



# RESEARCH ASSESSMENT PROCESSES: GATHERING EVIDENCE FOR A SCIENCE EUROPE INITIATIVE FOR MUTUAL LEARNING

KATHARINA WARTA AND MARIA DEL CARMEN CALATRAVA MORENO

DOI: 10.22163/fteval.2020.487

## ABSTRACT

This paper discusses a collective approach in the design, implementation and validation of a study commissioned by Science Europe on research assessment processes of research funding and research performing organisations. The collective approach is based on the involvement of its member organisations and Science Europe itself at different stages of the study for the mutual learning of all stakeholders and the community in general.

This paper describes the study, including the purpose, methodology and findings, and discusses the importance of its findings and recommendations for research funding and performing institutions, as well as the singularity of its approach from the perspective of evaluation practices.

## INTRODUCTION

Assessment of research is conducted in a wide variety of situations, such as the review of research output, decisions on future research, the appraisal of researchers and also of entire research units and organisations. This is reflected by a wide range of approaches and criteria, often dependent on the academic discipline, institution profile, etc. The challenges, however, are often similar, such as a high number of applications, limited resources, difficulties to differentiate among applications, the trade-off between excellence and relevance, etc.

In 2019, our institution conducted a study on behalf of Science Europe (SE) to explore practices used for the selection of research proposals in competitive research funding programmes and the selection of researchers for promotion within research organisations. The study deliberately puts the focus on generic programmes, “to establish a thorough and comparable knowledge base of the current and developing assessment processes.” (SE 2020a, p. 9). The study was part of a broader exercise leading to a set of recommendations on research assessment processes published by SE in July 2020.

Although the entire process must not be taken as an evaluation, it has interesting aspects in common, most importantly, the establishment of a sound and comparable evidence base with the objective of institutional (and policy) learning. The exercise was partly motivated by the broad

discussion on new approaches to research assessment – it was then somewhat surprising that the evidence gathered in our study showed that generic programs are widely based on standard processes. This article presents our approach to capture dynamics in a slowly moving context. Before that, we provide an introduction to the broader framework of our study, as we are convinced that this is a promising setting to feed empirical results into design processes, and thereby reinforce the above-mentioned dynamics.

## SCIENCE EUROPE DIGGING INTO RESEARCH ASSESSMENT PROCESSES

Science Europe (SE) represents the major public research organizations in Europe. It has 36 members from 27 European countries, of which 31 are research funding organizations and five are major research performing organizations. Founded in 2011 in Brussels, it provides a collective voice for its member organisations (MOs) to advocate and shape science policy and funding. “Ensuring the quality of science” is one of the long-term objectives of SE, and the improvement of research assessment practices is one of the derived priorities, next to cross-border collaboration, EU framework programs, open access, research data, research infrastructure, and recently also COVID-19. Based on consultations, events, and studies, SE produces a variety of publications, like responses/reactions to (European) policies, briefing papers, brochures, factsheets, joint or position statements, as well as survey reports to give examples.

In July 2020, SE published a position paper with recommendations on research assessment processes, based on a study launched by SE and conducted by our institution in 2019 and a broad consultation among Science Europe MOs and stakeholders from the research community in 2020. These recommendations shall provide a framework for further development and optimization of processes and aim to promote knowledge sharing and mutual learning between research organizations.<sup>1</sup> They are linked to previous and ongoing work of other international initiatives, like

1 See Science Europe (2020a), p9.

the San Francisco Declaration on Research Assessment (DORA), the Leiden Manifesto for Research Metrics (Hicks et al, 2018), Global Research Council (2018) Statement of Principles on Peer/Merit Review, the Joint Statement on Research Assessment, jointly released by Science Europe and the European University Association (EUA) and various publications of Science Europe. The recommendations address the following dimensions and themes:

- Approaches used to assess and select proposals and researchers: (i) transparency of research assessment processes, (ii) evaluation and monitoring the robustness of research assessment processes;
- Challenges faced during assessment processes: (i) discrimination, bias, and unfair treatment in research assessment practices, (ii) cost and efficiency of research assessment processes, and applicant investment of time and effort;
- Current developments in the assessment of proposals and researchers.

These recommendations are primarily about assessment processes and methods, and not so much about criteria.<sup>2</sup>

In this article, we focus on the process of this undertaking, to share the experience of contributing as external consultants to a collective evidence gathering and analytical exercise, in the following four sections: First, we present our role as external consultants in the broader context, which is different to our “classical” role as external evaluators (section 3). Section 4 provides a summary of the key findings of our study, on research assessment practices. In section 5, we present our methodology to grasp change, section 6 provides some conclusion on the relevance of this undertaking for the research policy evaluation community.

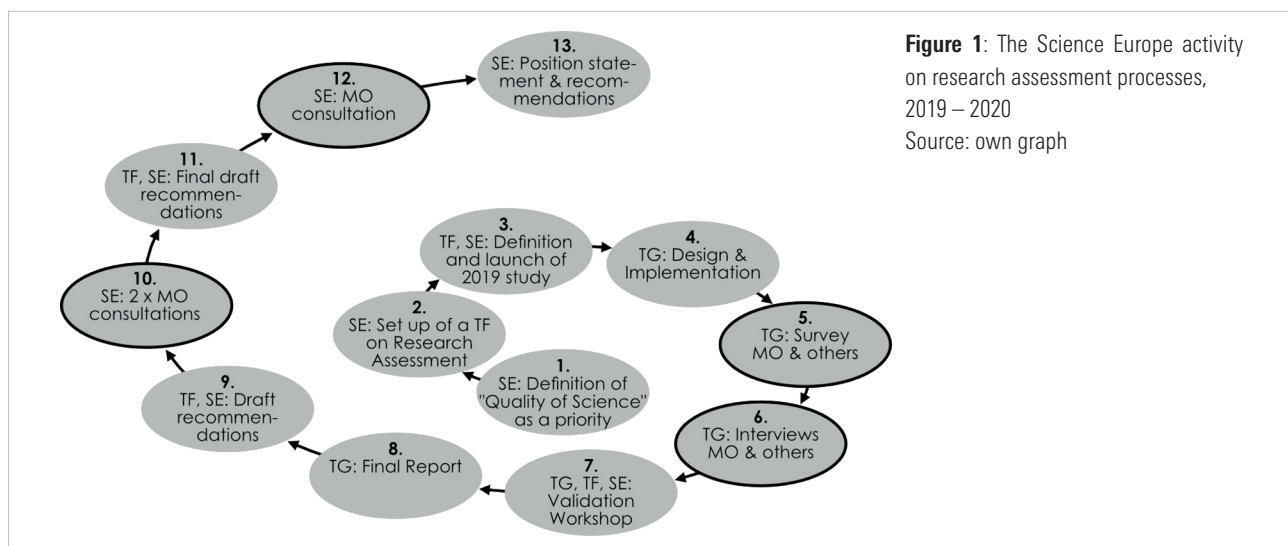
## GATHERING EVIDENCE FOR SCIENCE EUROPE: THE EMBEDDEDNESS OF THIS STUDY IN A BROAD PARTICIPATORY PROCESS

In July 2019, following a competitive call for tenders, our institution was commissioned a study on research assessment processes. The objective of the study was “to investigate ways to find out how SE member organisations (MOs) processes for research assessment lead to selecting the best projects for funding and researchers for their career progression.”<sup>3</sup> The findings are based on the analysis of policy documents and documentation specific to the participating RFOs and RPOs<sup>4</sup>, an online survey covering their ‘generic competitive funding’ or ‘generic researcher promotion’ scheme<sup>5</sup>, followed by 20 semi-structured in-depth interviews with a subset of these organisations, and a validation workshop with the Task Force on Research Assessment (TF) set up by SE, as well as representatives of the Science Europe Office.

The following questions have guided the study:

1. What approaches are used to assess and select proposals and researchers in a robust, fair and transparent manner?
2. What are the challenges that research organisations face during the assessment processes?
3. What are the current developments in the assessment of research proposals and researchers?

It is important to note that our study is one piece in a broad participatory and collective exercise, involving member organisations of SE as well as other invited organisations at several points in time as shown in Figure 1.



**Figure 1:** The Science Europe activity on research assessment processes, 2019 – 2020  
Source: own graph

<sup>2</sup> Ibid.

<sup>3</sup> See Science Europe (2020b), p2.

<sup>4</sup> These include annual and final reports of funding actions, publications of calls for applications, regulations of the processes in research assessment exercises, and guidelines for applicants and reviewers participating in the assessments.

<sup>5</sup> The survey provides a broad overview with particularly good coverage of funding organisations that are members of Science Europe, and additional information of some non-members as well as non-European RPOs and RFOs. It was completed by 38 organisations (33 RFOs, 4 RPOs and 1 organisation that functions both as RFO and RPO), with an overall response rate of 86%.

In the evidence-gathering phase, participants responded to the questionnaire survey (step 5) and some of them were also interviewed (step 6). In the concluding phase, member organisations were involved in broad consultations (steps 10 and 12)<sup>6</sup>, based on a first draft and then revised draft of recommendations and a first draft of conclusions in form of the Position Paper prepared by the TF and SE.

In addition, the TF which is composed of representatives from seven member organisations, played a crucial role in the design and launch of the study (step 3), in the discussion and validation of study results (step 7), and drafting and finalising conclusions (steps 8, 10, and 12). Therefore, the study was designed as a truly collective exercise in which external consultants engaged with all other types of stakeholders for the different phases of the study in different roles: with MOs as participants (step 5 and 6), MOs representatives as members of TF, TF and SE as validation partners (step 7-10).

Our institution mainly had the role of a service provider for the professional collection of evidence and data analysis. Following the terms of reference of the study, we developed and used an appropriate methodology and tools to carry out the study, including suggesting the necessary amendments to the proposed questionnaire to ensure that the online survey is robust, and then collected and analysed the data gathered through the online survey and subsequent targeted interviews.

## KEY FINDINGS ON ASSESSMENT PRACTICES

The key findings of the study can be summarised as follows:<sup>7</sup>

### DIVERSITY OF ORGANISATIONS WITH WIDELY SHARED BASIC PRINCIPLES ON RESEARCH ASSESSMENT

Although the organisations participating in this study are of diverse nature, have a different focus and implement a variety of programs, they have common well-established practices for the assessment of research and researchers, primarily the use multi-stage research assessment processes, external single-blind peer reviews and panel reviews. Other approaches such as rankings, external open reviews, and internal single-blind reviews are also used but to a lesser extent. The least common approach is double-blind reviews, although it is used by one participating RFO with satisfactory results to make the research assessment more objective.

Transparency has received considerable attention in the design of the research assessment processes of the participating organisations not only after the assessment process has been concluded (i.e. by providing feedback from reviewers) but also prior to it (i.e. the publication of the assessment criteria, description of the process and actors) and during its implementation (i.e. through the introduction of rebuttal phases).

## CHALLENGES DURING THE ASSESSMENT PROCESS

The mandate to ensure that the assessment process successfully selects the best projects for funding and researchers for promotion was discussed with the participating organisations. Reliance on competitive systems, peer review, multi-stage evaluation processes, written assessment guidelines and qualitative evaluations were discussed by most participating organisations as the key elements for ensuring robust assessments in this regard. Additionally, measures to prevent and detect discrimination and bias are in place in most organisations. The most scrutinised potential biases are gender and discipline, followed by affiliation in the case of RFOs and seniority in RPOs. Generally, the regulations or guidelines for assessment established by the organisations raise awareness on this topic and 68% of the surveyed organisations form reviewer panels with diverse profiles to minimise potential discrimination or bias.

Limited research funds and academic positions set more pressure on research and promotion assessment processes. Particularly challenging is distinguishing and ranking proposals and candidates for promotion when they are of similar quality and worth funding/promoting.

The cost and efficiency of the research assessment are also discussed, particularly in evaluations that do not rely on quantitative indicators. Moreover, the balance between the quality and cost of the research assessment is of critical importance not only for the organisations but also for the scientific community whose members are involved as reviewers and for the applicants. Approaches for improved efficiency for these three stakeholders were discussed by the participants. Some of these approaches aim to optimise the assessment and application efforts, for instance, through the introduction of a scoring system that translates the qualitative assessment to a quantitative scale that facilitates the ranking of candidates, or the introduction of multi-stage evaluation processes to reduce the effort invested by both reviewers and applicants, or the streamlining of funding schemes and standardisation and the standardisation of application processes.

### CURRENT DEVELOPMENTS IN RESEARCH ASSESSMENT AND ALTERNATIVE METHODS

Most organisations rely on a qualitative assessment of research, some of them in combination with quantitative approaches (i.e. the number of publications in high-ranking journals) but most of them give higher importance to the qualitative assessment than to any other quantitative approach. Some interview participants elaborated on recent updates on guidelines for assessment to inform reviewers of the importance of the qualitative assessment and to discourage the use of metrics.

Experimentation with alternative assessments systems and tools takes place at a rather incremental basis and in selected small programs. Drawing lots, sandpits, double-blind assessments are being piloted by some organisations, while in others these are already in place mostly for specific programs and purposes. Several organisations are considering

<sup>6</sup> In step 10, the first consultation was in written format, the second was organised in way of two virtual meetings: "A first draft of policy recommendations, based on the knowledge gathered from the 2019 Study, was formulated by the SE Office and the TF in January 2020. An online consultation survey was run between 4 February and 11 March for all SE MOs, whether they had participated in the first phase or not. A remote video-conference workshop was held on 24 March for nominated experts from SE MOs to further discuss the recommendations and topics." At that point, commonalities and divergences in SE MO strategies were identified, these diversities "were the subject of a second virtual discussion during an SE MO consultation event on research assessment processes on 24 March." (SE 2020b, p3). Based on that, the SE Office and TF developed a first draft of the Position Statement (step 11) which was sent to SE MOs for a third and final written consultation (step 12.)

<sup>7</sup> This section quotes parts of the executive summary of the study (Technopolis Group 2019).

the use of altmetrics, while some others do not use it but recognise a broad format of research outputs.

Although non-academic impact and significance are often not considered in large generic research funding programs and promotion schemes, evidence was gathered on several RFOs creating mission-oriented funding schemes to prioritise such kind of research. These programs are adapting their research assessments with different or extended criteria and reviewers to better assess this kind of research.

## OUR APPROACH TO GRASP DYNAMICS

According to the terms of reference, the study should identify trends, gaps and new directions with regard to testing robustness of selection processes, assessment tools and pilots and experiments. However, related to their generic programs – which are the focus of the study –, most organisations regularly revise and refine small aspects of their research assessment methods on a more-or-less incremental basis. Despite a broad discussion of challenges and resource limitations, major changes

could hardly be observed in practice. In order to grasp these dynamics, we opted for two ways of questioning:

The first option was to ask respondents to indicate whether their organisation has implemented changes in the way research proposals or candidates for promotion are assessed, or whether they plan to do so.<sup>8</sup> This allows a differentiated understanding of tools and practices. As shown in Table 1, for most organisations the assessment of the research content of scholarly publications is either a long-standing practice, a recent change or a planned change. The broadening of the range of quantitative tools used to assess research is considered by a significantly lower proportion of organisations. In fact, most organisations have reduced or are planning to reduce the use of journal-based metrics. However, it is difficult for RFOs and RPOs to verify whether reviewers do not use quantitative tools or criteria in their assessment.

As a second option, along with the information about current or past use of an element and considerations to use it in the future (or not), respondents with the experience of using it (currently or in the past) were asked to assess its importance. This approach was for instance used to identify aspects of research that reviewers are required to assess in research assessments (Table 2).

	Long-standing practice	Made this change	Planning to make this change	Not made this change and not planning to do so in the future	Do not know	Not applicable
Reducing the use of journal-based metrics	8 (21%)	13 (33%)	3 (8%)	7 (18%)	4 (10%)	4 (10%)
Eliminating the use of journal-based metrics	6 (15%)	9 (23%)	4 (10%)	9 (23%)	6 (15%)	5 (13%)
Broadening the range of non-publication research outputs required to assess	4 (10%)	14 (36%)	5 (13%)	8 (21%)	5 (13%)	3 (8%)
Broadening the range of quantitative tools that are used to assess research impact	1 (3%)	6 (15%)	6 (15%)	14 (36%)	7 (18%)	5 (13%)
Considering qualitative indicators of research impact, such as influence on policy and practice	6 (15%)	10 (26%)	5 (13%)	11 (28%)	5 (13%)	2 (5%)
Considering the research content of the scholarly publications <sup>9</sup>	17 (44%)	7 (18%)	6 (15%)	3 (8%)	3 (8%)	3 (8%)
Being explicit about the criteria used in the assessment	29 (74%)	2 (5%)	4 (10%)	1 (3%)	2 (5%)	1 (3%)

**Table 1:** Long-standing practices, changes and plans for changes in research assessments Source: own data based on the survey answers of the organisations participating in this study (n=39).

8 See Technopolis Group (2019), p. 18f.

9 The assessment of the content of the scholarly publications was often conducted on a selection of publications provided by the applicant as relevant previous work.

	Currently using	Used in the past	Never used but considering using in the future	Never used and not considering using in the future	For organisations (RFOs and RPOs) that are using or have used the respective aspects:		
					Very important	Moderately important	Less important
<b>Soundness of the proposed methodology</b>	32 (100%)	0 (0%)	0 (0%)	0 (0%)	29 (91%)	1 (3%)	0 (0%)
<b>Feasibility of the proposed research</b>	33 (100%)	0 (0%)	0 (0%)	0 (0%)	29 (88%)	2 (6%)	0 (0%)
<b>Resource allocation in line with objectives</b>	31 (97%)	0 (0%)	0 (0%)	1 (3%)	17 (55%)	11 (35%)	1 (3%)
<b>Feasibility of research in relation to applicants' expertise</b>	33 (100%)	0 (0%)	0 (0%)	0 (0%)	27 (82%)	4 (12%)	0 (0%)
<b>Complementary expertise of researchers</b>	28 (97%)	0 (0%)	1 (3%)	0 (0%)	16 (57%)	11 (39%)	0 (0%)
<b>Dissemination plan</b>	28 (88%)	1 (3%)	1 (3%)	2 (6%)	11 (38%)	13 (45%)	3 (10%)
<b>Novelty of the research question</b>	33 (100%)	0 (0%)	0 (0%)	0 (0%)	24 (73%)	7 (21%)	0 (0%)
<b>Potential econ. and soc. Impact</b>	22 (69%)	1 (3%)	2 (6%)	7 (22%)	6 (26%)	14 (61%)	1 (4%)
<b>Potential transfer/commerc.</b>	19 (59%)	2 (6%)	2 (6%)	9 (28%)	2 (10%)	13 (62%)	3 (14%)
<b>Potential contribution to public policies</b>	17 (55%)	2 (6%)	2 (6%)	10 (32%)	4 (21%)	10 (53%)	2 (11%)
<b>Ethical considerations</b>	32 (100%)	0 (0%)	0 (0%)	0 (0%)	25 (78%)	5 (16%)	1 (3%)

**Table 2:** Aspects of research that reviewers are required to assess in research assessments.

Source: own data based on the survey answers of the organisations participating in this study (n=39).

This allows the analysis of trends both in terms of use and relevance of aspects reviewed in research assessments. For example, the study provides evidence that the majority of the programs already ask reviewers to consider the potential economic or societal impact in their assessment. However, only a minority ranks this as highly important. This might be surprising as research policy is increasingly considering the need to stimulate research in directions that provide knowledge relevant to tackle societal challenges. In fact, the 'generic programmes' organisations participating in this survey indicated they mainly focus on scientific criteria and have a high level of stability of assessment criteria, as not many changes have taken place in the past, nor are considered for the future.

## CONCLUSION

Science Europe's (SE) engagement to analyse research assessment processes and formulate related recommendations provides several learning opportunities for the evaluation community, even if it is undoubtedly not an evaluation itself.

First, it brings more light into research assessment processes and SE's member organizations' approaches to evaluate these processes. This was not the focus of this article, but in the position paper, SE recommends that "All organisations should conduct evaluations of the robustness of their assessment processes.", and "Organisations should

re-evaluate their processes at fixed intervals, whenever broad reforms to assessments are implemented, or when problems are identified.” (SE 2020a, p. 13). As a matter of fact, this kind of evaluations gain importance, and fteval will also devote an event on this topic.<sup>10</sup>

Second, the approach clearly aligns with some of the defining principles of evaluations, namely “a transparent and systematic procedure, based on empirically obtained data; distinct from everyday assessment procedures” (fteval 2019, p6). However, here, in contrast to evaluations, the entire process of the study constituted a collective exercise, in the sense that a community of actors engaged in the formulation of recommendations for themselves. They are organized in an association, with administrative support, a task force and the support of an external service provider for the collection of evidence. This collective approach naturally complies with the 3rd principle of RTI Evaluations (fteval 2019, p. 11), addressing participation, and seems particularly promising concerning the 4th principle, namely utilization and benefits: “The benefits of an evaluation are generally enhanced if relevant interest groups are involved in the evaluation process, if specific evaluation questions are formulated and responded professionally, and if coherent recommendations are communicated as a result of the evaluation” (p.12).

In return, the 7th principle of independence needs to be looked at more closely, it states “the evaluation is not materially influenced or manipulated by political interests, the client, programme managers or those affected, nor by any possible bias of the evaluators themselves.” In this kind of study, we would argue that the inclusion of all member organisations in the analytical phase (not only for providing data and information), provided transparency and re-iteration and ensured that the result has no bias. We would, however, argue that the involvement of external consultants with sound experience in (independent) evaluation procedures helped to ensure that the evidence base has clear priority over any individual interest.

As a matter of fact, evidence on research assessment processes in generic programs shows that they are more stable than we would have expected, given the challenges, growing constraints, but also new technical opportunities. Sound questioning approaches allowed to get a differentiated view on past and present experience, intentions and the importance of a broad range of aspects. Many of these were taken up in the SE’s recommendations published in July 2020. Of course, their implementation can only be assessed after a while, but the active involvement of the member organisations in their formulation is both promising and exemplary.

## REFERENCES

- fteval - Austrian Platform for Research and Technology Policy Evaluation** (2019): Evaluation standards for research, technology and innovation policy, Vienna.
- Global Research Council** (2018): Statement of Principles on Peer/Merit Review. [https://www.globalresearchcouncil.org/fileadmin//documents/GRC\\_Publications/Statement\\_of\\_Principles\\_on\\_Peer-Merit\\_Review\\_2018.pdf](https://www.globalresearchcouncil.org/fileadmin//documents/GRC_Publications/Statement_of_Principles_on_Peer-Merit_Review_2018.pdf)
- Hicks et al.** (2015): Bibliometrics: The Leiden Manifesto for Research Metrics. <https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>
- San Francisco Declaration on Research Assessment (DORA)** (visited in September 2020). <https://sfedora.org/>
- Science Europe** (2019): Joint Statement on Research Assessment. <https://scieur.org/se-eua-assessment>
- Science Europe** (2020a): Position Statement and Recommendations on Research Assessment Processes. <https://www.scienceeurope.org/our-resources/position-statement-research-assessment-processes/>
- Science Europe** (2020b): Research Assessment Processes, Methodology of the activity. <https://scieur.org/ra-methodology>
- Technopolis Group** (2019): Science Europe Study of Research Assessment Practices. <https://scieur.org/ra-report-2019>

## AUTHORS

### KATHARINA WARTA

*technopolis group | austria*

Rudolfplatz 12/11

1010 Wien

T: +43 664 8407232

E: [warta@technopolis-group.com](mailto:warta@technopolis-group.com)

<https://orcid.org/0000-0002-5937-0044>

### MARIA DEL CARMEN CALATRAVA MORENO

*technopolis group | austria*

Rudolfplatz 12/11

1010 Wien

T: +43 664 8846 7985

E: [carmen.calatrava@technopolis-group.com](mailto:carmen.calatrava@technopolis-group.com)

<https://orcid.org/0000-0001-6493-2251>

## KEYWORDS

Research assessment; collective evidence validation; assessment dynamics; Science Europe; research funding organisations; research performing organisations.