

DEVELOPMENT OF A PERFORMANCE ASSESSMENT SYSTEM FOR SCIENTIFIC RESEARCH ORGANISATIONS, AS A PART OF THE ONGOING RAPID REFORM

I. Dezhina

The performance assessment system for scientific research organizations, which is currently being developed in Russia, has become an integral part of the ongoing reform in this country's scientific research complex. The assessment principles and criteria are determined by the RF Government's Decree issued in November 2013; the opinion poll among the directors of research institutes was conducted by the Federal Agency for Research Organizations in April and May 2014. This gave rise to a new wave of debate concerning the parameters to be applied in the assessment. An analysis of the main trends in this debate shows that the issue as to the purposes of the assessment has been overlooked.

The period April–May 2014 saw intensification of the efforts aimed at determining the principles and procedures for assessing the performance of scientific research organizations. The necessity of such an assessment and the main rules for its conduct are outlined in the RF Government's Decree No 979, of 1 November 2013, 'On Introducing Alterations into Decree of the Government of the Russian Federation of 8 April 2009, No 312'¹. The principles on which the assessment should be based, in accordance with the Decree, are as follows:

- the assessment should be conducted by independent organizations;
- scientific research organizations should be arranged into reference groups irrespective of their departmental subordination, with due regard for their specific field and the types of scientific research conducted by them;
- the assessment should rely on the same indicators as are applied for assessing the performance of scientific research organizations in the developed countries².

In early 2014, it was planned that the assessment procedures would be further adjusted with due regard for the opinions expressed by the scientific research community, and in particular the directors of institutes subordinated to the Russian Academy of Sciences (RAS). Since then, scientists, on their own initiative, have been putting forth proposals concerning possible alterations to the assessment procedures. However, it was only at the second conference of representatives of scientific research personnel of the RAS on 25 March 2014 that head of the Federal Agency for Research Organizations (FARO) Mikhail Kotiukov finally confirmed that RAS personnel would indeed

be involved in developing the proposed performance assessment system for scientific research organizations³. Later, on 16 April 2014, the FARO sent a letter to directors of scientific research organizations, asking them to submit, by 25 April, their comments on various assessment parameters – that is, effectively within 10 days. This caused another surge of criticism aimed at the government. Such a response originated by no means only from the fact that too little time was allotted for the discussion of such serious an issue. The FARO's letter asked for comment not only on the assessment parameters or procedures – it addressed, in fact, the fundamental principles on which the assessment was to be based. Thus, it looked as if the FARO had decided to launch the entire project anew and complete it within a record short time. The Agency was interested, among other things, in answers to the following questions:

- frequency of performance assessments;
- the reliance on only quantitative data, or the necessity to back them up with experts' estimation assessment;
- the feasibility of assessing the performance not only of scientific research organizations, but scientific research groups;
- the principles for enrolling members in the experts' boards⁴.

In accordance with the approved government plan, the development of assessment procedures should be finalized by 1 July 2014⁵, which means that even the

3 FANO privlechet sotrudnikov RAN k razrabotke sistemy otsenki effektivnosti institutov [The FARO Will Involve the RAS Personnel in the Development of the System for Assessing the Performance of Institutes. RIA Novosti, 25 March 2014. <http://ria.ru/science/20140325/1000989328.html>

4 Letter of the FARO of Russia to directors of scientific research organizations, 10 April 2014, No 007-181-07.

5 Second letter by Academician Alexey Parshin on experts' assessment of scientific research organizations to the FARO on 27 April 2014. See <http://www.saveras.ru/archives/9059>

1 <http://pravo.gov.ru:8080/page.aspx?67047>

2 For more details concerning these parameters, see Situation in Russian Science and Innovation Sector // Russian Economy in 2013. Trends and Outlooks. (Issue 35) – M.: Gaidar Institute 2014. P. 367–369.

FARO will have too little time at its disposal for pooling and processing the submitted proposals, and for making the final decision – given the fact that even institutes that share the same field of study may offer different points of view. The short timelines for decision-making in this case remind us of the way the reorganization of three state academies was carried out – swiftly, without a discussion, and without assessing its possible consequences. This time, the same administrative approach is applied in developing the performance assessment procedures for research institutes.

The main target for the criticism voiced by the scientific research community against the government performance assessment project was the choice of the level of an organization and not that of a laboratory, based on the principle of division of the organizations into reference groups, as well as the reliance on bibliometric data – whose limitations are very well known. Scientists are almost entirely unanimous in their opinion that an assessment at the institute level will result in data distortions, especially because there are many multidisciplinary research institutes in this country, as well as institutes with unique specialization, whose performance it will be nearly impossible to assess in the framework of such an approach. The procedures for conducting the assessment and selecting the experts were also discussed, including the possibility of attracting international experts. Opinions differ – with a slight prevalence of those against the involvement of foreign experts.

There was no single opinion concerning reference groups, either. Some believed that reference groups represent an approach that provides formal grounds for closure of research institutes – and these would be by no means always poorly performing ones. Others considered reference groups to be important, but only as a basis for comparing laboratories, and not entire institutes. According to many of the participants in the discussion, institute branches must not be closed down – instead, the assessment results may show the ways to upgrade and support the weak ‘links’. However, even a comparison made on the basis of bibliometric data at the laboratory level may be biased towards those who work in more popular fields, where citations can be noted more often.

The argumentation against preferential reliance on bibliometric data was supported by references to foreign experiences. The most popular in this connection has become the recent experience of the UK, where the government introduced a new achievement assessment methodology in the field of science (Research Excellence Framework)¹. Indeed, this initiative

had been discussed for several years before finally being implemented, it is well-grounded, and each field applies its own measurement criteria and assessment procedures. The key elements of this system are assessments at the level of departments and laboratories, and the use of bibliometric data only as a secondary tool. Moreover, the journal impact factors – as well as the overall number of works published by a given scientific research group – are not taken into consideration. Bibliometric data may be applied only for the purpose of supporting and adjusting experts’ estimations. Experts, in their turn, assess the quality of publications of each research laboratory (or group) on the basis of four best publications over the last 5 years. The number of citations and other types of bibliometric analysis are not applied in the assessment of research results in social studies in humanities. In the field of economics and econometrics, available citation data are taken into consideration whenever they are necessary as supplementary information, and the absence of citations from a given study does not influence the results of its assessment. And finally, one more important consideration is the purpose of introducing such a system: the assessment results are used as a basis for redistributing the financing flows between administrative structures and for determining the number of additional jobs that can be created in one or other department of a given institute.

The UK experience appears to be convincing, it has already been tested in a pilot mode, and so the Russian government – who have proclaimed their reliance on methods for assessing ‘the performance level of scientific research organizations in the developed countries’ – can take it into consideration. This experience becomes even more relevant if we point out the fact that, in contrast to the UK – where no reforming takes place in the field of science, and so the discussions as to its possible improvement may be lengthy – in Russia the situation is such that we cannot afford to spend years on the discussion of government plans.

The consolidated opinion of active representatives of the scientific research community is reflected in the letter of the Council of the Society of Researchers to the Chairman of the RF Government ‘On the Performance Assessment of Scientific Research Organizations’². The letter states that ‘the core assessment target must become laboratories and research groups, and not entire institutes. And the assessment must essentially rely on expert opinions, and not simply on quantitative performance indicators. The subdivision of all scientific re-

1 Research Excellence Framework 2014. Panel Criteria and Working Methods. http://www.ref.ac.uk/media/ref/content/pub/panelcriteriaandworkingmethods/01_12.pdf

2 Letter of the Council of the Society of Researchers to the Chairman of the RF Government ‘On the Performance Assessment of Scientific Research Organizations’ of 28 April 2014. See <http://www.saveras.ru/archives/9102>

search organizations into three categories (the leaders, the stable ones, and those with no prospects for development) on the basis of their quantitative parameters cannot be recognized as acceptable'. In this connection, 'it is necessary to adjust Decree No 979 without delay'. Thus, the main message of the letter is that the proposed key approaches must be rethought once again without haste, because the more rapid procedures envisaged by the FARO may result in irreparable mistakes.

On 22 May 2014, an expert session took place, where the directors of institutes formerly enjoying an 'academic' status attempted to develop a common viewpoint. The outcome of this meeting is yet another confirmation of the fact that it is not easy to achieve a consensus even inside the scientific research community. In fact, no majority-approved decisions with regard to such issues as the feasibility of creating reference groups, as well as the status of experts and the methods of their selection were made at the caucus. Among the positive results, we may point to the inclusion in the draft resolution of the provisions stipulating that an assessment should be based on experts' estimations (with due regard for quantitative data), and that institutes must independently select the data that will be then submitted for the consideration by experts.

At the same time, the government has not released any statements – and this aspect remained practically unmentioned – as to what specific purpose an assessment should be tailored to. The purposes may be different, and they will ultimately determine the assessment principles. Thus, for example, an assessment may have the purpose of determining the number of personnel, organizations and institutes, in order to make proper cuts. In this case, some targets must be offered for the planned reduction. For example, the targets were clearly set when the PRND system (performance indices for scientific research) was introduced for a three-year period in the RAS. All the institutes were told to reduce the number of their research personnel by 20% within three years, without any consideration for the actual performance level of each institute. The results were monitored, some optimization was achieved, the salary level somewhat increased, and so the clearly stated purpose was

achieved (this is not the place to discuss whether that purpose was reasonable). Thus, if the goal is to optimize the existing network of scientific research organizations subordinated to the FARO, some targets – are least approximate – must be offered, for instance 'reduction of the scientific research personnel by no more (or no less) than by x%'.

In order to achieve this goal, an initial assessment would be advisable, with some modest targets, so as to understand the actual scale of the forthcoming personnel cuts. For example, it can be estimated how many specialists, in the last 5 years, published less than 5 articles in the journals entered in the Higher Attestation Commission's list. This could serve as a sort of initial, rough baseline.

However, if the goal is to more efficiently distribute available funding, and so, to merge institutes working in the same field and undertake some similar measures, – another approach will be necessary, with reliance on assessments at the level of laboratories, in the framework of each specific field of research. The first step can be a preliminary discussion, on the basis of groups consisting of representatives of institutes operating in adjacent fields of research, aimed at elaborating a consolidated opinion on the assessment principles to be applied in each given field of research. It must be emphasized that the unit to be assessed must be a field of research, and not an institute or its structural subdivision.

The degree of specificity of each field can be different. Thus, for example, if the approximate number of redundant workforce is to be assessed, a more general division into fields of science may be possible. However, if the goal is to optimize the performance of the existing system without necessarily reducing the number of institutes, the by-field subdivision must be more detailed, and give consideration also to the new areas of research.

And finally, as the year-long moratorium on transactions involving the Academies' property and alteration to the personnel structure is to be over by the year-end 2014, the remaining months can be used for preparing for final approval the proposals concerning assessment procedures. ●