According to a growing number of stakeholders, ‘traditional’ ranking instruments are causing too many negative effects.

With an international research team Frans van Vught and Frank Ziegele developed the new tool U-Multirank, funded by the European Commission and additional philanthropic actors. ‘It does not have the intention to sell information, consultancy services or newspapers.’

The blessings of U-Multirank

Why it is better and more usable than other rankings

Frans van Vught & Frank Ziegele

Transparency in higher education and research is of growing importance worldwide. Universities are increasingly expected to make a crucial contribution to the future life chances of growing numbers of individuals as well as to the innovation and growth strategies of nations and regions around the globe. Obtaining valid information on the characteristics and performances of higher education institutions within and across higher education systems has become increasingly important.

Various types of transparency tools with different purposes exist, particularly at a national level but increasingly also on international and even global levels. These include classifications, rankings/league tables, study guides, benchmarking instruments and quality assurance and accreditation systems. Rankings in particular have become widespread transparency tools, also on a global scale. Well-known worldwide rankings are the Academic Ranking of World Universities (ARWU, Shanghai Jiao Tong University), the Quacquarelli Symonds (QS) ranking, the Times Higher Education ranking and the US News & World Report ranking. However, these rankings have rightly being criticised (Hazelkorn, 2011, Van Vught and Ziegele, 2012, EUA, 2014, Bekhradnia, 2016). As an alternative to these ‘traditional’ league tables, we have developed and implemented a multi-dimensional and user-driven ranking tool, which we will explain below. But let us first formulate our criticisms of the ‘traditional’ rankings.

Our criticisms

- Traditional rankings focus on only one or very few dimensions of the broad spectrum of functions of higher education and research institutions – primarily the research function. This is particularly problematic when ranking information is presented in the form of league tables. Because traditional rankings reduce their ‘message’ to ‘University X is number Y in the world’, users are led to believe that such a statement refers to the overall performance of a university.

- Traditional rankings appear to have a negative effect on the diversity of higher education systems. Because of their preoccupation with research, they tend to stimulate imitative behaviour on the part of institutions focused
This imitative behaviour triggers a ‘reputation race’ in higher education and research worldwide. On one single profile: the large, comprehensive, internationally orientated research university. This profile in fact relates to a ‘top research university’ but in the end becomes synonymous with ‘top university’ or ‘world class university’ in general, at the expense of other important higher education activities, dimensions of performance and successful organisational models.

- This imitative behaviour also triggers a ‘reputation race’ in higher education and research worldwide, stimulating politicians, policy-makers and university leaders to implement a range of policies and major research investments specifically designed to achieve higher ranked positions for their institutions in the league tables, with prejudicial effects on other important areas of potentially improved performance (Van Vught, 2008). In addition, levels of resource inequality between institutions increase. ‘Successful’ institutions are able to generate additional resources on the basis of their position in the rankings and thus achieve further success. This pattern further enlarges academic performance gaps between institutions and adds reputation and resource fuel to vertical academic stratification processes.

- Traditional global rankings do not respond adequately to the different information needs of various stakeholder groups. The weighted composite scores of traditional league table rankings impose the value judgements of the ranking producers on the users, instead of allowing them to choose criteria corresponding to their preferences. The use of composite indicators blurs differences in performance across particular dimensions and indicators and hence decreases transparency. Furthermore, empirical studies show that the weighting systems underlying composite indicators are anything but robust.

- Because most rankings focus only on the level of the institution as a whole, they ignore differences in performance across different disciplinary fields within the institution. In bibliometric terms, traditional rankings are not sufficiently sensitive to varying publication and citation cultures across different disciplinary research fields. In addition, averages across fields are of little use to many users, especially students looking for an academic programme in a specific subject, and can be highly misleading.

- The league table approach of the traditional rankings tends to exaggerate differences between universities (‘number 57 is better than number 61’). Analysis shows that in many league tables small differences in the numerical scores on the indicators can lead to relatively large differences in league table positions.

- Traditional league tables have been shown to trigger strategic behaviour by institutions by providing strong incentives for them to ‘game the results’ to boost their performance positions. The incentives for such strategic behaviour increase with the inadequate use of rankings at the political level, for instance when league table outcomes are applied in funding decisions.

- In their selection of indicators traditional rankings appear to focus on what is easily measurable rather than what is relevant for reflecting performance across the different functions of higher education and research. In addition, they do not take into account major contextual differences between higher education systems (language, culture and varying regulatory frameworks).

Traditional rankings are often characterised by non-transparent, unspecified and volatile procedures in terms of indicator construction, calculation and aggregation.

We argue that an enhanced understanding of diversity in the profiles and performances of higher education and research institutions at a national, European and global level requires a radically different transparency approach that addresses the major shortcomings of the traditional ranking
Rankings: a mixed blessing

Given this argument, rankings should be based on the interests and priorities of their users.

Instruments. To this end we have, over the last seven years, developed and implemented an alternative ranking tool: U-Multirank. The next sections describe how this new instrument was designed and its various features.

Design principles

Based on clear epistemological and methodological positions we have formulated a set of design principles which has created the basis for the new, multi-dimensional ranking tool U-Multirank (Van Vuith and Ziegeler, 2012).

- The most fundamental epistemological argument is that as all observations of reality are theory-driven (formed by conceptual systems); an ‘objective ranking’ cannot be developed. Every ranking will reflect the normative design and selection criteria of its constructors.

- Given this argument, rankings should be based on the interests and priorities of their users – they should be user-driven. This principle ‘democratises’ the world of rankings by empowering potential users (or user groups). Different users and stakeholders should be able to construct different sorts of rankings, according to their own wishes and priorities.

- Another important principle is multi-dimensionality. Higher education and research institutions are predominantly multi-purpose, multi-mission organisations involved in different mixes of activities. Rankings should reflect this multiplicity of functions and not focus on one function only (research), to the exclusion of all else.

- An obvious corollary to this principle is that institutional performance on these different dimensions should not be aggregated into a composite overall ranking. There is neither theoretical nor empirical justification for assigning specific weights to individual indicators and aggregating them into composite indicators.

- A further design principle concerns the comparability of institutions. In rankings, institutions and programmes should only be compared when their purposes and activity profiles are sufficiently similar. Comparing institutions or programmes that have very different purposes does not create better transparency. In our view this principle implies a two step process: first, institutions with similar profiles have to be identified by ‘mapping’ their activities; secondly, rankings can then only be created for institutions with similar ‘activity profiles’. This connects the description of the horizontal diversity of activity profiles to the assessment of the vertical diversity of performance profiles.

- Higher education rankings should reflect the multi-level nature of higher education. With very few exceptions, higher education institutions are combinations of stronger and less strong faculties, departments and programmes. Producing only aggregated institutional rankings disguises this reality and does not produce the information often most valued by students, potential students, their families, academic staff and professional organisations. This does not mean that institutional level rankings are not valuable to other stakeholders and for particular purposes. The new instrument should allow for comparisons at the level of the organisations as a whole but also at the level of the broad disciplinary fields in which they are active.

- Finally, we include the principle of methodological soundness. The new instrument should refrain from methodological mistakes such as the use of composite indicators, the production of league tables and the lack of transparency of methods. In addition, it should minimize the incentives for strategic behavior on the part of institutions to ‘game the results’.

During the last years some traditional rankings have started to accept some of these principles (which we formulated seven years ago), for instance by making the indicators behind their composite scores somewhat more transparent, by addressing the field/disciplinary level and by publishing rankings with a specific user-focus. However, they continue to focus on research, to produce league tables and to patronize users with composite scores.

Innovative and transparent

The design principles have underpinned the conceptualization and design of U-Multirank, which is a user-driven, multi-dimensional and methodologically robust transparency tool. This instrument enables its users to identify comparable institutions and programmes, to create both institutional and subject level performance profiles of individual institutions, and to undertake comparative performance analyses of institutions and programmes.

In operational terms U-Multirank consists of:

- Five performance dimensions (teaching and learning, research, knowledge transfer, international orientation, regional engagement).
Rankings: a mixed blessing

- A range of indicators that are used to compare institutional performance on these five dimensions at the institutional and/or subject level.

The selection of these dimensions and indicators has been based on two processes:

- **Stakeholder consultation process**; a strong stakeholder orientation is a cornerstone of our approach given the centrality of the principle of rankings being user-driven. The process is never complete as we are continuously looking to develop new indicators and to optimise data presentation according to users’ needs.

- **Methodological analysis** of the validity of the selected indicators, the reliability of the information to be gathered, and the expected feasibility of the use of the dimensions and indicators (availability of data; the extent of extra data collection required from institutions). These methodological issues have also been continuously addressed during the last years. As a result, some of the original indicators were discontinued while for some others the definition has been adapted.

During the design process all potential dimensions and indicators were clearly described and discussed in stakeholder workshops. After a first validity and reliability check, we suggested comprehensive lists of possible indicators derived from the literature and from existing practice (including from areas beyond rankings). In addition we designed a number of new, sophisticated indicators, particularly bibliometric indicators for the research and knowledge transfer dimensions. Currently U-Multirank includes 35 indicators at the institutional level and up to 36 indicators at the subject level across the five dimensions. All indicators are size-normalised.

Given its design principles and its methodological position, U-Multirank clearly differs from the traditional global rankings. It does not aggregate data into composite scores, so it does not need any weighting of data. It uses ordinal scales with bands, called the ‘grouping method’, so it is not confronted with measurement problems at ratio scales. U-Multirank also is completely open about its data use. The data included in U-Multirank are drawn from a number of sources: data from international bibliometric and patent databases and other publicly available sources, information provided by the institutions themselves and surveys completed by more than 100,000 students at participating universities – one of the largest international student samples in the world. Some performance measures (indicators) within U-Multirank have to rely on information collected directly from the institutions as this information is not available in international or national databases. Unlike several traditional rankings, U-Multirank does not use any ‘reputational data’ collected by surveys in which academics are asked to ‘assess’ the reputation/quality/prestige of individual institutions.

**Carefully checked**

Self-reported data in U-Multirank are carefully verified. Data verification in U-Multirank includes several steps and procedures; data delivered by institutions are tested for consistency and plausibility and ‘outlier’ results are carefully checked. The process includes both ‘manual’ and automated checks and a series of direct communication loops with institutions, also regarding contextual details.

In addition U-Multirank uses data from existing (multi) national databases such as the European Tertiary Education Register (ETER) in the EU, the Integrated Postsecondary Education Data System (IPEDS) in the USA and the Higher Education Statistics Agency (HESA) database in the UK. These data are carefully checked for their comparability with the U-Multirank indicator definitions and then used for the ‘pre-filling’ of the questionnaires that are being sent to the institutions for verification and commentary. The major reason for using these databases is to minimise the administrative burden of institutional data collection. On the basis of data gathered on the various indicators across the five performance dimensions, U-Multirank provides its users with the online functionality to create two general types of rankings:

- **Institutional rankings**: rankings on a selection of the indicators of one or more of the five performance dimensions at the level of institutions as a whole;

- **Field-based rankings**: rankings on a selection of the indicators of one or more of the five performance dimensions in a specific field/discipline in which institutions are active.

In 2010/11 U-Multirank was piloted in a worldwide project involving around 200 universities. Since 2014 U-Multirank has published its global rankings every year. The fourth release (2017) features nearly 1,500 universities from 99
Rankings: a mixed blessing

countries, with more than 3,280 faculties/departments and over 10,500 study programmes. It includes over 42,000 performance scores at the institutional level and another 65,000 in 13 subject areas. Given these numbers, U-Multirank has quickly become the largest global ranking tool and performance database in the world.

In addition to its yearly release, U-Multirank regularly publishes so-called ‘ready-made rankings’. These ready-made rankings are offered as example sets of rankings for which the U-Multirank team itself acts as the ‘user’ and thus selects a specific subset of indicators from several dimensions. Currently there are ready-made rankings for ‘research and research linkages’, ‘economic engagement’ and (at the field level) for ‘teaching and learning’.

A crucial feature of U-Multirank is its interactive webtool (see: www.umultirank.org). The basic idea of the webtool is to create an intuitive and flexible instrument for users to work with, instead of a static presentation of information. It offers different tracks into the wealth of data, taking into account the typical searches different stakeholder groups have indicated their interest in. The ‘for students’ track helps students to find universities that perform best in the areas that matter most to them. Students can search for the institutions and programmes that best ‘match’ their individual wishes and priorities. The ‘compare track’ is an entry point into the system for university leaders and administrators who want to analyse and benchmark their university’s performance with other institutions. Accessing the webtool using the ‘at a glance’ feature directly leads to the detailed performance information about a single university, as well as links, maps and qualitative statements by the university about the features of their programmes. On these institutional pages can also be found the U-Multirank ‘sunburst charts’, a graphical overview of all the indicators for a university for which there is information available in the database.

The webtool offers special ‘how-it-works’ videos explaining the use of the different tracks. Furthermore, the website includes the indicator handbook, the methodological descriptions and the questionnaires, guaranteeing the full transparency of the instrument. Since 2016 U-Multirank has also been available for tablets and smartphones. Although basically the same information and functions are offered as in the webtool version, this mobile format allows for a faster and a ‘snapshot’ approach to the rich U-Multirank database.

**A stakeholders’ instrument**

Because of the way it is designed, U Multirank does not have an implicit preference for specific university performance profiles, for instance those of research-intensive universities. In U-Multirank research performance is one of the five dimensions on which performance can be shown (using various bibliometric and other indicators). However, the tool itself does not include any ordering of dimensions, which implies that research performance is not assumed to be a more relevant performance category than other dimensions. And within the dimensions different indicators are available to users. The 2016 release for example showed three well-known US institutions as the best performers on the ‘citations rate’ indicator but three European Universities of Applied Sciences as top performers on the ‘co-publications with industry’ indicator.

U-Multirank is a transparency tool for various categories of stakeholders. The first three releases of U Multirank show that the tool indeed is able to serve the needs of a variety of stakeholder groups.

- U-Multirank is a transparency tool for students (and their parents). It allows them to look for specific institutions and programmes, and to compare these with a set of criteria based on their own preferences and priorities.
- U-Multirank allows different types of universities to show their profiles and specific strengths. It offers the opportunity to select and use the specific indicators relevant to them for benchmarking, strategic decision-making, quality management, networking and/or marketing.
- U-Multirank can also be used as a policy tool at a higher education system level. Governmental policy-makers, university associations/consortia/networks and client organisations (including student organisations) at regional, national and international levels can obtain and analyse data in order to develop their policies and arguments.

**Further developments**

U-Multirank sees itself as a (not for profit) public service tool. It does not have the intention to sell information, consultancy services or newspapers. We want to continuously
To create this ‘public good’, U-Multirank will continue to be funded by the European Commission.

provide a comprehensive global performance database, freely accessible to stakeholders to serve their needs. To create this ‘public good’, U-Multirank will continue to be funded by the European Commission and additional philanthropic actors.

U-Multirank will continue to develop. Important aspects for future development are further increases in the global participation of universities; expanded usage of publicly available databases; continuous development of data collection instruments, indicators and presentation modes; and intensified communication with users and institutions.

Frans van Vught
is a high level expert at the European Commission. Previously, he was Rector and chairman of the board of the University of Twente and director of the Center for Higher Education Policy Studies (CHEPS)

Frank Ziegele
is executive director of the Centre for Higher Education (CHE) in Germany and Professor for Higher Education and Research Management at the University of Applied Sciences Osnabrück

Frans van Vught and Frank Ziegele are the joint project leaders of U Multirank, a project sponsored by the European Commission and several international Foundations.

References