Chapter 11
Concluding Remarks

Frans A. van Vught and Frank Ziegele

In this book we have addressed the general topic of rankings in higher education and research as well as the development of a new multidimensional ranking tool. We looked at the various issues surrounding the ranking debates, and analyzed current practices and their impact. We have been critical of some of the current ranking practices and methodologies and have developed our own approach. In Part I of this volume we discussed the current practices in general and drew a number of conclusions with respect to a new and better methodology. In Part II we expanded on this new approach, which we call U-Multirank. U-Multirank is intended to address the weaknesses in the existing approaches and to offer a multidimensional and user-driven perspective to ranking. We present U-Multirank as a new ranking tool, completely different from existing global ranking instruments.

This book is the result of almost 2 years of intensive work on all facets of international rankings by a team of researchers who conducted the analyses of current ranking approaches and designed and tested the alternative new multidimensional instrument. Several have also contributed to this volume, in which ranking issues are addressed on three levels:

- We analyzed the 'state of the art' of existing rankings, identifying their features, strengths and weaknesses as well as their influence.
- We drafted a new concept for international rankings, labeled 'U-Multirank'.
- We carried out empirical testing of the new multidimensional concept via a worldwide pilot study.

F.A. van Vught (✉)
University of Twente, Enschede, The Netherlands

F. Ziegele
Centre for Higher Education (CHE), Gütersloh, Germany

F.A. van Vught and F. Ziegele (eds.), *Multidimensional Ranking: The Design and Development of U-Multirank*, Higher Education Dynamics 37,
DOI 10.1007/978-94-007-3005-2_11, © Springer Science+Business Media B.V. 2012

This final chapter presents some concluding remarks on ranking in higher education and research in general as well as on the applicability of our new multidimensional approach.

In any ranking the basic normative ideas should be made transparent. We have formulated a set of normative positions for our specific approach to ranking: user-drivenness, multidimensionality and multileveledness, a participative approach.

In the introductory chapter we described our epistemological and conceptual normative ideas regarding ranking. We introduced three basic ideas.

First of all we suggested that in our view there is no such thing as 'an objective ranking' and that the notion of what should be seen as 'good performance' behind any ranking is always related to the subjective assumptions of the ranking producer. These subjective positions about what is and what is not 'good performance' are not always transparent in existing rankings, leading to the risk that the subjectivity is hidden and a false impression is created of a so-called 'objective performance list'.

A hypothetical solution would be to create and accept an 'authority' that would define the 'right' indicators following the idea of an ideal university. However, this proves to be impossible in higher education, since the diversity of university profiles and the diversity of stakeholders' preferences doesn't easily allow consensus about a definitive set of criteria defining the best performance for all stakeholders. The only way to deal with these diversities is to take the normative position of a user-driven approach, accepting the subjective character of a ranking as a design principle and leading to the empowerment of its users This also implies a multilevel approach: some situations in which stakeholders' decisions could be supported by rankings refer to the institutional and some to the field level.

The user-driven approach does not exclude the option that certain 'authorities' would create their own rankings, claiming that their choice of indicators reflect the most relevant aspects of performance in higher education and research. As a matter of fact these 'authoritative rankings' are a special form of the application of the principle of 'user-drivenness', allowing specific organizations, representative bodies, client groups or institutions to present their specific normative positions as convincing and attractive views on what should be seen as relevant and less relevant performance.

Our analysis of the existing global rankings showed that these rankings only cover a small percentage of the total number of higher education institutions worldwide Moreover, they only address a very special higher education institution profile: the 'globally active, comprehensive, research-intensive university', which is presented as the most attractive general 'world brand' because of its research-based performance and reputation in the international context. All other institutional profiles are not addressed in these current rankings, simply because their characteristics are not covered by the indicators applied.

To make up for this deficiency – and as a second normative starting point – we suggest taking a multidimensional approach to ranking. A multidimensional approach allows a large variety of institutional profiles to be included in rankings, thus paying attention to the horizontal diversity of institutional missions and profiles. In addition, the multidimensional approach offers the opportunity to distinguish the various 'functions'

of higher education and research institutions and to assess the performances according to these various functions, rather than forcing institutions to all strive towards a dominant profile of research-intensiveness. Finally, the multidimensional approach opens up the possibility to compare sets of institutions with similar missions and profiles, which appears to be more useful than ranking institutional profiles that are very different and can hardly be compared.

A third normative idea behind our views on ranking regards the 'participative approach'. So far a participative approach has hardly been used in global rankings. The idea to involve the users of the rankings in the processes of selecting the indicators and compiling the data is relatively new in the ranking world. We suggest that the application of feedback loops with users leads to a higher level of usefulness for these users, while also creating a better chance of having access to data. Experience shows that stakeholders often have strong feelings about the relevance of indicators, and are eager to interpret the outcomes of rankings in the context of their personal ideas about quality in higher education and research. A participative approach to ranking emphasizes the principle of user sovereignty and stimulates users' reflections on the relative importance of indicators and performances.

We offer our basic normative ideas in order to be as transparent as possible about our views on ranking. These ideas are based on our analyses of the current ranking instruments and their results and impacts. But they remain normative positions; our normative positions.

Quality assurance activities and rankings in higher education and research are related, but not similar.

In our view quality assurance activities and rankings are both transparency tools. Both are information tools designed to communicate information on higher education and research institutions' efforts and performances to external and internal stakeholders. But quality assurance activities first of all aim to provide 'proof of quality' to stakeholders, and their information provision function is secondary to this objective. Rankings (and other transparency tools, like classifications and league tables) are instruments that intend to create transparency about the activities and performances of higher education and research institutions. But, by doing so, these instruments often imply an implicit view on the relevance of the efforts and outcomes of these institutions. As a matter of fact, the choice of indicators, criteria and data presentation modes in transparency tools reflect an, often implicit, definition of quality. This is a main reason why, in our approach, we not only try to be as transparent as possible about our own choices but also emphasize the importance of a user-driven approach: it should be left to the stakeholders/users to decide which indicators, and hence which aspects of quality, should be the focus of a certain ranking.

Quality assurance activities provide 'proof of quality' for two main reasons: accountability and quality enhancement. The accountability function leads to an externally focused perspective on quality assurance, while the enhancement function is mainly internally focused. In both orientations the provision of information of course plays a major role, but this role is largely limited to reassuring stakeholders that the quality is satisfactory (as in accreditation) and/or collegially controlled

(as in peer review systems). An active focus on the support of the decision-making processes of stakeholders is usually not found in quality assurance activities.

Rankings intend to bring transparency to the performance of higher education and research institutions and to provide information on their performance to a variety of stakeholders. As such, rankings are decision support instruments that seek to assist stakeholders in forming their own judgments on the basis of relevant information. Rankings address the problems of information deficiency and asymmetry regarding higher education and research resulting from the fact that, in economic terms, the activities of higher education and research institutions are to be seen as 'experience goods' or 'credence goods'.

Quality assurance activities and rankings are nevertheless clearly interrelated. The provision of information is a major aspect of any quality assurance activity and hence also rankings can play an important role in quality assurance. In particular when external actors are to be involved in judging the quality of performance of higher education and research rankings could become a highly useful instrument. In addition, rankings support the decisions of a variety of clients of higher education and research institutions and thus inspire these institutions to communicate their qualities in the best possible ways. Rankings in this sense stimulate the internal quality cultures of higher education and research institutions, and invite them to present their results according to their specific missions and profiles.

Quality assurance and rankings are not to be seen as competitive transparency tools. They have different functions and orientations, but are also clearly interrelated. Both are crucial instruments for the further development of higher education and research worldwide.

Although several methodological flaws exist in their current applications rankings nevertheless appear to be attractive to many stakeholders and have major impacts.

Our overview and analysis of the state of the art in rankings in Part I of this volume showed that an inventory of the methodological problems regarding rankings produces the following list:

- Rankings are not always clear about their specific clients and target groups. They often appear to assume that whatever information is provided should be relevant to all potential clients. Moreover, regularly the implicit assumption appears to be that the availability of indicators also defines the relevance of indicators.
- Most rankings only address institutions for higher education and research as a whole, and appear to ignore the internal diversity within these institutions. Differences in performance between faculties, departments, centers and other units within the institutions are not taken into account, and neither are differences between academic fields.
- Most rankings appear to focus on a very limited part of the activity profiles of higher education and research institutions, in particular on research productivity and research reputation. At the same time these rankings appear to suggest that they address the overall quality of the institutions, implicitly limiting the concept of quality to the dimension for which (bibliometric) data are most easily available.

The other dimensions of the activity profiles (teaching & learning, knowledge transfer, international orientation, regional engagement) are largely ignored.

- Many rankings provide composite overall indicators in which sets of weighted indicators are combined into a single performance measurement. Composite indicators are highly problematic because they lack the conceptual base from which they should be calculated and its designers cannot provide the theoretical and empirical arguments for assigning particular weights to the constituent parts. In addition, the choice for certain indicators and weights imply an implicit definition of the 'ideal model' of a higher education and research institution. Furthermore, composite indicators appear to be far from statistically robust and they tend to patronize users and clients since, by providing fixed combinations and weights, they imply choices about the relevance and appropriateness of certain indicators.
- The many rankings that provide league tables ignore the statistical problems related to the characteristics of methodological scaling and the existence of standard errors in data. League tables have to assume continuous ratio scales and by doing so exaggerate differences between institutions ranked in these tables.
- Most rankings are unable to address the differences in performance that are the result of cultural, language and other contextual factors. This is particularly problematic in the bibliometric assessment of research performance, where the effect of differences in publication cultures is clearly visible. The existing international bibliometric databases are still facing the challenges of publication cultures that are not focused on traditional academic, international, English-publishing journals, and of including research institutions that are not part of university organizations.
- Rankings often are insufficiently transparent about their methodologies, and regularly appear to adapt these methodologies without being explicit about it. The outcomes of rankings are not always replicable because of methodological and/or statistical changes.

Yet, while rankings are often criticized – and usually rightly so – their impact is nevertheless large. Several categories of stakeholders are heavily influenced by ranking results, although they are not always willing to publicly admit so. Institutional leaders react to the outcomes of rankings in their institutional strategies and communication behavior. Students appear to take ranking results into account when making their choices for enrolling into institutions and programs. Policy-makers use ranking outcomes to design and adapt national higher education and research policies (including funding, merging and excellence policies). Employers appear to pay attention to rankings in their recruitment and contracting policies. Journalists report on ranking outcomes to the general public, thus creating an impact on institutional reputations.

Rankings also have system-level effects. They fuel the higher education 'reputation race'. They create public images of assumed quality. They contribute to academic stratification and institutional wealth inequality. And they trigger institutional behavior of 'gaming the results' (see Chap. 5). The various impacts of the outcomes

of rankings make it clear that there is sufficient reason to take rankings seriously and to try to improve their conceptual and methodological bases.

Improving the current approaches to ranking is highly needed but offers some major challenges.

As just noted, our analysis of the various higher education and research rankings around the world pointed out a number of shortcomings. It also should be noticed, however, that some ranking organizations are taking initiatives that intend to improve their existing methods and to make them more transparent. In addition, the 'Berlin Principles' designed by the International Ranking Export Group (IREG) and the suggestions by a special expert group (AUBR Expert Group) set up by the DG Research & Innovation of the European Commission show that there is an increasing international awareness regarding the need to strengthen the conceptual and methodological foundations of rankings. Multidimensionality and a clear and targeted user-focus are mentioned as important aspects of the further development of ranking in higher education and research.

As may have become clear in Part II of this volume, these new aspects of ranking are not easy to develop. With respect to multidimensionality the challenge is first of all the availability and international comparability of data. If we move beyond the traditional focus on bibliometric data, rankings largely have to rely on institutional data provision. Multidimensional rankings that want to take the variety of institutional missions and profiles into account cannot be realized without the application of institutional and student surveys. Therefore these rankings have to succeed in convincing higher education and research institutions to invest time and energy in data-collection and reporting. This makes multidimensional rankings vulnerable: if they don't see clear benefits from the ranking outcomes, institutions may not be inclined to get involved in data provision.

Another challenge is the potential risk of a limited attractiveness of multidimensional rankings in comparison with monodimensional league tables and composite indicators, particularly to the general public. Simple league tables are often striking, and are easily taken up by the media. Multidimensional rankings that address a variety of target groups may offer more elaborate information, but cannot be reduced to an overall list of winners and losers. Multidimensional rankings need to invest in presentation modes and communication processes, explaining to clients and stakeholders how the various outcomes can be interpreted. In order to be effective in these communication processes multidimensional ranking producers will have to analyze the decision-making processes of user groups (such as students, parents, institutional leaders, policy-makers, business leaders) and the information needs in these processes. In our view, these needs can be revealed by intensive stakeholder dialogue; what we have called 'a participative approach'.

The user-driven approach to ranking presents another specific challenge. If a ranking is based on the user's selection of institutions and indicators, the ranking result is not a unique performance list such as the ones that normally are the outcome of the existing rankings. In a user-driven approach users can produce their own 'personalized' rankings. Eventually these personalized rankings may become

'search engines' that present information ('hits') based on combinations of search terms (indicators). Such search engines will be based on smart technologies (of indexing and storing links) and on the surfing behavior of large numbers of users, resulting in visually attractive and user-friendly information provision. Ranking information will thus become integrated in new communication tools based on internet and social media. The release of a new ranking outcome will not the publication of an updated list, but the integration of a data update in the ranking database, allowing a variety of users to produce a large number of their own personalized rankings in an interactive way.

We nevertheless still call such a multidimensional, user-driven methodology a 'ranking' since it remains a tool to render vertical diversity transparent. Also multidimensional ranking results show high and low performances and position institutions/programs in the context of the performance of their peers and competitors. But multidimensional ranking results also offer differentiated pictures of strengths and weaknesses of institutions and programs. They show differentiated performance profiles to a variety of users.

The challenges of further developing the methodology of ranking in higher education and research are substantial, but – we feel – must nevertheless be addressed. Rankings do exist in higher education, and will not easily lose their impact. Criticism of rankings is relevant, but not sufficient to create better approaches. New instruments must be designed and tested. U-Multirank is the result of such efforts to design and develop a new approach. While U-Multirank cannot immediately resolve all the methodological problems of the current rankings, it at least addresses a number of these challenges.

U-Multirank is a new ranking tool, based on a coherent set of assumptions and ideas regarding multidimensional and user-driven ranking.

U-Multirank is a transparency instrument offering multiple ranking options to users. It is based on our normative positions regarding ranking: user-drivenness, multidimensionality, multileveledness and a participative approach. U-Multirank recognizes that higher education and research institutions serve multiple purposes and perform a range of different activities at different levels. It is a tool that allows a number of different rankings according to the selection of dimensions and indications by users.

U-Multirank is user-driven: it is *you* (the client/stakeholder/user) who is enabled to rank comparable profiles according to the criteria important to *you*. The pilot project during which we designed and tested U-Multirank has specifically been focused on this multiple ranking concept. Taking this concept seriously, we not only distinguished five different dimensions regarding the functions performed by higher education and research institutions, we also addressed two levels regarding these functions (institutional and field level) and incorporated the user-driven approach of a multitude of potential users. The result is a truly multidimensional ranking tool that allows the comparison of a multiple set of different activity profiles, thus creating the possibility for a large variety of higher education and research institutions to compare themselves to organizations with similar or related profiles. U-Multirank

does not limit itself to a single, dominant profile of only one type of higher education institution, i.e. the research-intensive, comprehensive research university. It also allows regionally focused institutions, bachelor degree awarding colleges, polytechnics, art schools, music academies, specialized research centers and many other types of higher educations and research organization to appear in international rankings and to benchmark themselves at an international level with counterpart institutions that may have similar orientations on user-defined dimensions.

U-Multirank intends to serve the needs of a broad variety of users, allowing them to select dimensions and indicators according to their own criteria and preferences. Different users can create their own 'personalized rankings' focusing their own specific rankings at the topics regarding higher education and research that they judge to be most relevant. In addition, U-Multirank offers the option to present 'authoritative rankings', in which a specific selection of dimensions and indicators is pre-defined and selected on the basis of the 'authority' of a certain organization, institution, association or network. Authoritative rankings can be produced and published on behalf of higher education membership organizations, specific associations of higher education institutions, national or international public authorities, client representation organizations, independent foundations, etc. The only condition is that these organizations define (and motivate) their selection of dimensions and indicators.

U-Multirank also has an eye for the empirical fact that higher education and research institutions perform differently in different fields. Faculties, departments, centers and various other units within higher education institutions often have their own view on relevant performance in their specific disciplinary or interdisciplinary fields. U-Multirank offers the option to produce rankings at two different levels of activity, the institutional level and the field level. By doing so, U-Multirank addresses the internal diversity in higher education and research institutions.

In addition, U-Multirank intends to allow the adaptation of indicators to the specific characteristic of fields. An important aspect of the participative approach is the involvement of field experts and stakeholders in the process of defining and selecting indicators for field-based rankings.

'Version 1.0' of U-Multirank shows that a multidimensional, user-driven ranking tool is feasible at a global level.

The U-Multirank pilot project proved that a user-driven, multidimensional ranking tool is feasible at world scale. During the pilot project a broad variety of feasibility aspects was explored and tested. We analyzed the conceptual clarity of the sets of indicators; we tested the availability and consistency of data for these indicators. We studied the feasibility of the data collection instruments. And we explored the potential for up-scaling the pilot application to both a global scale and a broad spectrum of fields.

The pilot test shows that the number of feasible indicators is more limited in some dimensions than in other. In particular in the dimensions 'knowledge exchange' at the field level and 'regional engagement' at both institutional and field levels feasible and applicable indicators appear to be only limitedly available. The future challenge certainly is to design and develop more and generally acceptable indicators in these areas.

Regarding the up-scaling to a global level, the pilot project results are encouraging. There appeared to be a strong expression of worldwide interest to participate in the pilot sample, although in some parts of the world the recruitment of institutions for participation in the pilot project proved to be difficult. We concluded that there is a broad stakeholder interest in the further development and implementation of U-Multirank and we expect that substantial numbers of higher education and research institutions from all over the world will be willing to participate in multidimensional global rankings.

The extension of U-Multirank to a broad variety of disciplinary and interdisciplinary fields may also be expected to be feasible. The set of field indicators applied in the pilot study may be regarded as a solid and useful base for such an extension, although it also should be noted that in order to allow a broader coverage of fields, specific field indicators will have to be developed. As mentioned before, for this the participation and commitment of field experts and stakeholders will be highly important.

U-Multirank offers some innovative ideas to the international debate on and the state of the art of ranking.

The characteristics of U-Multirank, in particular its emphases on multidimensionality and a user-driven approach, appear to already have influenced the international debates on ranking in higher education and research. Various other international rankings have introduced new elements into their own approaches that are rather similar to the basic approach of U-Multirank. The expansion of data collection beyond bibliometric data, the development of field-based rankings and the introduction of user-driven weights in indicator selection processes are examples of recent adaptations in existing ranking methods that might be triggered by our U-Multirank methodology. But a coherent and comprehensive ranking methodology that addresses the broad variety of functions of higher education and research institutions, and that allows both personalized and authoritative rankings is so far only found in U-Multirank. U-Multirank offers a new epistemologically sound and conceptually and methodologically transparent approach to global ranking.

In addition U-Multirank brings some specific new elements to the state of the art of international ranking, potentially leading to substantial progress in ranking methods. A first new element is the two-step approach of combining a mapping and ranking transparency tool. By using U-Map, the horizontal diversity of higher education and research systems is addressed and the various activity profiles of higher education and research institutions are made transparent, allowing the identification of institutions with similar or related activity profiles. By applying U-Multirank to groups of institutions with (partially) similar activity profiles multiple rankings of groups of comparable institutions can be created and specific performance profiles can be shown. A second new element regards the design and implementation of a number of innovative bibliometric indicators, analyzing co-publications (of academic organizations with respectively industrial, international and regional co-authors) as a way to report on the performance in the dimensions ‘knowledge transfer’, ‘international orientation’ and ‘regional engagement’. A third new element concerns the introduction of a global student satisfaction survey instrument, which when tested proved to be feasible in a global context. Finally, the introduction

of field-based rankings offers the option to root the rankings in the academic community and to increase their acceptance as relevant and useful transparency tools.

For the further development and implementation of U-Multirank a number of issues will have to be seriously addressed.

Now that 'Version 1.0' of U-Multirank is available, its further development and international implementation can be taken up. However, in order to make an effective international rollout possible, a number of conditions will have to be fulfilled.

First of all, the further development of applicable and widely acceptable indicators will have to be stimulated. In particular in the dimensions 'regional engagement' further discussions and testing will be needed to allow a growing international consensus on feasible indicators. Similarly, at the field level a debate will have to take place on the relevant indicators for 'knowledge transfer'. In addition, in order to allow the expansion of the number of field-based rankings, field-specific indicators will have to be selected and added to the base set of field-based indicators.

Secondly the availability of international comparative data needs to be improved. So far international databases comprise only limited data at the level of higher education and research institutions. Even regarding the crucial dimension of 'teaching & learning' comparable data on for instance labor market success of graduates appear to be nonexistent. A concerted international effort to improve the data-availability will be crucial for the further development of international transparency tools. The international harmonization of data-collection standards, the integration of national databases into joint international databases and the combination of international data-sets are highly important aspects of such a concerted international effort.

Thirdly, 'user-friendly' and attractive presentation modes of the outcomes of rankings will be needed. Both experienced and 'lay' users should be enabled to make use of performance rankings. The presentation modes should include attractive graphical presentations (like the 'sunburst chart' applied in U-Multirank) and make use of symbols and colors (like in the 'grouping approach') to create clear and coherent impressions at first glance. A web-application should provide clear guidance and explanation, and in particular address the needs of specific user-groups. A differentiated information provision format should be an integrated part of the web tool. The presentation modes should refrain from simplistic and risky methods (like league tables) and be based on sound methodological principles.

Fourth, given the fact that international databases are limited to bibliometric and patent data, data-collection from higher education and research institutions will remain necessary. Data delivery should therefore be sufficiently attractive for these institutions. The costs of collecting and delivering institutional data should be out weighted by their benefits such as the ranking outcomes. On the costs side, 'prefilling' of questionnaires with externally available data and coordination of data collection processes (now often organized as separate tracks) will reduce the workload for the institutions involved. On the benefits side, offering benchmarking opportunities with comparable institutions and tailor-made ranking outcomes applicable in internal planning & control processes may stimulate the willingness to deliver data.

Generally speaking, for institutional data-collection to be successful the organization of the data-collection processes should be clearly focused on the costs/benefits balances of the higher education and research institutions involved.

Finally, a crucial condition for a successful international implementation of U-Multirank will be its institutionalization. The 'authority' of the actor organizing the ranking processes and the 'ownership' of the data are sensitive issues in the world of ranking and should be carefully approached. In our view, U-Multirank should be independently institutionalized, with extensive advisory and communication facilities for experts and stakeholders. There should be no direct decision-making authority for political bodies, governments or interest groups, and there should be a highly transparent governance structure to safeguard the independent character of the ranking outcomes. Funding could come from independent foundations and from sponsoring public and private organizations, as well as from the sales of standardized products and services (such as data visualization, benchmarking support processes, SWOT analyses). Interested parties could be invited to create and publish their specific 'authoritative rankings'.

The future of U-Multirank and of the further development and implementation of multidimensional ranking in general to a large extent depends on how the various issues just mentioned will be addressed. Multidimensional and user-driven rankings in higher education and research have been proven to be feasible. The coming years will show whether they will also be internationally realized.