

# How Good is Good?

## **BŁAŻEJ SKOCZEŃ**

Faculty of Mechanical Engineering, Kraków University of Technology  
blazej.skoczen@pk.edu.pl

## **MARIA LEWICKA**

Faculty of Psychology, University of Warsaw  
marlew@psych.uw.edu.pl

## **MACIEJ ZABEL**

Institute of Histology and Embryology, University of Medical Sciences,  
Poznań  
Institute of Histology and Embryology, Medical University, Wrocław  
mazab@ump.edu.pl

## **AGNIESZKA POLLO**

Department of Physics, Astronomy and Applied Informatics  
Jagiellonian University, Kraków  
National Centre for Nuclear Research, Warsaw  
Academia, Magazine of the Polish Academy of Sciences  
apollo@fuw.edu.pl

**On 30 September 2013, the Ministry of Science and Higher Education announced the results of the latest evaluation of Poland's scientific institutions. A new categorization method, aiming to build on past experiences, has been used**

Scientific institutions around the globe are subject to periodic evaluation. In Poland, this happens every four years. The reason is simple: the resources dedicated to science and research are generally limited, so we need an algorithm to help us channel funds to those institutions that use them the most effectively. Evaluation may also be used to help develop the country's scientific policies by indicating the desired areas of growth and promoting the strategic fields of research.

It's worth noting here that there has never been a perfect system, and it is unlikely that one will ever be created. Almost by definition, science is difficult to quantify, and it tends to be unpredictable: today's breakthrough may turn out to be a dead-end in a few years' time, while inventions that now seem beyond our wildest dreams may prove popular with our great-grandchildren.

There are two main approaches, which are somewhat at odds with each other: expert evaluation, where scientists are assessed by their peers, and parametric evaluation, based on numerical criteria and indicators. The first method may be

unreliable due to the bias that might be shown by experts, while the second involves the difficulty of selecting the right parameters, which may not actually be directly convertible into evaluation points. It is a well-known phenomenon that as soon as an evaluation measure becomes publicly known, it automatically becomes subjective, since the individuals or institutions being assessed will strive to obtain the best results according to that particular parameter. This year's evaluation of Polish institutions was based on numerical criteria, although it also took into account statements presented by the institutions following the experiences of previous years. In other words, parametric evaluation was supplemented by elements of expert evaluation.

Additionally, inspired by methods of multi-objective analysis (multiple criteria decision making), the principles of pairwise comparison were also adopted, following superiority/inferiority relationships. Scientific institutions were assessed on the basis of four key criteria: (I) scientific and creative achievements, (II) scientific potential, (III) economic outcomes of scientific activities, and (IV) other achievements. Criteria I and III are not directly related to institution size: in these cases, point-based evaluation is divided by the number of research staff members. Criterion II is not scaled the same way, which means that larger institutions generally obtain better results. Finally, criterion IV is based purely on expert opinions, and involves the assessment of ten achievements nominated by the institution itself as representing its finest (outstanding) work during the evaluation period.

After being split into joint assessment groups, scientific institutions are compared on a pairwise basis for each criterion in turn; the point value assigned to each institution is defined using the superiority/inferiority relationship. This means that in each assessment, the stronger institution can attain a maximum score of +1 point, and the weaker a minimum score of -1 point. After all pairwise comparisons are conducted in each joint assessment group, institutions receive points for each of the four criteria. The final result for each institution is a certain objective function, whereby individual criteria are awarded different weightings. To take into account the specificity of the given scientific community, different weighting systems have been adopted depending on the particular scientific field, as well as on whether the unit being evaluated is a PAS scientific institute, a university faculty, other R&D institute, or a different type of institution.

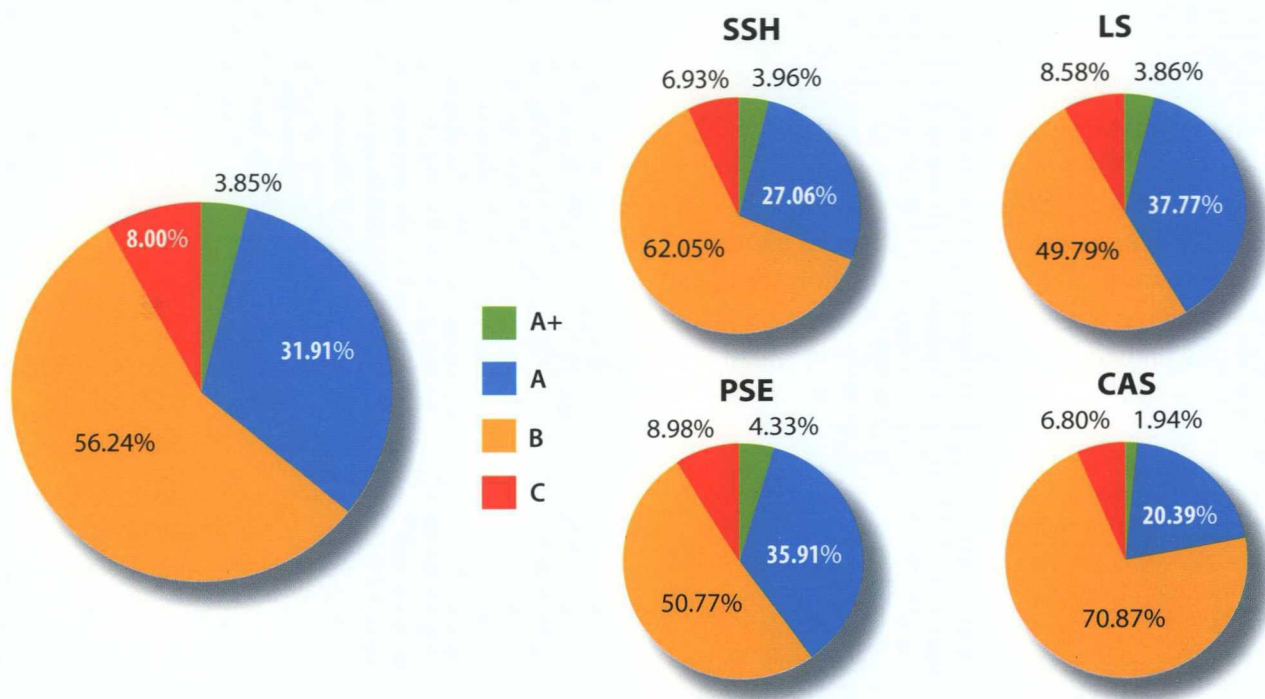
To be ranked in category A or category B, each institution had to beat the "reference" institutions for that category (virtual units defined according to a common principle linking

their level to the median of the top 15% in each joint assessment group, to reflect typical point levels in each scientific field). From the outset, there were not going to be many “outstanding” A+ institutions: this category was awarded when a category A institution fell in the top 25% of its joint assessment group, and – even more importantly – when it stood out in terms of citation rate, prestigious awards, and notable achievements.

Developing a robust system for evaluating scientific institutions is an ongoing process. One of the as-yet unsolved problems is posed by interdisciplinary institutions, as well as those whose profile does not clearly fit into any of the joint

of individual scientific disciplines as found in international databases (such as Web of Science, Web of Knowledge and Scopus). The process has only partially been successful, although we hope that this aspect will be better addressed during the next parametric evaluation. We are working on the assumption that the strategic aim of parametric evaluation and categorization is to improve the quality and effectiveness of Poland’s scientific institutions as compared to their equivalent organizations in Europe and around the globe. Parametric evaluation should be more than just an attempt to create a relatively objective map of Polish scientific institutions; it should also be instrumental in influencing our own

Błażej Skoczeń/Paweł Adamów



Percentage distribution of categories A, B, C and A+ in all evaluated institutions and by different scientific fields: social sciences and the humanities (SSH), life sciences (LS), physical sciences and engineering (PSE), and creative and artistic sciences (CAS)

assessment groups. It is also necessary to analyze the question of how to make comparisons between joint assessment groups, which for all intents and purposes is not provided for in the current system. On one hand, many differences are directly linked to the specific peculiarities of each field and the different methods of practicing science, so direct comparisons even between related joint assessment groups frequently result in misleading conclusions. On the other, it is patently clear that Poland has certain very strong fields and others that are simply weak, therefore institutions assigned to category A but in different fields may in fact vary greatly in terms of their scientific excellence.

The aim of the evaluation was to tackle this issue by creating reference institutions based on international standards

scientific circles, and act as a tool for improving their position relative to strong international partners.

Of the 962 institutions and university faculties under evaluation, 307 (32%) were included in the top category A, 541 (56%) were categorized as B, and the 77 (8%) weakest ones were assigned to category C. For the first time, a set of 37 model institutions were awarded the A+ category. It should be stressed that – given Poland’s modest level of research funding from the state budget (around 0.4% of the GDP) in comparison to other European countries – our science and research are in relatively good shape, and in some fields they can be described as being in robust health. ■

*The authors are members of the Polish Scientific Institution Evaluation Committee*