Liu, N. C. and Cheng, Y. (2005). "Academic Ranking of World Universities: Methodologies and Problems". *Higher Education in Europe* 30 2, pp. 127-136.

Merisotis, J. P. (2002). "On the Ranking of Higher Education Institutions". *Higher Education in Europe* 27 4, p. 361.

van Raan, A. F. J. (2005). "Fatal Attraction: Ranking of Universities by Bibliometric Methods". *Scientometrics*, 62 1, p. 133.

The CHE-Ranking of German, Swiss and Austrian Universities

By Detlef MÜLLER-BÖLING and Gero FEDERKEIL

Introduction

This study describes the ranking methodology developed by the Centre for Higher Education Development (*Centrum für Hochschulentwicklung* – CHE) in Gütersloh, Germany, which started comparing German universities in 1998. In the context of the Bologna Process, by 2010 with the goal of creating a European higher education area the need for comparative information on European universities is growing. Therefore, the CHE decided to internationalize its ranking. The strategy and methodology of this internationalization is also described in this present study.

Quality assessment has become one of the most prominent discussion topics on higher education, both for science and higher education policy. The necessity for quality assessment becomes all the more evident when the following points:

- considering competition among universities has significantly increased, both on a national and an international scale; and,
- diversification in the higher education sector has created an incredibly rich variety offering of courses, programmes and degrees, again both on a national and an international scale. (In Germany, for example, there are about 9,000 degree programmes in higher education.)

Academic rankings and league tables are an important and useful instrument to create transparency in what one might be tempted to call

the "university jungle". Rankings are a way of compiling information about universities, programmes and research as well as teaching activities in order to provide orientation to specific target groups, ranking from high school graduates who want to go to university, to students who want to change their field of study or their university to members of the department or the university management who want to assess their strengths and weaknesses in order to stay competitive. Today, throughout Europe and the world, a number of different university rankings can be found with different methodologies, scopes and target groups – and of different quality, too.

In order to satisfy the variety of needs and expectations, rankings must be based on a scientifically founded methodology. Validity and reliability of data are indispensable for serious and honest rankings that merit publishing and consulting.

The CHE-Ranking of German Universities

As with other instruments and procedures of quality assessment and evaluation, Germany was when a latecomer with regard to rankings, too. For decades, the German higher education system has cultivated the myth that all universities are of equal quality. Together with a strong notion of university autonomy, this is the main reason why Germany remained a latecomer in the "quality assessment movement" in higher education. When other countries were already characterized as "evaluative states" (Naeve 1988), evaluation was still new territory in Germany (Cave *et al.*, 1997). Up to the 1980s many stakeholders within the higher education sector opposed by notions of competition and quality assessment. However, in a period of tight resources, issues of accountability, competition and quality control gained more public attention. At the same time, there was a growing sense of differences in quality between German universities, which in the beginning were discussed in terms of "profiles". But it was no earlier than 1989 that the first broader ranking of German universities was publishing, by the weekly magazine *Der Spiegel*, asking "Which university is the best?" During the 1990s, a number of other magazines started rankings of higher education institutions, some for single subjects only.

Following these initiatives, the CHE started its ranking after a two-year period of preparation in 1998 in co-operation with the *Stiftung Warentest*, a national consumer protection foundation testing goods and services. From 1999 to 2004, the ranking was published in co-operation with the weekly magazine *Stern*. Since 2005, it has been published in co-operation with the weekly newspaper *Die Zeit*, an important public space for discussing issues of science and higher education. Competencies are strictly separated between the two partners: the CHE is exclusively responsible for the concept and the data, whereas the *Die Zeit* holds responsibility only for publication, marketing and distribution.

The CHE was founded in May 1994 by the German Rectors' Conference and the Bertelsmann Foundation. The Centre's purpose is to initiate and assist reform in Germany's institutions of higher education. The CHE defines itself as a "think tank" and consulting group for higher education. As a non-profit institution, the CHE formulates impartial political objectives, develops integrated concepts, existing options for future development and explores through pilot projects in close co-operation with academic and government institutions. Creating transparency about German universities by means of a ranking was one of the major founding tasks of the CHE.

The CHE-ranking focuses on selected academic subjects offered by a substantial number of universities: it includes about 36 subjects, which are updated in clusters within a three-year cycle. Starting with economics and chemistry in 1998 and with the last extension into geography/geology in 2005 the ranking will cover the subjects of almost 80 percent of all beginners at German universities. Besides a print version of selected results, all results and all indicators are presented in the internet-version of the ranking, which is freely accessible (www.che-ranking.de)[1]. Here, the interactive possibilities of the medium can be used to make a personal ranking by selecting and weighting indicators according to individual priorities and preferences. Since the ranking's 2002 edition, comparisons over time can also be made.

Methodological Principles of the CHE-Ranking

Target Groups of the Ranking

One of the first prerequisites for designing a ranking is to reach clarity concerning the objectives and hence the main target group(s) of the ranking, for this has decisive implications for the ranking design and the presentation of results. In most rankings, the main target groups are university entrants – with perhaps different factors influencing their decision – "degree movers" (in Germany typically after the *Vordiplom* (first level graduate degree), *i.e.*, after two years) as well as students looking for a post-graduate degree. And, of course, universities themselves are also users, if not exactly a "target group" of the ranking. But while universities are interested in detailed and highly sophisticated information particularly on research, German university entrants are confronted with some 9,000 courses at more than 300 universities. Therefore, academic rankings have to find a balance between diverse expectations. The orientation towards university entrants has implications concerning the concept of a ranking, the indicators and the way of presenting results.

General Approach

In three central aspects, the CHE-ranking follows methodological principles that distinguish it from many other ranking approaches.

- The main target group of the ranking is university entrants. They have to decide on a specific subject or programme at a university, not so much for a university as such (it should be kept in mind that there are no Liberal arts colleges and Bachelor's programmes in Germany). Therefore, the ranking does not rank whole universities, but strictly refers to single subjects. This approach is supported by the theoretical argument that universities comprising many disciplines and programmes are far too complex to be ranked as a unit. In addition, empirical evidence suggests that there are great differences in performance between different subjects within a university. A university may rank high in physics and at the same time rank very low in literature. The information, that a specific university is ranked in the middle, which inevitably will be the result of ranking the whole university, would not have any relevance to a freshman in physics. And, it would be of no use for universities in terms of analyzing their strengths and weaknesses, too.

- Moreover, even within a single subject, the CHE-ranking does not calculate an overall value out of single, weighted indicators. There is neither a theoretical nor an empirical basis for such weighting procedures. With regard to the orientation towards the students as the main target group of the ranking, the heterogeneity of preferences within the target group has to be taken into consideration. Some students are looking for a university with high research activities (measured by research grants, publications, citations, etc.), while other students may look for a small university with close contacts between students and teachers, good mentoring and a short study duration.[2] Calculating an overall score, which inevitably has to give general weight to single indicators, is to patronize the target group. The internet with its interactive features offers new opportunities to the users of rankings: in the CHE-ranking, users can make their own personal ranking by choosing and weighting indicators according to their own preferences.

  Furthermore, calculating an overall score ignores the fact that within single subject universities have different profiles with specific strengths and weaknesses that will be overlooked by an overall score. University "A" may give particular attention to good tutoring and mentoring and close contacts between professors and students; university "B" may have strength with regard to an early introduction of students to research. By using an overall score a ranking will either produce a judgement with regard to which profile is more valuable or – by equal weights to both aspects – level-out profiles to average values. Hence, the CHE-ranking is multidimensional by ranking each indicator separately and leaving the decision about the relevance of an indicator to the user.

- Most rankings order universities in league tables with individual rank positions. This approach suggests that each difference in the numeric value of an indicator marks a difference in the entities ranked. This inevitably involves the danger to misinterpret small differences in the numeric value of an indicator in terms of differences in performance or in quality. For example, in the 2001 edition of the *U.S. News and World Report* ranking of national universities, the difference between rank-13 and rank-22 is only six points on a 100-point scale. In many cases, data are not precise enough to establish clear cut and unambiguous league table positions in a reliable way. Alternatively, in statistical terms, such a procedure ignores the existence of standard errors. Instead, the CHE-ranking orders universities in three groups: The best universities are ranked into the "top" group, the worst into the "bottom" group and the rest is considered to be in an "intermediate" group.

The grouping procedure varies according to two categories of indicators: facts (as for example staff/student ratios or number of publications) are grouped according to quartiles. The upper quartile and the lowest quartile are ranked first respectively bottom group, the middle-two quartiles are ranked intermediate. In the case of subjective indicators based on survey data, *i.e.*, judgments by students and professors, the grouping procedure takes into account the diversity of judgments within universities compared to the overall score and the number of cases. A university is ranked "top" if the confidence interval of the mean (using a scale from 1 "very good" to 6 "very bad" corresponding to German school marks) is completely below the overall mean of all universities (in a particular subject). At the other extreme, a university is ranked "bottom" if its confidence interval is completely above the overall mean. Accordingly, a university is ranked into the middle group, if the mean is "intermediate" or if judgments are controversial, *i.e.*, dispersion is high, making the confidence interval large and hence being neither completely below nor completely above the overall mean.

Indicators

Crucial to rankings is the choice of indicators. Rankings can be distinguished according to the data sources they refer to or to the quality (relevance, validity) of indicators.

Indicators should be of relevance to the target group(s) of rankings. In an almost two year preparatory phase, CHE tried to identify relevant indicators with the help of an advisory board (including evaluation experts and members of professional and university associations) and by group discussions with school-leavers and students. Those discussions were repeated regularly in order to adjust indicators to changing demands for information within the target group. Empirical evidence of other studies suggests that the CHE-ranking covers all issues important to students for their choice of a university.

Out of this process, a "model of decision" was derived containing nine components relevant in the decision process (Figure 1). Each component comprises several indicators, all in all some 35 (depending on subjects). The components range from general information on towns (*e.g.*, mean rents) and the university (size, year of foundation, type), student characteristics (*e.g.*, male-female, foreign), central issues of courses and teaching (course organization, mentoring, libraries, laboratories, etc.), aspects of employability and research to some overall judgments by professors (including reputation indicators) and students.

Figure 1. A "Model of Decision" in institutions

students	**study outcome**	**internationalization**
study and teaching	**resources**	**research**
employability	**overall assessment (students, professors)**	**city and university**

The CHE-ranking follows a multi-perspective approach. First, each component is comprised of indicators from *different data sources*. To take

"research" as an example, calculated indicators based on data delivered by the faculties (*e.g.*, research grants, number of PhDs) were taken in making a bibliometric analysis on the basis of various data banks (for instance, the *Science Citation Index* and *Social Citation Index*, but also some specific German databases for single subjects) and indicators based on the professor survey (*e.g.*, research reputation) were used.

Second, the set of indicators comprises *facts* as well *subjective judgments*. In the component "teaching" for example, there are factual indicators as student/staff ratios or mean study duration (which varies tremendously between German universities, in some diploma courses up to three years) as well as judgments of students and alumni, *e.g.*, on course organization, contact between professors and students, libraries, computer facilities, etc.

Regarding the international comparison of rankings and their indicators, one important issue to be addressed: *The appropriateness of indicators is heavily dependent on national higher education systems*. An indicator may be highly expressive in one country and make (almost) no sense in another. The study duration is one example: in the German higher education system with the (in theory) five-year Diploma degree, mean study duration differs tremendously between universities and hence is an important factor for the decision, for university. In sociology, *e.g.*, the median ranges from nine to 16 semesters, which means that half of the students at some universities take three-and-a-half years longer to obtain their degree than at other universities. In a higher education system with structured courses in which the target study duration is also the empirical norm, this indicator does not have much validity.

Another example is the selectivity of universities, usually measured by the percentage of applicants accepted. This is important information giving an insight into the degree of elitism of a programme/university and, from the student's perspective, of the probability to be accepted. In the German higher education system, universities have no right to choose their students. The number of places in a programme is determined by state – to say bureaucratic – rules with almost no scope for the universities. Either there are enough places for all applicants, and then all applicants will be accepted. On the other hand, there are not enough places (as *e.g.*, in medicine and psychology), then a central authority, the

Zentralstelle für die Vergabe von Studienplätzen (ZVS – Central Office for the Allocation of Places of Study) distributes applicants to universities. Besides school marks, the nearness of a university to the applicant's place of residence is the most important factor. Recently, universities were allowed to select a portion of students: But this has been a "negative selection" process, as universities could choose applicants only after the first 80 percent were distributed by the ZVS. Accordingly, only a small minority of universities made use of this possibility – a good example of an unproductive reform.

Data Sources

As the CHE-ranking follows a multi-perspective approach, data are collected from different sources in order to allow different points of view. Wherever possible, data from one source are counterbalanced by respective data from other sources. This is particularly the case of subjective opinions which are contrasted by facts whenever possible. The ranking is based on several data sources and surveys:

- *Universities*: Information on fees, accommodation, students, central services (up to now more than 300 universities have been included).
- *Departments*: Information on programmes, teaching, staff, research, resources (up to now about 2,000 departments have been included).
- *Professor survey*: Each year questionnaires are sent to all professors in the disciplines involved that year (until 2005, almost 30,000 professors responded). The survey asks for the universities with the highest reputation in a field with regard to study and to research.
- *Student survey*: A random sample of 300 students per degree programme and university is included. Up to now, questionnaires of about 200,000 students have been analyzed. Students were asked about teaching, study organization, mentoring, libraries, IT and so on. In addition, they are asked for some information on living conditions.

- *Bibliometric analyses* are carried out in those subjects where adequate data are available. Methods and databases vary according to subjects. Whereas in the sciences the "Web of Science" (SCI, SSCI) is an accepted database, in other disciplines, *e.g.*, the humanities, specific databases that are sometimes restricted to national publications must be considered.
- *Analysis of patents*: In engineering the CHE-ranking made the first comprehensive and comparative analysis of patents of German universities.
- *Graduates*: At the moment, the authors are aiming to carry out comparative graduate surveys. German universities do not have good alumni contacts. Most universities do not have systematic information on their alumni, nor do they pursue addresses of their alumni systematically. In 2003, the CHE conducted a nationwide survey in medicine where the sample (of five graduation cohorts) was drawn by Medical Chambers, from which each person who wants to work as a physician has to register. In 2004, the authors conducted a survey in business studies.

Quality Assurance

Rankings may be used by universities as an instrument of quality assurance but they themselves have to put into place mechanisms of quality assurance. The issues are, in order of importance, transparency of the methodology, data sources, and ranking procedures. Details on the CHE-ranking can be read in a methodology report (Berghoff *et al.*, 2006).[3]

In the process of data collection, CHE sends all data delivered by universities and departments back to them for control before computing indicators. Hence, universities have the possibility to complete missing data and to correct misinterpreted data.

Each year the university establishes an advisory board for the respective subjects. Members are representatives of faculties' associations and/or of subject associations. The board gives advice on specific particularities of the subjects regarding indicators, questionnaires and other methodological issues. Furthermore, the board's experience and

knowledge is used for checking the plausibility of results. This leads to the last instrument of quality assurance, which is represented by plausibility controls. In addition to the professional experience of the advisory board, statistical procedures are employed, in order to identify extreme cases and inconsistencies in the data. The established policy is to publish only valid and reliable data. Consequently, much data were collected that, in the end, are not published.

Presentation of Results

In order to get a high circulation of the ranking, results are published in co-operation with a media partner, the weekly magazine *Die Zeit*, which, since 2005, has had a high acceptance within German academia. There is a strict division of competencies: the responsibility for the concept, methodology and data of the ranking is held exclusively by CHE, whereas the *Die Zeit* is responsible for publication, distribution and marketing.

Selected and exemplary results are published in the weekly print version of the *Die Zeit*, together with a yearly special issue called "Student's Guide". All data with detailed information on universities and programmes are freely accessible in the Internet version of the ranking (http://www.das-ranking.de/che6/CHE6). On the Internet, users can interactively produce their "personal ranking" by selecting and weighting indicators according to their own preferences.

Effects of Ranking

Effects of rankings may refer to students as well as to universities themselves. The effects of the ranking on students are considerable. According to survey data, about a third of prospective students use rankings for orientation, which is a considerable fraction in the German context, where one still finds the myth that "all universities are equal" and where rankings, as a result, were not accepted in the scientific community for a long time. The ranking helped to raise awareness of the differences in the quality of teaching and research. The proportion of students using the ranking varies across subjects: it is higher in

engineering, medicine and law and lower for students in the humanities. Generally, it can be said that it is particularly the achievement-oriented students who make use of the ranking.

A good example of effects can be shown for psychology, which was first included in rankings published in 2001: in the year after publication, the number of applications at the recommended universities increased notably while it remained stable on the whole. The rise was about 19 percent for the universities that had been recommended for the "researcher" student (who wants to get the most of our available courses and professors) and about 13 percent for the "normal" student (who wants to study rapidly and efficiently with adequate monitoring). What was also observed was that good ranking results had more effects on applications than bad ones.

On the institutional level, it was observed that universities and departments take the ranking as a starting point for the analysis of their strengths and weaknesses. In this context, the authors offered detailed analyses of the student survey for single departments that went beyond the published indicators. After a first phase in which poorly ranked departments often expressed fundamental criticism of the ranking, now many positive replies have been received, even from those departments who came off badly (or at least by some professors or vice-deans who are engaged in matters of teaching) telling the authors that they want to make use of the results for an analysis of problems and for reforms.

Internationalization of Ranking

In the context of the Bologna Process, student mobility within Europe is growing and will probably grow further in the next years. Accordingly, information for students about programmes in an international perspective will become more important. In an intermediate perspective, CHE is striving for a European ranking. In a first phase, the universities of Austria and Switzerland have been included in the ranking as their higher education systems are most comparable to the German system with regard to programme structures and – at least for the majority of Swiss universities – also with regard to language.

In 2002, a first pilot project was undertaken to test the methodology to include mathematics and chemistry at Austrian universities; the results were not published. In 2003 for the first time, a comparative ranking with results for English/American Studies and Electrical Engineering at Austrian universities was published, and co-operation with Swiss universities also begun, initiated by the Swiss Rector's Conference. With the 2005 ranking, Austrian universities have been included in the normal cycle of the CHE-ranking. Also in this year, business studies, economics, sociology and political sciences in both Swiss and Austrian universities have been included into the normal cycle of the ranking. Recently, results of the 2006 ranking exploring sciences and medicine were published.

The internationalization strategy is determined by two goals:

- first, the ranking should gain a high acceptance within the higher education system of the countries included;
- second, the comparative ranking must – concerning its methodology and the choice of indicators – take into account specific characteristics of the higher education systems and the academic culture of the respective countries. Otherwise, the comparison will not be able to produce valid information on those countries. For example, one has to check carefully the availability of adequate databases for comparative bibliometric analyses in order to avoid biases disadvantaging any country.

To reach both goals, CHE is co-operating with qualified partners in Austria and Switzerland, who have solid knowledge of the higher education system of their country. Furthermore, CHE lays stress on a strong commitment to the ranking by the national universities associations. The idea of a joint ranking is supported by the Austrian Ministry of Higher Education as well as by the Austrian Conference of University Rectors respectively by the Swiss Ministry and the Swiss Conference of Universities.

To reach these goal concepts, indicators and questionnaires had to be adapted to the Austrian and Swiss higher education system. This adaptation and data collection were done in close co-operation with two partners, the Austrian Quality Assurance Agency (AQA; <http://www.aqa.ac.at>) and Swiss Up: Foundation for the Excellence of Education in Switzerland (www.swissup.com). In both countries, national advisory boards were set up

to adapt methodology, indicators and questionnaires to the national situation.

By this approach, the CHE-ranking differs from those "world rankings" that mix and analyze commonly available data on different countries without regard to differences in the structure of higher education or in academic cultures.

Finally, the aim is to produce a European ranking of universities. As a first step to further internationalization, the inclusion of selected universities in other European countries such as France, the United Kingdom, Italy, the Scandinavian countries, the Netherlands, Belgium and Luxembourg could be considered in ranking processes. This could lead to a ranking of European "top" universities – but, according to the methodological standards set (to treat single disciplines in a multi-dimensional fashion) instead of simply calculating overall scores and with absolute rank groups instead of league table positions.

Notes

[1] An English translation of the ranking is prepared by the *Deutscher Akademischer Austausch Dienst* (DAAD – German Academic Exchange Service) on <www.university-ranking.de>

[2] Study duration is a very important issue in German higher education. Up to now, the German diploma-degree still predominates. This degree's length is five years–in theory. In practice, however, the average duration in many cases does six or even seven years, with great differences existing between universities.

[3] <http://www.che.de/downloads/Methoden_2065.pdf>.

References

Berghoff, S. *et al.* (2006). *Das Hochschulranking. Vorgehensweise und Indikatoren*, CHE-Working Paper No. 75, Gütersloh. <http://www.che.de/downloads/Methoden_2006.pdf>.

Cave, M., *et al.* (1997). *The Use of Performance Indicators in Higher Education. The Challenge of the Quality Movement*. London: Jessica Kingsley.

Federkeil, G. (2001). "Hochschulranking – Elemente eines methodischen Standards" [University Ranking – Elements of a Methodical Standard]. *Wissenschaftsmanagement – Zeitschrift für Innovation* 2, pp. 7-11.

Neave, G. (1988). "On the Cultivation of Quality, Efficiency and Enterprise: an Overview of Recent Trends in Higher Education in Western Europe, 1986-88". *European Journal of Education*, 23:7-23.