaluation phase that he attended. However, from his observation of this phase, the observer believes that experts were all perfectly aware of and understanding the call context, the annual work plan 2020, the topics, the CFM and OC types of proposals, the various action types (CSA, IA, RIA) for this call, the Innovation Programmes (IPs).

They also appeared familiar with and understanding important issues such as eligibility of proposals and Conflicts of Interest, their own roles, the grouping of proposals in 7 different panels, the evaluation methodology and process, and the award criteria and sub criteria.

The “no negotiation“ principle was well understood by the experts who refrained from suggesting “improvements” to the proposals, even though there is room for some flexibility in the interpretation of this principle as stated in the briefing to the experts

Other important aspects clearly well understood by evaluators were the need to avoid penalizing twice (or rewarding twice) a proposal for the same reason under two different criteria/ sub criteria, the need for “commenting before scoring”, the required consistency of the comments with the scores, and the need to pay attention to the wording of the consensus report in order to (1) avoid encouraging claims for redress and (2) ensure that applicants get an accurate, clear and fair feedback from the evaluation (the ESR).

- the allocation of experts to proposals: balance (gender, geographic, sector), relevance and balance of expertise, ...

Experts were selected primarily on the basis of their expertise in relation to the topics and therefore to the proposals to be evaluated. In addition to the “standard” experts/evaluators, financial experts were selected (from the existing expert database) to make a separate review

of financial aspects of CFM proposals which will be funded (if successful) under the lump sum pilot scheme.

It was noted that there were 46 experts in total, of which 30 % were women. One third of the experts were new and had never been involved in earlier evaluations pertaining to the Horizon 2020 Programme. Experts were of 11 different nationalities and residences, the largest number (over 40%) coming from Spain and Italy.

Experts were from a wide range of organisations, more than half of them coming from private-for-profit, or high education organisations, 13% from public organisations, and 8 % from research organisations

- the process of the individual evaluations and the actors involved

Not applicable in this report, as the independent observer was not involved in the individual evaluation phase

- the process of the consensus meetings and the actors involved

Consensus meetings were all performed remotely, using the Webex platform. The consensus meetings went smoothly thanks to a careful planning of the work for the different panels. As already mentioned, there were large differences in the workload of the different panels, but the relatively low number of proposals to be evaluated (43) was such that the complete consensus evaluation could be completed within less than 5 working days (actually from 1 day for panel 7, up to 4.5 days for panel 6)

There was a common approach in conducting consensus meetings with however some differences depending on how the roles are being shared within the tandem moderator/recorder. Some moderators are actively managing the consensus meetings whereas others are less pro-active and tend to let the recorder have a more active role in running the meeting. In all cases the recorder and the moderator always displayed rather good complementarity in their respective roles. The observer believes that this complementarity is very important and the observed small differences in conducting the consensus meetings are well within the flexibility margins allowed by the rules and do not have any adverse impact on the outcome of the evaluation.

In all cases, recorders refrained adequately from expressing their views on the proposals and from suggesting scores

There was sufficient flexibility in the planning to allow discussions to take sometimes significantly longer than planned, and therefore there was no detrimental effect of these long discussions on the planning. On a few occasions, the discussion lasted for more than 1 hour on one single criterion. The observer believes it is the responsibility of the moderators to act as a timekeepers.

It was observed that moderators and recorders encouraged the evaluators to qualify their negative comments as shortcomings, or minor shortcomings, or weaknesses, etc. which is extremely useful when it comes to put scores against the comments

In one particular case, a panel agreed to discuss and comment all 3 criteria, before scoring all of them together at the end of the meeting. The observer would like to suggest that this should be avoided. The score should be given while the discussion on the criterion is still fresh in the mind of the experts rather than 1 or 2 hours later when one or 2 other criteria have been discussed. Scoring all 3 scores only at the end of the consensus meeting brings the risk of mixed-up & overlapping views about the 3 criteria and therefore penalising (or rewarding) twice a proposal for the same reason. On the contrary, scoring each criterion immediately after the discussion of this criterion has the advantage of closing the discussion on a criterion before starting the discussion on the next criterion, thus reducing the risk of overlaps.

- criteria and scoring scheme: appropriateness, completeness, relevance, clarity, consistency in application, ...

The applicable 3 main criteria, along with the set of sub criteria (specific for each action type: IA, RIA or CSA) were explained in detail to the experts during the briefings. They constitute a simple and robust evaluation system which all stakeholders are used to work with. The 3 criteria are rather independent from each other and do not overlap, thus minimizing the number of situations where comments could apply to 2 different criteria.

There were no specific difficulties expressed by evaluators in understanding the scoring system nor the significance of thresholds. The score interpretation table was extensively used in every consensus meeting before scoring each criterion against the agreed comments. And negative comments were systematically qualified in terms of severity using the specific wording of the scoring table (minor or major shortcoming, etc.). This proved to be very helpful to ensure that scores were indeed reflecting and supporting the comments. The systematic use of this methodology contributes to the consistency in applying the evaluation scores.

The observer noted that, in all meetings he attended, the word “weakness” was not used, and all negative points were just “shortcomings” (minor, neutral, or major). This may explain why, in the end, 90% of the proposals were above thresholds. Of course, this may reflect that 90% of the proposals were good or better, but not using the full range of “severity” does not seem to reflect well the recommendation to use the full range of scores.

- final panel meeting and the actors involved

Each topic had its own budget and all proposals pertaining to any given topic (1 proposal per topic for each CFM topic and from 1 to 5 proposals per topic for OC topics) were evaluated within the same panel. As a result of this small number of proposals per topic, proposal ranking and final panel meetings were rather straightforward formalities in each panel. The observer noted only one occurrence (in panel 6) where 2 equal score proposals (RIAs) had to be ranked: they were easily ranked on the basis of the difference in their “excellence” score as explained by the moderator, in accordance with Annex H of the general annexes to the Work Programmes Horizon 2020. There was no situation of minority views to be resolved which was brought to the attention of the observer.

Due to the remote nature of the consensus phase and the need not to send out the panel reports (for confidentiality reasons), at the end of the panel meetings moderators asked the evaluators

to agree verbally on the ranking list and on the panel reports; then an e-mail would be sent to evaluators to request them to confirm their agreement in writing, by return mail.

- occurrence and handling of specific issues (if any) such as conflicts of interest

To the extent possible, conflicts of interest (CoI) situations were resolved at an early stage in the evaluation process, immediately after receipt of the proposals, in parallel with the eligibility check, once the applicants are known. For that purpose, more experts than strictly needed were invited to express their interest and availability ahead of the proposal submission deadline, anticipating that some of them will have to be eliminated due to CoI situations.

CoI situations detected during the consensus phase were extremely rare and handled simply by excluding an expert with a CoI from the evaluation of a proposal and replacing him (her) by another expert without a CoI for that proposal.

There was one case (in panel 6) when the ineligibility of a proposal was confirmed at the end of the consensus phase when it was established that one beneficiary had close links with an associated member of S2R (which is not allowed for an OC type of proposal). By disregarding this beneficiary, the minimum number of participating countries, as required by the Horizon 2020 rules of participation, was not met. Once declared ineligible, this proposal was immediately discarded from the ranking process

- quality of evaluation summary reports

All participants are made aware from the start of the process of the importance of producing high quality CRs (which eventually become ESRs) in order to ensure a clear, fair and useful return to the applicants and avoid opening the door for complaints.

At the consensus stage the use of dedicated recorders with good writing skills ensures a good quality level of the reports. In addition, 2 quality checkers from the S2R JU Team were tasked to check all CRS and come back to the panel moderators to request corrections/clarifications/additions or deletions when needed.

As a result of this process the evaluation summary report quality is generally very high. However, the presence of quality checkers should not be seen by the expert panels as an encouragement to pay less attention to the quality of the initial write up of the CRs

- overall conduct of staff: responsiveness, hospitality, competence, ...

Throughout the evaluation period attended by the observer, the professionalism and support of all S2R staff involved in the evaluation were outstanding. All participants ( call coordinators, moderators, quality checkers) responded to all requests from the observer in a prompt, efficient and cordial manner and demonstrated a high degree of openness and transparency. The observer was given a comprehensive schedule of all consensus and final panel meetings and was free to select any meeting he wanted to attend.

- infrastructure and working conditions for evaluators

During both the individual phase and the consensus phase of this evaluation the infrastructure and working conditions for the experts and the observers were their own home conditions, due to the remote character of this evaluation.

Only a few members of the S2RJU evaluation team were able to work from the S2RJU premises in the White Atrium building in Brussels

- workload and time given to evaluators for their work

The observer is of the opinion that the allocated period to perform individual remote evaluation (29 May -18 June 2020) was sufficient for all experts and all proposals (about 6 proposals per expert, on average), even for those who were the most loaded (panel 6 with 9 proposals/expert)

The same applies for the onsite consensus phase: the number of proposals per panel ranged from 1 (Panel 7) to 9 (Panel 6). Five working days were originally planned but it appeared that the whole exercise could be completed in 4.5 days. No comment/complaint was heard by the observer about the workload.

- remuneration of evaluators (in relation to workload)

The observer did not note or hear any specific comment about the remuneration in relation to the workload for this particular S2R JU call 2020.

#### **4. Recommendations**

On the basis of the observations made during the “virtual central” evaluation via Webex, the observer would like to express the following ideas/suggestions

- the concept of a dedicated rapporteur (or recorder) with good writing skills is to be continued as it contributes to the quality and the efficiency of the evaluation. The specific recorder briefing is a valuable tool towards achieving high quality CRs (and ESRs)
- some effort should be put into harmonizing the language between all parts of Horizon 2020: for example, it should be decided once and for all to use the word “subtopics” instead of “work streams”. Although it may sound like a minor issue, the observer would like to submit that this may lead to some complaints if proposals with the same score are prioritised according to their coverage of a “workstream” not covered by a higher ranked proposal... the applicants may argue that this is against the rule in Annex H which mentions exclusively “ topics, or sub-topics”. The word “workstream” is never used
- The observer would like to recommend that the number of experts per proposal should preferably be an uneven number (3 or 5 or 7) and that even numbers like 4 or 6 should be avoided simply because, in case of lasting disagreements, if there is a need to vote, there will always be a majority with an uneven number.
- The involvement of dedicated “quality checkers” for ensuring quality, homogeneity and coherence in the reports is a very good practice. One has to be careful that the

quality check does not become a bottleneck in the evaluation process. Also, the presence of quality checkers should not be a reason for experts and recorders to “delegate” their own responsibility to the quality checkers: the quality of CRs and ESRs is the responsibility of ALL and not only the responsibility of the quality checkers

- The observer would like to suggest that the score for each criterion should be agreed immediately after agreeing the comments related to that criterion (before going to the next criterion) and not at the end of the consensus meeting. This would reduce the risk of mixing up and overlapping comments and of double penalising (or double rewarding) a proposal for the same reason
- Finally, one suggestion coming from the fact that this consensus phase was entirely remote and experts never met: S2R JU may want to consider inviting all selected experts to send a short résumé (5 to 10 line maximum) which would be distributed to all co-evaluators within a panel, ahead of the consensus phase, in such a way that each of them has an immediate and quick overview of the specific background of his colleagues.

In closing, the observer would like to express his view that the evaluation process was conducted thoroughly and very professionally by all participants throughout the complete consensus phase of the process. Judging only from the procedural aspects (which is the remit of his mandate) the observer is convinced that all proposals received adequate, fair and equal treatment. The outcome of the exercise is a set of ranked lists of projects where the very best ones have been identified and will be funded within the budget limits.

The observer would also like to express his thanks and gratitude to all staff of the S2R team involved in the evaluation process, especially the call coordinators Nadia Debza and Neil Griffin and all panel moderators Lea Paties, Judit Sandor, Monique van Wortel, Esther Bravo, Gorazd Marinic, Sébastien Denis and Javier Ibanez de Yrigoyen, for their transparency, openness and friendly support throughout the whole period. This support contributed to maintaining some social dimension in this “virtual local” evaluation exercise, thus making it quite enjoyable.

Jo Prieur

Independent Observer

