



IREG Observatory on Academic
Ranking and Excellence

ASSESSMENT AND RANKINGS ARE DIFFERENT TOOLS

*IREG position paper
on Agreement on Reforming Research Assessment*

December 2022

IREG welcomes the broad debate on research assessment expressed in the “Agreement on Reforming Research Assessment” by the Coalition for Advancing Research Assessment (CoARA). Different forms of research assessment are used all over the world of higher education. And while they can have significant impact on individuals and organisations, they do not always apply the most valid and responsible methodologies. IREG agrees that systems of research assessment should not rely solely on – a few, and often problematic – quantitative bibliometric indicators.

The proposals aim at responsible methodologies to assess research on various levels, from individuals to entire universities. From an IREG perspective the agreement helps to make a distinction between research assessment and rankings, between their purposes and approaches. Indeed, the agreement lists a number of principles that should guide rankings, too. Rankings, however, play a different role than research assessment.

The agreement sees a necessity to defend institutional autonomy against external pressure exposed by rankings: *“It will help the research community and research organisations regain the autonomy to shape assessment practices, rather than having to abide by criteria and methodologies set by external commercial companies.”* The argument ignores that some bibliometric indicators that are criticised (for example journal impact factors) have been widely used within the HE system before the emergence of (global) rankings. Moreover, every institution is free to choose which, if any, rankings to consider in evaluating its performance or taking strategic decisions. As an expert group we have never observed any limitation to universities autonomy by birth of new rankings, nor (to our best knowledge) penetration of ranking results to the assessment of individuals.

Several principles for assessment criteria listed in the agreement are relevant to rankings, too. The recognition of diversity of institutions for example has already been acknowledged in the **IREG Berlin Principles on Academic Rankings** accepted in 2006. IREG welcomes the agreement as it should stimulate rankings to further refine their methodologies and indicators, for example the postulate *“Use assessment criteria and processes that respect the variety of scientific disciplines...”* is but a call for rankings by subject and to critically review and develop their specific sets of indicators.

Already in 2014 IREG published the **Guidelines for Stakeholders of Academic Ranking** articulated recommendations for a responsible use of rankings. The Guidelines clearly pointed out that *“rankings differ from other instruments of quality assurance within higher education, such as peer review, which is based on evaluation and accreditation”*, and explained that *„one should expect rankings to identify strengths and weaknesses, but not to prescribe remedies or explain successes.”* We must admit that these explanations have not always been heard or understood.

The agreement argues in favour of peer reviews as the gold standard for qualitative research assessment, which could be supported by quantitative indicators. Peer review is, indeed, an adequate and proven method for assessing research of individuals, groups, and institutions. But the agreement largely ignores weaknesses and limitations of peer review. It just states “*the research community re-assesses and improves peer review practices regularly*” to address its biases and imperfections. But this is a normative statement rather than an empirical argument, which can be with equal validity used in the context of rankings.

Concerning the argument in favour peer review, IREG wants to express two major concerns from a ranking perspective.

First, the agreement discusses research assessment as such, and does not sufficiently distinguish between levels of assessment that have implications for the methodology of assessment, namely individual persons, small academic units (alike departments, faculties), entire university and (national) university system. Peer review is a highly resource intense endeavour (largely manpower) and its costs increase with the level of analysis. While peer review is an established system in the assessment of researchers, research proposals and academic units, only few large higher education systems have regular and comprehensive system wide peer reviews of research performance. As far as we know, and for a good reason, there is no single global nor international research assessment based on peer review. It would be completely out of range to expect global rankings to do it!

Second, and most important, most of the peer review and research assessment systems in general are an internal quality assurance tool within the HE sector, largely run by experts and addressed to experts. Rankings have a completely different purpose: **In contrast to research assessment, rankings communicate information on higher education institutions to specific user groups and to the broader public.** A major target group of rankings, prospective and mobile students, are lay users of information about HEIs with limited (prior) knowledge about higher education. Hence, they need a user-focussed reduction of the complexity of information, while research assessment systems aim at the highest level of complexity. Hardly any prospective students will understand highly sophisticated peer review reports, nor have the resources to read six or eight on his/her shortlist of institutions among which to choose.

Rather than a long report, encompassing many aspects of the university performance in a multitude of its activities, the ranking audience is expecting an aggregated information from the fields most relevant for that group particular interests. By virtue, such an approach introduces simplifications, but they make the results easier to understand by the addressed target group. Rankings are, by definition, systems using mainly quantitative indicators to compare units of higher education. Quantitative indicators may be supplemented and contextualized with qualitative information, but no ranking can run large scale peer reviews. Let us repeat: **Rankings and research assessment are two instruments with different purposes and target groups!**

The agreement indicates a need for changes in research assessment; it must be said that university rankings are changing and evaluating, too. A trend to include indicators related to the UN sustainable development goals (SDG) is a good example of rankings evaluation. The agreement fails to recognize the complexity of the ranking world. Global ranking may be getting all the attention in the media but the largest group of stakeholders, prospective students and their parents turn to national rankings for information while debating which university to choose.

The growing number of by subject or specialized rankings are a clear sign that rankings also take measures to improve. Fair and reliable rankings are constructed in line with the core commitments of the agreement: they use the bibliometric/numeric indicators wisely and responsibly, following specificities of the fields, disciplines, or regions. Reliable ranking, scrutinized and commented by international experts, never aim to be a basis for the assessment of individuals.

IREG Observatory on Academic Ranking and Excellence

December 2022

www.ireg-observatory.org