Ranking of universities and higher education institutions for student information purposes?
Ranking of universities and higher education institutions for student information purposes?
Published by the Swedish National Agency for Higher Education 2009
Högskoleverkets rapportserie 2009:27 R
Contents: Swedish National Agency for Higher Education, Department of Analysis, Ellen Almgren
Graphic design: Swedish National Agency for Higher Education, Department of Communications
Printed by: Swedish National Agency for Higher Education, Stockholm, November 2009
Printed on environmentally labelled paper
Contents

Summary 5
Introduction 7
The ranking phenomenon – an overview 11
Summary 11
Ranking – a brief history 11
Ranking – what is it? 13
The ranking debate 24
The effects of ranking 27
International trends in ranking 39
Survey of existing ranking systems – some conclusions 43
Summary 43
Existing rankings – a rich flora 43
International ranking systems 45
The United States – a wide range of commercial rankings 48
Canada – a long and stormy tradition of university ranking 63
Australia – from student questionnaires to ranking 69
Great Britain – newspaper rankings for student choices 81
Germany – multidimensional and interactive rankings 92
Sweden – limited range of university and higher education rankings 96
Rankings as student information 109
Summary 109
Information to students 109
Ranking as information for students? 111
Student indicators 116
Data sources 128
Student information in Sweden – concluding remarks 141
References 149
Summary

Ranking of universities and other higher education institutions has become increasing common in recent years. The purpose of rankings is frequently to make complex circumstances more transparent, especially for students. In this report, we describe the ranking phenomenon in general and, more specifically, a number of existing ranking systems. We also discuss the prerequisites for and the pros and cons of ranking in terms of information for students. This report is a presentation of an assignment by the Government to Högskoleverket [the Swedish National Agency for Higher Education] to survey and analyse the ranking of universities and other higher education institutions. The report comprises three main sections. In the first chapter, we describe the ranking phenomenon in general terms, in the second we survey a number of existing ranking systems, and in the third and final chapter we discuss ranking in terms of information to students.

The ranking phenomenon has expanded in line with the supply of higher education. Rankings are sometimes described as a way of simplifying and clarifying a complex scenario for students and other interested parties. Even if the ranking concept may cover many aspects, most rankings have a common factor in that they present indicators of quality – explicit or implicit – that are weighted to produce an outcome which, in its turn is ranked in comparison with all other such results. In other words, ranking is an attempt to measure the quality of higher education or research, but it differs from many other forms of quality assessment because it is relative – there are no absolute criteria or norms for what may be regarded as “minimum quality”. This is partly because the ranking designers are often commercial entities – newspapers or magazines – with no direct responsibilities for the higher education sector.

There is extensive discussion of the ranking concept. Ranking is frequently defended on the grounds that it is useful for students and for other (less well-informed) stakeholders. There is, however, widespread criticism of what factors are measured and the way measurements are carried out. It is hard to determine the extent of the impact of rankings on institutes of higher education and on students. The – relatively few – studies that have been conducted do not provide any unequivocal information. There are both positive and negative effects on higher education institutions. Some students use the rankings when making their choices, but most do not. The international agenda involves, for example questions regarding the quality of rankings, European challengers versus the international ranking system, and multidimensional, interactive rankings.

Universities and other higher education institutions are ranked all over the world. A few rankings are international, in the sense that they rank higher
education institutions in several countries. Such international rankings are perhaps subject to the toughest criticism, but they have considerable impact.

We have decided to study a number of national rankings that are designed to provide information for students in more detail. There is a wide range of ranking systems in the United States with a relatively long tradition, but considerable protests against ranking have also been voiced. This also applies to Canada. Australia has a long tradition of extensive student questionnaires, which have ultimately become both a source of ranking information and a basis for the allocation of resources. In Great Britain, higher education institutions are ranked by several major newspapers, with the object of informing students before they make their choices. Germany represents perhaps the most well-known example of multidimensional, interactive ranking. Sweden has a rather short and very limited history of ranking universities and other higher education institutions.

Students invest both time and money in their education. As a result, it is important for potential students to have access to comprehensive and relevant information about higher education before they choose. Based on our survey of rankings in general, and a couple of specific ranking systems, we conclude that most rankings do not provide all-round, relevant or reliable information about the quality of higher education.

Weighted rankings of entire higher education institutions are particularly problematical. Quite simply, they provide too little information. They also assume that all the potential target groups have a common and very precise definition of quality – a definition which may be seriously questioned in terms of indicators, methods, reliability and – in particular – relevance. Multidimensional rankings, or rather student information systems, in which students or other stakeholders are themselves permitted to define quality, are a more attractive alternative. Quality and the selection of indicators also present problems in this context too, however.

There is a great deal of information about higher education in Sweden, and much of this information is of satisfactory quality. But the quality of information to students does not merely depend on quality-assured data sources and the reliability of the information. It is also important that the information is used for the purpose for which it was intended. Very little quality-assured information is suitable for ranking of the kind that we have studied.

It is obviously desirable that the information is transparent and accessible for students who are interested in comparing different higher education institutions or different programmes, and there is certainly potential for further development in this area. But further investigation of a number of issues is required, however. The question of what the target group – the students themselves – want is also crucial. An awareness that simplification of information also involves a loss of information is also required, and that information simplified to the extent called for in weighted ranking lists is hardly information worth the name, in the context of the quality of higher education.
Introduction

In May of 2008, the Government assigned the Swedish National Agency for Higher Education (HSV) to survey and analyse ranking systems for universities and higher education institutions. Based on this survey and analysis, HSV is to assess the prerequisites and the advantages and disadvantages of a ranking system for Swedish universities, institutions of higher education and education programmes. In this assignment, HSV is to study

- various ranking entities, for example academic and political entities, central government authorities, and entities in the private and media sectors,
- various purposes of rankings, for example consumer information to students and other interested parties, and the need for brand development on the part of universities and higher education institutions,
- the various focuses of rankings, for example education or research, entire higher education institutions or defined education programmes, labour-market aspects and student-welfare issues,
- various methods, indicators and time perspectives for ranking, and the results of the choice of method, indicators and periods,
- the possible impact of rankings on students’ higher education choices,
- differences and similarities between quality evaluation and ranking,
- the advantages and disadvantages of ranking as an incentive for quality development on higher education, and
- international trends in the development of ranking systems.

This assignment is to focus on education at the first, second and third cycle levels, but research is also to be taken into account.

In accordance with this assignment, the Swedish National Agency for Education has consulted the Swedish Research Council (VR), the Swedish Council for Working Life and Social Research (Fas), the Swedish Council for Environment, Agricultural Sciences and Spatial Planning (Formas) and the Swedish Agency for Innovation Systems (Vinnova). There has also been consultation with the Association of Swedish Higher Education (SUHF), the Swedish National Union of Students (SFS) and the Swedish School-students' Council (Svea). This report is the Swedish National Agency for Higher Education’s presentation of this assignment.

---

1. The Swedish word “rankning” is a clumsy attempt to cover the English word “ranking”. The real meaning of ranking is “set in order of precedence”, but it does not exactly cover the ranking concept as applied in an international context, where ranking is often applied in connection with “league tables” – an expression which seems foreign from a Swedish viewpoint. Furthermore, not all activities described as ranking lead to league tables – and not always even to a strict order of precedence – as will be shown subsequently in this report.

2. The Government decision for the assignment is denoted U2008/3731/UH.
The background to this assignment is the increased interest attracted by various forms of ranking of university and higher education institutions in recent years, especially in an international context. However, in parallel with this growing interest in ranking, there has also been greater criticism of existing ranking systems. The background description of the assignment states, amongst other things, that there are currently no international ranking lists that focus on the quality of the education provided.

The reason for assigning HSV to survey and analyse various aspects of ranking is to provide increased information about the extent to which ranking may improve students’ access to relevant information about differences in quality between various higher education programmes and institutions in Sweden. Ranking systems have been justified in these terms in many countries: potential students need information about quality differences between higher education institutions, since higher education studies involve an appreciable investment for the individual, both in terms of time and money. Ranking may give students better information in their choice of institution and education programme. In addition, it is hoped that ranking encourages competition between higher education institutions which, in its turn, provides incentives for further quality development of their programmes. Companies and employers may also be regarded as a target group for ranking, if this helps them to recruit of students from institutions offering the highest standards of quality.

In performing its assignment, and based on the reasons stated by the Government for the assignment, the Swedish National Agency for Higher Education has primarily focused on questions concerning ranking in the context of the quality of higher education. The basic assumption has been that the target group for ranking in this assignment is primarily students. This does not mean, however, that we neglect indicators covering research activities, or that we totally ignore other potential target groups.

Quality is the key factor in this assignment, and in most of the subsequent discussion. The basic assumption is that, in principle, rankings always try to monitor some form of quality, even if this is an inadequate measure, incorrectly defined or simply implicitly stated. As a result, a substantial proportion of HSV’s tasks in this assignment have involved discussion of the way in which quality is measured and assessed, quality indicators and the problem of defining quality.

This report consists of three major sections. The first section describes the ranking phenomenon for universities and other higher education institutions from a general perspective – history and current trends. Various aspects of what ranking is and what it is not are discussed and, as a result, there is a discussion of the boundaries between ranking and other types of quality assessment. In addition, some aspects of the ranking controversy are presented, and of the criticism voiced in recent years. There is also discussion of the impact of ranking, both on student choices and as an incentive for quality development in higher education.
The second section of the report is concerned with existing ranking systems – both at the international and the domestic levels. Some of the most well-known ranking systems used to rank universities and higher education institutions all over the world are described in brief. A selection of other ranking systems – often of a national character – is also analysed. We describe in greater detail rankings conducted in the United States, Canada, Australia, Great Britain and Germany. It is not easy to provide a full overview of the rich flora of the current rankings of institutions of higher education, but we attempt to describe some of the most influential and/or interesting examples. Finally, there is an account of the Swedish ranking tradition, which is not as extensive and comprehensive as its international counterparts.

The third section of the report discusses ranking in terms of student information and the prerequisites for ranking in Sweden. In this context, there is a discussion of ranking as a source of information for students, in which we make use of our previous findings. What types of rankings are appropriate for informing students about quality differences in higher education in a comprehensive and relevant manner? In addition, we discuss the quality indicators that are commonly applied in the rankings studied, their strengths and weaknesses, together with brief comments on the possibilities of finding similar indicators in Sweden. We also discuss different types of data sources. The report concludes with a discussion of more general prerequisites for a system for the ranking Swedish universities and higher education institutions in the context of information for students, and the problems involved.

This report should be primarily regarded as a survey of ranking at both the international and national levels. It is based on academic literature in the field and on direct studies of a number of ranking systems. As is always the case in surveys of this nature, we have been forced to select what we have studied. We have primarily tried to concentrate on literature that is as current as possible, but we have also studied particularly interesting references in the course of our reading. The ranking systems which we have studied in particular detail are those that have a long history, have an international impact, or are especially relevant from a student perspective. Unfortunately, language aspects have affected the objects of study in this report. Apart from German and Swedish rankings, all the other rankings are from English-speaking countries. We hope, however, that readers who are interested can analyse this topic in greater depth as a result of our references.
The ranking phenomenon – an overview

Summary
The ranking phenomenon has expanded in line with the supply of higher education. Rankings are sometimes described as a way of simplifying and clarifying a complex scenario for students and other interested parties. Even if the ranking concept may cover many aspects, most rankings have a common factor in that they present indicators of quality – explicit or implicit – that are weighted to produce an outcome which, in its turn, is ranked in comparison with all other such results. In other words, ranking is an attempt to measure the quality of higher education or research, but it differs from many other forms of quality assessment because it is relative – there are no absolute criteria or norms for what may be regarded as “minimum quality”. This is partly because the ranking designers are often commercial entities – primarily newspapers and magazines – with no direct responsibility for the higher education sector.

There is extensive discussion of the ranking phenomena. Ranking is frequently defended on the grounds that it is useful for students and for other (less well-informed) stakeholders. There is, however, widespread criticism of what factors are measured, and the way measurements are carried out. It is hard to determine the extent of the impact of rankings on institutes of higher education and on students. The – relatively few – studies that have been conducted do not provide any unequivocal information. There are both positive and negative effects on higher education institutions. Some students use the rankings to help them make their choices, but most do not. The international agenda involves, for example, questions regarding the quality of rankings, European challengers versus the international ranking systems, and multidimensional, interactive rankings.

Ranking – a brief history
The starting point for the ranking of universities and higher education institutions is normally regarded as the early 1980s, when the U.S. News and World Report magazine published a ranking of American universities. The fact is, however, that the ranking of higher education institutions may be observed much earlier than this. Such institutions were already being classified in 1870 in the United States, and various rankings were subsequently performed sporadically throughout the 20th century. The first media ranking of universities and higher education institutions was published by the Chicago Tribune in 1957. However, the U.S. News and World Report’s rankings in 1983 and the
Fiske Guide to Colleges in 1982 marked the start of considerably more extensive ranking activities in the higher education sector.\textsuperscript{3}

There has been a very substantial expansion of the ranking market in recent decades, and not only in the university and higher education sector. However, only a very few rankings are international in the sense that they cover universities and higher education institutions in many parts of the world but, on the other hand, they are the most well-known. The vast majority of rankings are national – sometimes regional – ranging from higher education institutions in an entire country to the ranking of certain specific education programmes.

The expansion of rankings in higher education has occurred in parallel with a very considerable growth in the number of organisations in the higher education sector. It was estimated that there were slightly less than 17,000 such organisations in mid-2007. The flow of students – both in the western world and elsewhere – has also grown dramatically. As a result, it is hardly surprising that rankings have become increasingly popular, since they represent a way of organising and simplifying a complex reality by classifying higher education institutions in terms of one or more measurable criteria. In its turn, this makes it considerably simpler to compare the various institutions with one another, although the relevance of these comparisons obviously entirely depends on the indicators on which the ranking is based. One consequence of the growth of this sector is that competition on an increasingly global higher education market has become much keener. Higher education and research institutions are competing at the international and national levels for the “best students”, the “best researchers” and, in particular, for funding. Potential students, in their turn, want to know where they should preferably invest their time and money to get the “best” education.\textsuperscript{4}

At the same time, the governments of many countries have increasingly emphasised the importance of quality in higher education and research, and the links between research, innovation and economic growth. This has often resulted in greater requirements for opportunities to make demands on higher education and research institutions. And as a result, there is also increased demand for information about the quality offered by the various institutions concerned.\textsuperscript{5}

Accreditation of programmes, periodic assessments, evaluations carried out by external experts, inspections, audits, performance-based contracts, benchmarking and the assessment and evaluation of research are all different forms of such demands for responsibility. Some of these measures are initiated by the organisations (higher education institutions) themselves, others are carried out externally by those who provide funding, quality assurance authorities, government bodies, vice-chancellors’ councils and so on. A wider

\textsuperscript{3} Salmi & Saroyan (2007).


\textsuperscript{5} Hazelkorn (2007) and Salmi (2008).
group of stakeholders – for example the general public (via the media), student bodies and employer organisations – also conducts various assessments of higher education and research, Ranking is commonly said to be an instrument for the latter group.\textsuperscript{6}

**Ranking – what is it?**

What, then, is ranking? It is not easy to define ranking in a clear-cut way. The simplest answer – but perhaps not the most exhaustive – is that it is a collection of indicators.\textsuperscript{7} A fuller explanation is that:

League tables, also referred to as institutional rankings and report cards […], are constructed by using objective and/or subjective data obtained from institutions or from the public domain, resulting in a “quality measure” assigned to the unit of comparison relative to its competitors.

* Salmi & Saroyan (2007)

The above definition is still rather broad, however. Rankings share a number of common characteristics, but there is also a considerable range in the character of ranking systems. Some components are crucial, however, even if their application may vary from one ranking system to another:\textsuperscript{8}

- **Rankings involve a collection of indicators**
  
  Ranking is always based on one or more indicators, but there is considerable variation in the indicators employed, as described in the next chapter. Ranking often entails quantification of resources (personnel, physical or economic resources), “input indicators”, or results (of education programmes or research) – i.e. “output indicators” – but there are also several other types of indicators.

- **It is assumed that the indicators measure some form of quality**
  
  The priorities implied by a given ranking are not always explicitly stated, but the fact that the entity ranked first is regarded as the “best” implies that the ranking designers consider that the indicators represent something that is important or positive. Naturally, these important or positive factors vary, depending on the indicators selected, and on the interpretation of the ranking results. As a result, ranking is closely linked with discussion of, and definitions of, the quality of higher education, as clearly demonstrated subsequently in the section covering the public debate on ranking.

---

\textsuperscript{6} Salmi & Saroyan (2007).

\textsuperscript{7} “Rankings are simply collections of indicators”, according to Alex Usher in an address given at a conference on *Does Size Matter? Universities Competing in a Global Market*, Reykjavik, 6 June 2008.

\textsuperscript{8} The following section is based on several different sources, including Salmi & Saroyan (2007), Finnie & Usher (2005), Usher & Savino (2007) and Dill & Soo (2007).
• Indicators are based on some source of information, but this may involve very different types of data – ranging from subjective experiences to official statistics

The data used in the various ranking systems varies considerably. The use of several different types of sources for the various indicators in one and the same ranking is common practice, as further described in the survey of ranking systems in the next chapter.

• The indicators describe a specific unit, but the type of unit may vary – ranging from entire higher education institutions to departments, disciplines or education programmes

Some of the rankings that are perhaps the most familiar, cover entire higher education institutions – in other words, the indicators measure all the characteristics of an entire university or comparable institution. Several rankings divide higher education institutions into various categories, however, and rank each category separately. There are also rankings that simply concentrate on certain types of institutions (for example business schools). In addition, there are also rankings describe and rank programmes or academic disciplines at higher education institutions.

• Indicators are sometimes – but not always – weighed up to provide a cohesive, aggregated result

In many rankings, the various indicators are weighed up to achieve a cohesive, aggregated result. This may take more or less sophisticated forms, applying more or less sophisticated statistical techniques. This often means that various indicators are given different weightings, depending on which indicators are regarded as the most crucial. There are also multidimensional rankings, however, in which the indicator results are reported separately, with no attempt to weight them to produce an overall result.

• Finally, ranking involves a listing in order of precedence of the units covered, involving comparison of the results achieved

The ranking results resemble the results of a competition (league table) in which the various units compete for the best rating, as opposed to measurement and evaluation of results in terms of absolute criteria and quality standards.

Defining quality

The fact that ranking systems list units from best to worst indicates that what is measured is something desirable. The designers of ranking systems often state explicitly that they are measuring quality, although sometimes this is not clearly indicated. In any event, the choice of indicators and the aims are to be

---

regarded as some sort of definition of quality. As regards external stakeholders (users), the ranking system more or less deliberately determines the limits for the definition of quality.\footnote{Usher & Savino (2006).}

As a result, the indicators that are actually included in a ranking are a key issue in an evaluation of the relevance of such a ranking for a specific user. There are, of course, an almost infinite number of potential indicators for higher education or research characteristics, but only a relatively limited number of them are employed in current ranking systems. This will be further elucidated at a later stage in this report.

Distinctions are often made between different categories of higher education ranking indicators (and other quality assessment systems). The classifications may vary to some extent, but for the most part the following categories are involved:

- **Characteristics of new students (beginning characteristics)**\footnote{“... represent the characteristics and abilities of incoming students as they start their programs...”, Finnie & Usher (2005:19).}

  Characteristics describing students commencing higher education at a given unit (institution, subject, education programme, etc.). This may range from background variables such as gender, family and social or ethnic background to more intangible characteristics such as learning propensity, etc. In cases in which this type of indicator is applied, this normally takes the form of the qualifications of new students in terms of grades or test results (commonly expressed as “SAT” scores in the United States).

- **Resources for education or research (inputs)**\footnote{“... embody all the factors that reflect or determine students’ educational experiences and are at least theoretically controllable by the educational system (at the institutional or some other level)...”, Finnie & Usher (2005:20). Finnie & Usher do not discuss inputs in a broad sense, however, but confine themselves to what they call “learning inputs” – that is to say indicators that involve higher education rather than research in their definition, although “inputs” are also included in the ranking of research, or the commonly applied but hard-to-define concept of “world-class universities”. We have adopted the somewhat broader “resources” term to depict this category, since the English term inputs cannot be translated in a more appropriate form.}

  This involves the resources supplied by the institution or department, etc. in broad terms, and may involve personnel resources in terms of the number of teachers or, more specifically, the number of teachers with Ph D qualifications. Financial and physical resources are also taken into account, for example as regards disbursement per student or access to computers and library facilities.
• Results of the education programme or research (outputs)\textsuperscript{13}

Indicators of this nature indicate the direct outcome of higher education (or research), and may encompass measures such as student throughflow the number of degrees and student credits. They may also cover factors that are less amenable to measurement, for example students’ analytical abilities.

• Final outcomes\textsuperscript{14}

Indicators describing the long-term benefits of higher education (or research). The most common measures are students’ status in the labour market and future incomes, but it may also involve more indirect outcomes such as job satisfaction, good citizenship, general education, etc.

• Academic reputation\textsuperscript{15}

Most ranking systems also include an indicator that – consciously or unconsciously – denotes the reputation or prestige of the institution (or department) concerned. This often involves indicators based on student questionnaires or, more commonly, surveys of persons in the academic world, such as professors and vice-chancellors. Some rankings incorporate success in international rankings or successful attraction of research grants as an indicator of academic reputation.

• Processes

It has been widely suggested that there are certain indicators that fail to cover the above rather well-established categories. Some people consider that there is something that lies between resources and results – a process that describes the actual learning process, for example in the form of the quality of teaching, educational methods, the interaction between teachers and students, the actual content of the programme and opportunities for training placements.\textsuperscript{16}

• Value-added

Indicators (or methods) that capture the added-value of an education programme are applied to a limited extent, but are nonetheless frequently sought after. Instead of applying beginning characteristics in their own right, these characteristics are used as a form of control variable in a statistical model. If entrant characteristics are used as indicators of quality, they are interpreted in terms of the attractiveness of the institution (department or programme) –

\textsuperscript{13} “...represent any skills or characteristics of graduates which the educational system potentially produces…”, Finnie & Usher (2005:20). Once again, it is important to note that Finnie & Usher refer specifically to “learning outputs” in this context. We have chosen to refer to this category as results – a term which does not correspond exactly to outputs in English, although it describes satisfactorily the category for which these indicators apply.

\textsuperscript{14} “...represent the ‘ultimate ends’ to which the educational system may contribute…”, Finnie & Usher (2005:20).

\textsuperscript{15} Usher & Savino (2007) and Dill & Soo (2005).

\textsuperscript{16} Dill & Soo (2005).
in other words the competence of the students that the institution manages to recruit. But, from a value-added perspective, it is more relevant to look at the outcome related to the students’ initial characteristics. In other words, an attempt is made to measure how the educational programme has benefited the students, and not what the students have managed to achieve previously. Additional sophisticated measures in the same spirit are sometimes proposed, for example measurement of the outcome, taking into account both beginning characteristics and resources. Hence, this involves a measure of efficiency: roughly speaking the degree of added-value achieved per currency unit invested.

*Student welfare factors*

Factors that are not immediately related to the actual education programme or institution are a category that is applied less frequently. This may involve a broad range of student welfare aspects such as access to accommodation, health care services, cultural, sports and leisure activities, or the attractiveness of the location of the institution itself. This type of indicator should not necessarily be regarded as a separate category. Student welfare resources may be included on the same lines as physical resources at the institution concerned (access to libraries, IT facilities, etc.) are regarded as educational resources. On the other hand, many of these factors can only be affected by a higher education institution to a limited extent and, furthermore, they involve resources in the student’s overall situation during the education period rather than immediate resources in the education process.

This classification of indicators is not completely self-evident, and this is partly because different indicators may be regarded from different angles, depending on the potential target group and who has designed the ranking system. Sometimes, for example, indicators covering research activities (such as citation indexes or external research grants) may be regarded as a form of resources. Such resources may benefit students (and doctoral students in particular), but if the ranking is designed instead to measure the quality of research at the institution concerned, such indicators should obviously be regarded as reflecting results. Some people even treat research indicators as a separate category that is independent of both resources and results.

The most common indicators applied in the various rankings are measures of resources, results and reputation, but they are often supplemented by one or more indicators of entrance characteristics and final outcomes. Measures reflecting the actual learning process are seldom applied, and indicators of an added-value or student welfare nature are even more uncommon.

This heavy concentration on resource, result and reputation indicators may be interpreted as suggesting that there is not much dispute about the definition of academic quality. But if we examine exactly which indicators are applied in greater detail, and the way they are measured, the interpretation must be the opposite – that there are significant differences in definitions of quality in the rankings. Seen in this light, the quality of higher education appears to be a somewhat arbitrary concept.  

The differences in quality definitions may be explained to some extent by the fact that different rankings give priority to different aspects (for example education versus research), and that the target groups envisaged also differ (potential students and their parents, the general public, future employers, university administrators, etc.). Cultural and geographical differences probably also have an impact on rankings – that is to say what measures are considered to be important. But, to a far from negligible extent, the differences in the various ranking lists reflect what indicators are available or can be readily gathered without undue expenditure of resources, rather than the result of deliberate choices. Ultimately, most ranking designers are commercial entities (often media) that are more interested in selling their product (often newspapers and periodicals) than in becoming deeply involved in definitions of quality in higher education.

“*One size fits all*”?  
Many ranking systems rate entire higher education institutions, although it is becoming increasingly common for such listings to be supplemented by rankings for different disciplines and subjects, departments or education programmes. In some cases (for example U.S. News & World Report in the United States and Macleans in Canada), different rankings of different types of higher education institutions are carried out, based on their different functions – for example a ranking of traditional “full-service” universities, and another ranking for specialist institutions. On the other hand, several current ranking systems – in particular British and international rankings – apply the same technique for all higher education institutions, often backed by the argument that they have similar assignments and prerequisites. In other words, all institutions are evaluated on the same basis and are compared with each other, irrespective of their mutual differences. This procedure is commonly termed “one size fits all” ranking.

There are similar consequences for a policy of weighing up all indicators to produce a single aggregated result, as in the case of ranking all higher education institutions in terms of a single measure. The weighted result does not provide any indication of what the various indicators reveal for a specific institution. A “satisfactory” result (high ranking) may be achieved by being

---

relatively “good” for all the indicators covered, or by being unusually “good” in some respects, but considerably less satisfactory in others.

Some form of weighting and/or “normalisation” in connection with the averaging of indicators is often applied. Weighting occurs if it is considered that one or more of the indicators should have a higher rating than others (i.e. are more important from an overall quality aspect). Sometimes the ranking designers apply weightings on the basis of some theoretical argument, because a specific factor is particular important for quality. In other cases, an indicator is given a higher weighting because the data is very reliable or – on the contrary – it is given a lower weighting because the data is considered to be highly unreliable. But in many cases the weighting is relatively routine, and is simply based on previous practice.24

“Normalisation” is another type of weighting employed by ranking designers. This involves taking differences between higher education institutions into account, often in terms of their differences in size so that certain indicators are weighted on the basis of the number of students at the institution concerned. Sometimes an attempt is made to weight results in accordance with the focus applied by the institution concerned, or on the basis of the various disciplinary areas covered. The clearest example is to take into account the extent of an institution’s involvement in various disciplines in terms of bibliometric indicators, since the humanities and social sciences are inadequately reflected in international citation studies.25

Normalisation is, of course, intended to avoid some of the problems of the “one size fits all” concept, but it is not easy to weight out all types of differences between institutions or education programmes. As a result, the overall weighted rankings – particularly for entire higher education institutions – are relatively one-dimensional in their approach to quality. There is an underlying assumption that, in principle, the same quality definition fits all types of higher education institutions and programmes.

Data sources and ranking
Different ranking systems take their indicators from a number of different sources. These sources are sometimes classified in terms of three different types:26

• Questionnaire and interview surveys
Questionnaires and telephone interviews with various stakeholders – students, staff or potential employers for example – are a common source used to measure the quality of education.

26. The following section is based on Usher & Savino (2006)
• Independent third parties
This is the widest and probably the most common source category. This often involves information collected and published by government authorities responsible for the area concerned, and it is usually a question of administrative data in the form of official, frequently register-based statistics, although it may also involve financial information from research councils, for example, or accreditation or quality evaluation data from the public authorities responsible for the area concerned.

• Higher education institutions
Much of the information used in the various ranking systems is obtained directly from the institutions concerned. To some extent, of course, this is the richest source of information, but it is not subject to quality assessment in the same way as information supplied by third parties, and it is not always easy to compare.

This way of classifying different types of sources involves the data collection methods employed by the designers of ranking systems. Other classification concepts may be envisaged, of course – for example objective and subjective sources, or quantitative and qualitative information. But no classification system can be simultaneously comprehensive and, at the same time, consist of mutually exclusive categories. Information about what sources have been used in a ranking – irrespective of how they are classified – is crucial, however, in assessing the reliability of the ranking concerned.

Ranking compared with other forms of quality assessment
If designers of ranking lists formulate a definition of quality via their choice of indicators – implicitly or explicitly – and then try to measure and rank it, the question of what distinguishes ranking from other types of quality assessment, such as accreditation of degree entitlements or education evaluations, obviously comes to mind.

Due to a rather vague terminology for ranking phenomena, it is difficult to say with any degree of certainty whether there is any difference between different types of quality assessment. Most of the factors that are measured and compared may, of course, be described as ranking (and this applies in some cases).27

But if we consider the current ranking systems, most of them have a common denominator that distinguishes them from many other kinds of quality assessment. Ranking is conducted in relation to the units covered, and not in

27. The terminology in this area is definitely confusing. When we refer to ranking in this report (league tables), we refer to the international use of this term – operations that apply this label, usually commercial entities, especially newspapers and periodicals. Hence, we use the ranking concept in a broader sense of “order of precedence”, which is a technique or method of presentation employed in ranking (and in other activities). We use the verb “to rank” and “state in order of precedence” indiscriminately to denote the same phenomenon – the rating of units in accordance with a grading system.
comparison with any absolute and clearly defined criterion for what is to be regarded as “good” or “bad” quality.28

In a Canadian report in 200529 on quality measurement in higher education, the authors distinguish between four types of approach:

- The Minimum Standards Approach
- The “Ranking/Indicators” Approach
- The “Learning Impacts” Approach
- The “Continual Improvement” Approach

There are no clear-cut distinctions between the different approaches, but they differ as regards the target group for the evaluation, what is to be evaluated, the method to be employed and what are the predominant types of indicators.30

**The minimum standards method**

This type of quality evaluation is the most commonly applied in OECD countries. The Swedish National Agency for Higher Education’ quality evaluations, which we will be returning to in the final chapter of this report, may be said to belong to this category. This method provides a basis for periodic evaluations of higher education, and it is almost universally conducted at the subject level, not at the institution level.

The minimum standards method is often applied by governments to require accountability from higher education institutions. The popularity of this method is no doubt because it is considered to be a legitimate central government function, in contrast with ranking. This may be illustrated in the form of a policy analogy: “the regulation of food services is accepted as a legitimate government responsibility, but the preparation of ‘Best Restaurant’ guides is not.”31

The minimum standards method usually implies a combination of self-evaluation on the part of the institution concerned, followed by external assessment, often carried out by university experts. The methodology is primarily qualitative, and one advantage commonly cited is that it encompasses key quality aspects in higher education. As the name indicates, this method primarily involves checking that the institution or discipline concerned has acceptable quality and, although the reports may contain much useful information at the internal level they are not particularly transparent for external stakeholders.

---

30. The following account is based on Finnie & Usher (2005), unless otherwise stated. An interesting argument on different ways of requiring accountability in higher education is presented in Salmi (2008), however, which also lists some of the above methods, including ranking.
The ranking and key performance methods

The Canadian report distinguishes between two types of approach under this heading. The ranking method is described as the technique normally used by private-sector entities, often newspapers or periodicals, that focuses on resource indicators (and often on reputation). There is another method, however, that focuses on results and outcomes — Key Performance Indicators. The key performance method has been used, primarily in the United States and Canada, in the same way that the minimum standards method is applied in other OECD countries, and has been used to require accountability on the part of higher education institutions, but normally with a focus on the entire institution, rather than on particular disciplines or subject-oriented departments.

The key performance method is not synonymous with the ranking method, although they have much in common. Apart from the fact that they focus on somewhat different types of indicators, the key performance method is closer to the minimum standards method in that a key number often constitutes some sort of norm or benchmark.32

Ranking and key performance methods are almost exclusively quantitative. Quality is measured in terms of statistical indicators of one kind or another — frequently resource indicators in the ranking approach, and result indicators in the key performance method. The ranking and key performance methods are considerably more transparent and easy to understand than the minimum standards method, but have met with considerable criticism on the grounds that the indicators used are often meaningless and do not really reflect quality.

The learning impacts method

The learning impacts method is, in many ways, a reaction to the ranking and key performance methods, especially as regards these methods’ focus on “the wrong” indicators. Since higher education institutions (particularly in the United States and Canada) considered that they were involved in teaching activities, they considered that it was unfair to evaluate quality on the basis of indicators that were not clearly correlated with the learning process.

In view of this, several major questionnaire surveys targeted at students were developed, in an attempt to ascertain their experiences and assessments of the learning environment at the higher education institutions concerned. These evaluations were primarily designed to assist the internal quality process at these institutions. The techniques employed in the learning impacts method are exclusively quantitative, but the indicators differ from those employed in the ranking and key performance methods.33 Instead of focusing on measures of resources and results, the learning impacts method tries to measure the actual learning process and also the value-added aspects of education. This


33. In several countries, for example Canada, Australia and Great Britain, ranking systems now often include indicators based specifically on such major student surveys.
means that this method is widely appreciated, since it tries to tackle factors in the actual education process which are difficult to measure. The disadvantage is that direct measurement is not possible – everything depends on students’ subjective experiences, and in this respect the results are not particularly transparent or easy to interpret.

The improvement method

The ranking, key performance and learning impact methods all emerged in North America in the 1990s. In Europe, the improvement method was developed as an alternative system, due to dissatisfaction with the minimum standards method, although this hardly represented a paradigm shift, since the improvement method is largely a variation of the minimum standards approach.

The improvement method is used by both governments and higher education institutions to require accountability for quality issues. This method is also normally based on self-evaluation, combined with external assessment by experts. The main difference is that quality is defined in terms of processes in which each organisation is itself responsible for goals, methods and the monitoring of quality development. In other words, this method evaluates the quality process in higher education institutions, rather than quality delivered in accordance with the minimum standards method.

The assumption is that the improvement method provides incentives for quality development within the institution concerned in a way that cannot be achieved with the minimum standards method. However, in many ways both methods apply the same qualitative methodology. The advantage in both cases is that they evaluate aspects that are relevant at the internal level, and the disadvantage is that they are not sufficiently transparent at the external level.

Although key-performance ranking is used for government control purposes in North America, this method is primarily confined to commercial entities, and this affects the type of quality assessment carried out. Newspapers and magazines that normally design and publish rankings of universities and other higher education institutions have no official responsibility for the quality of the education provided. As a result, the rankings do not apply any assumptions about causal factors; why the results look the way they do or what might be improved for the benefit of the quality of higher education. It is of course questionable whether the method employed in most of the existing rankings would support conclusions of this nature. The main purpose of ranking is (or perhaps should be) to promote transparency in higher education for external observers. From this point of view, quality development in higher education is, at best, a side effect of ranking.34

---

The ranking debate

“... valid questions are answered with invalid numbers and formulas” Gerhard Casper, President of Stanford University in a letter to U.S. News and World Report, 1996.

The quotation above is from a letter in which the former President of Stanford University compares attempts to arrive at a satisfactory ranking method with the search for the holy grail. And many people would still agree with him, more than 10 years later. The debate about rankings, their raison d’être, their methodological weaknesses, and so on is very extensive. Most of what is said and written about rankings in the debate in academic quarters is based on the assumption that rankings are an unwelcome nuisance, but that they are here to stay.

We will return to the ranking debate, and to arguments both for and against ranking, at several points in this report. For the time being, however, we would like to summarise at a simple level some of the most common arguments in favour of ranking, and some of the commonest criticisms.

Arguments in favour of ranking

Discussion of ranking may be said to largely consist of negative criticism and arguments against ranking and ranking systems. But it is also argued that ranking fulfils a function. Some of the most important arguments are summarised in the following:

- **Rankings simplify and clarify the student’s picture of higher education**
  Students invest considerable time and money (either in the form of fees or student loans) in higher education. As a result, the choice of the education option is important. Many people, however, consider that it is difficult to obtain satisfactory information about the actual differences in quality between different higher education programmes and institutions. Rankings provide relatively simple and direct information about the quality of higher education. Straightforward information of this kind is often considered to be particularly important for students with no experience of universities and the higher education sector.

- **Rankings also simplify and clarify matters for interested parties outside the higher education sector**
  There is also a demand and need for information about the quality of various education programmes and institutions outside the higher education sector. This may be a question of potential employers who need information about

which programmes or institutions provide education and training for people the employer wishes to recruit. And members of the general public may also be interested in information about the quality of the education services that they help to finance in the form of taxes.

• **Ranking gives higher institutions free publicity**

The considerable expansion of higher education makes it increasingly difficult to have a full picture on a worldwide basis of universities and institutions that provide higher education and conduct research. It is also difficult for the institutions concerned to inform the public about their operations. Rankings are one way of enabling higher education institutions and their operations to make themselves known – an efficient and cheap method, many people would claim. As a result, it is probably seldom easy to decide to withdraw from a ranking system by failing to supply the relevant information.

• **Rankings provide incentives for improving quality in higher education and research**

Ranking can create incentives for improving various aspects of the quality aspects measured, by making the institutions concerned aware of their position in the ranking lists, thus enabling them to improve their rating.

• **Rankings improve the quality of data collection in the higher education field**

If the designers of ranking lists indicate their methods clearly, ranking may result in an attempt by the institutions listed to improve the quality of the data on which the rankings are based. Information supplied directly by the institutions concerned is a crucial feature of many existing ranking systems, and the quality of data collection may be enhanced if the institutions cooperate.

• **Rankings provide a many-sided picture of higher education**

In many cases, rankings are influenced by the views of several different stakeholder and expert groups (both students and academic experts, for example), and are based on various statistical indicators, sometimes involving quality evaluations, etc. Hence, rankings collate information from many different sources and, as a result, provide a broader picture than specific indicators.

**Criticism of ranking**

Criticism of ranking is more common than its defence, however. Much of this criticism is at a fairly detailed level, and we will be looking at some of these

---

points in later chapters. But there is also some general criticism which may be summarised in the present context.37

- **Ranking provides a misleading picture of quality in higher education**

Perhaps the most common criticism is that rankings fail to detect quality, and also that the wrong indicators are selected for the ranking process. Ranking designers are also criticised because they only measure what is available, they are not interested in quality, and they fail to define it. It is even more common to find, however, that the specific definition of quality used in the various rankings is subject to criticism. This is often a question of an undue focus on reputation as an indicator, or on traditional resource or result measures.

- **Ranking provides a misleading and false picture of differences in quality**

While the previous item involved the wrong choice of indicators in the various rankings, in this case the criticism is directed at the wrong method. Some of this criticism is concerned with the problem of combining various indicators to provide a single result. This involves a simplification, per se, but a great deal of information about specific indicators is lost in the process. The weighting of the various indicators is also sometimes criticised on the grounds that an unduly heavy or light weighting is applied to one aspect or another or, more commonly, that the weighting is arbitrary or based on tradition — in other words that there is no underlying theoretical concept. This heading also includes criticism of the robustness of the statistical methods used, for example that they are sensitive to minor changes and that the differences between units (higher education institutions) are not always statistically significant.

- **Ranking provides a misleading picture of trends over time**

Since several ranking systems frequently change their methods or indicators from one year to the next, the trend in the ranking lists give a false picture of the real changes that are taking place. What may appear to be major changes for a higher education institution from one year to another (upwards or downwards in the list) may be entirely due to a change in the method employed, or because some particular indicator has been added or excluded.

- **Ranking lists encourage unreasonable behaviour on the part of students and institutions**

This argument is, of course, a consequence of the above criticisms. If ranking lists are unable to measure genuine quality, this may clearly have disastrous consequences if students base their choices on such rankings, or if higher education institutions endeavour to improve their position on these lists. In particular, this criticism is directed at “one size fits all” rankings which, since they

---

treat all institutions in the same way, do not allow for any variation in assignments, activities or results. This, in its turn, stifles profiling and variation in higher education and instead promotes a conformity that is seldom desirable.

The ranking debate will probably continue to become more intense. Perhaps it will also become more consistent than the current discussions. At the moment, the advocates of ranking and its critics are often talking about different things. The advocates refer to the benefits of ranking – which is discussed in the next section – while the critics primarily discuss the methods employed.

The effects of ranking

Arguments based on the assumption that ranking has certain effects are often voiced in the debate on the ranking of universities. One common view is that ranking affects the perceptions of both institutions and students in certain respects. Those in favour of the ranking of universities tend, for example, to claim that potential students use rankings to make a rational choice when selecting an institution or education programme, and that ranking encourages the institutions concerned to improve the quality of their programmes. On the other hand, the critics consider that the ranking of universities leads to irrational behaviour on the part of both institutions and students.

But what do we actually know about the impact of rankings on institutions and students? In order to study this question, we review some of the research in this area in the next two sections. First we discuss the existing information about the effects of ranking on institutions, and subsequently we discuss the research findings on the role played by rankings in potential students’ choice of education programme or institution.

The importance of ranking for higher education institutions

Although the impact of ranking on institutions is a widely debated topic, the empirical research in this area is relatively limited. A few investigations are frequently cited, but they are exclusively based on data from other countries, particularly the United States. We have not traced any scientific investigation concerning the way in which ranking affects Swedish higher education institutions. As regards the research that is available, it appears to be unanimously considered that ranking does affect the institutions in one way or another. On the other hand, there are differences of opinion about the extent of the impact of ranking, or what effects are actually involved.

Certain studies indicate that the ranking of universities has a considerable – and predominantly negative – impact on the behaviour of higher education institutions. A research overview by David Provan and Karen Abercromby suggests that several earlier studies show that American higher education institutions often employ more or less honourable methods to improve their ratings, in particular in the U.S. News and World Report’s ranking. This may
involve, for example, reducing the admissions ratio at the institution concerned, or manipulating information on the results of the SAT admissions test.\textsuperscript{38}

Marguerite Clarke shows, in another study, that ranking may have a negative impact on an institution’s recruitment from different social groups. This study is based on previous research, but also on media reporting from various parts of the world – mainly from the United States – and it indicates, among other things, that university rankings have a negative impact on the recruitment of students from minority and low-income groups. As a result of rankings, the institutions concerned had a vested interest in recruiting students who were regarded as an asset in terms of maintaining or improving the institution’s ranking – in other words students who were high-achievers. Conversely, such institutions were less interested in recruiting students from low-income groups or from ethnic minorities. According to Clarke, the results suggest that ranking systems should give credit to institutions for their ability to educate students to a greater extent than at present, rather than recruiting students who are already high-achievers.\textsuperscript{39}

Other research indicates that ranking has both positive and negative effects on the measures adopted by higher education institutions. In 2006, Ellen Hazelkorn carried out an extensive questionnaire survey covering more than 200 higher education institutions, primarily in Europe but also in other parts of the world. The results show that ranking is of considerable importance for both universities and other forms of higher education. Only 17 per cent of the institutions concerned stated that rankings were of no importance whatsoever. More than half the institutions surveyed indicated that university rankings had had a positive impact on their reputation and publicity, and that, as a result, this had also helped the institution to attract students, initiate academic cooperation and recruit prominent researchers and teachers. It also emerged that almost than half the institutions surveyed had used the ranking results in their own marketing.\textsuperscript{40}

The majority of the institutions surveyed had a formal process for evaluating ranking results, finding weaknesses in their own organisation and taking any measures necessary. Most institutions also monitored the positions of other universities and institutions in the ranking lists. About 40 per cent of the institutions surveyed stated that rankings were an important factor in determining which institutions they were prepared to cooperate with, and an even higher proportion thought that ranking played a part in other universities’ and higher education institutions’ choice of partners. Furthermore, a majority consider-

\textsuperscript{38} Provan & Abercromby (2000). Several examples of such actions have been revealed in the media. In 1994, the \textit{Wall Street Journal} discovered that several higher education institutions had manipulated their data in order to improve their ranking and, as recently as 2008, Baylor University in Texas achieved this objective by improving its SAT rating for new students. This university simply offered students a book voucher for USD 300 if they took the test. See Inside Higher Ed (15 October 2008).

\textsuperscript{39} Clarke (2007).

\textsuperscript{40} Hazelkorn (2007 & 2008).
red that rankings influenced other key groups, such as students, parents, the government, employers, present and future staff at the institution concerned and sources of research funding.

This survey also indicated that most higher education institutions considered that the rankings had several weaknesses, mentioning for example that they favour institutions that are already well-established, that they mainly focus on research rather than education, and that they contain many errors. For its own part, the specific institution concerned regarded the effects of rankings as either positive or negative, depending on its position in the ranking tables.

There is also research that indicates that university rankings have a very considerable impact on the prerequisites of the institutions, but that it is far from clear to what extent this affects their actions or choice of strategy. In a British study conducted in 2007, an Internet questionnaire was distributed to all higher education institutions in the UK, and more than 90 institutions responded – this represents a response rate of 68 per cent. Most institutions agreed with the statement that rankings reflect a strange view of academic quality, that they affect the reputation of academic institutions, and that higher education institutions have virtually no influence on the ranking methods employed. There was also a high degree of consensus regarding the statement that rankings affect the reputation of higher education institutions and their possibilities of recruiting students, forming alliances and establishing cooperation, recruiting academic staff and competing for financial resources. 41

On the other hand, many institutions denied that their behaviour or strategies had been affected by university rankings, at least in crucial areas. In the case of the institutions that reacted to rankings, this mainly involved analysing and adopting a critical attitude to the specific ranking results or the methods employed. Some institutions also said that they had modified their marketing to some extent as a result of ranking. On the other hand, only a few institutions stated that they had changed their strategies in core areas such as the recruitment of staff, the range of courses offered and course contents, or the research focus. In addition, only a minority of the institutions said that rankings had forced them to make changes which they would have preferred to avoid.

The importance of ranking for student choices

What impact does the ranking of universities have on the choice of education programmes by potential students? And what factors influence their choice of programme and higher education institution? In this section, we review both international and Swedish research in this field in order to obtain a picture of what we know about the impact of ranking on students choices. At the international level, there are both comparative studies of several countries, and also research for specific countries, in particular the United States, Great Britain

and Australia. In the case of Sweden, there are no studies of the way in which Swedish students are affected by rankings, so instead we look at studies of a more general nature involving the choice of education programme by potential students. We also present the results of a survey of the way in which international students studying in Sweden were affected by rankings.

When discussing the effects of ranking on potential students’ decisions concerning education programmes, it may be useful to distinguish between their selection of an education programme and their choice of the higher education institution at which they wish to study. As we shall see in the following, the choice of education programme tends to take priority over the choice of institution, even if these two decisions are clearly not completely independent of each other. The potential student’s decision process may be divided into three stages which, in practice are not always so clearly separated:

• In the first stage, the student decides whether or not he or she wishes to pursue higher education studies.
• In the second stage, information is collected and the student identifies a limited number of possible programmes, based on personal interest, suitability and what is regarded as feasible.
• In the third stage, the student takes a final decision about which institution to apply for.

In order to clarify the situation, we may note that university rankings may be a possible decision factor in the second or the third stages. On the other hand, it is less likely that rankings play any part in the decision to pursue higher education. In the following review of the research, we try, as far as possible, to make a distinction between the choice of education programme and the choice of institution when discussing the role played by ranking.

Students’ choices – the results of international studies

There are some studies that investigate students’ choice of education programme and compare the results in several different countries. The results of the role played by university rankings point in different directions, however, depending on the survey method employed and the student groups investigated. It appears, for example, that international students and students with parents with higher education make use of university rankings to a greater extent than other categories.

An Australian report discusses research results from countries such as India, Chile, Great Britain, Canada and the United States, and concludes that university ranking are of relatively limited importance for students’ choice of higher education institution. Instead, factors such as geographical location, the range of courses offered, study fees and social activities appear to be more important. In some case, the reputation of the higher education institution

affects the student’s choice to a considerable extent, but this is more likely to be based on information from their immediate environment rather than on rankings. The most important information comes from parents, friends, student counsellors and the institutions themselves.45

In another compilation of research from other countries, Margaret Clarke investigates, among other things, potential students’ choice of university or college. According to Clarke, previous research indicates that this choice is influenced by several different factors: opinions about academic quality and the institution’s or the programme’s reputation, entry requirements, geographical location, the existence of study fees, access to scholarships, the infrastructure, employment opportunities, social life, advice from the immediate surroundings and information provided by guidebooks and ranking lists.44

As regards ranking effects, Clarke indicates that the research results are ambiguous. On the one hand, there is research that suggests that ranking is not of any great importance for the majority of students. Studies in the United States, Great Britain and Germany, for example, indicate that students from high-income families are the principal users of ranking lists. On the other hand, there is research that shows that the ranking of a higher education institution in a given year does actually affect the number of applicants and the number of new students at such an institution in the following year. Patterns of this nature apply in the United States, but also for example in Germany. We will be returning to these results in the subsequent section on the United States.

In another international study, Ellen Hazelkorn interviewed a number of students in Germany, Australia and Japan. The results of this survey indicate that different student categories use rankings in different ways.

• National students at the undergraduate level normally study at the local university in their city or region, and only a limited proportion study elsewhere. Such students monitor regional ranking to some extent.
• International undergraduate students often travel abroad under the auspices of a programme, but they take reputation factors that are sometimes based on ranking lists into account to some extent.
• National students at the second or third cycle levels use rankings, but often in combination with other information about the various institutions concerned.
• International students at the second or third cycle levels make the greatest use of ranking lists – and not only the international lists but also national rankings, since they are often acquainted with a particular country, but are not familiar with its higher education institutions.45

44. Clarke (2007).
Education programme choices in the United States

In the case of the United States, there are several studies of the way in which university rankings affect potential students’ selection of education programmes. Once again, the results point in different directions. Some studies indicate that ranking is not particularly important in the choice of programme or higher education institution, while others suggest a significant impact.

Kathleen Galotti and Melissa Mark have conducted a study that indicates that ranking may not have any major impact. This survey involved interviews with more than 300 upper-secondary pupils on three occasions during their last year at school. The interviews indicate that their choice of education programme was largely influenced by people in their immediate environment, such as parents, friends and classmates. Pupils with well-educated parents were particularly likely to rely on parental advice. Written information in the form of university brochures and information from student counsellors did not carry the same weight, and was mainly used in the initial phase.46

A study conducted by Patricia McDonough investigated the importance of rankings in a more explicit manner. This indicates even more clearly that ranking was not a particularly important factor in students’ choice of higher education institution. This study was based on an extensive national questionnaire distributed to more than 220,000 new students at 432 colleges and universities. It posed, for example, the question of the role played by university rankings in the decision to study at a specific higher education institution. A full 60 per cent of the respondents stated that rankings had played no part in their decision, 30 per cent said that rankings had been of some importance, and only 11 per cent considered that they had been a very important factor. The students who had been most influenced by rankings were Americans of Asiatic origin and students with parents with a good education and high incomes. High-achieving students were also more likely to be influenced by rankings than students with weaker school grades.47

A survey carried out among new students in Los Angeles at the University of California in 2007 produced similar results. The two main reasons stated by more than half the respondents for the choice of higher education institution were that the university had a good reputation, with satisfactory subsequent opportunities for obtaining a job. On the other hand, slightly less than 18 per cent stated that the university’s ranking had been “highly important” in their decision.48

There are, however, other U.S. studies that indicate that rankings have a relatively significant impact on the student’s selection of education programme. In one study, Ronald Ehrenberg and Irving Ives looked at 30 private colleges and universities over an 11-year period to see how their positions

in the U.S. News and World Report rankings affected their ability to attract students. This study showed that a lower ranking in a given year had several different results in the following year. A higher proportion of applicants was accepted, fewer students took up their places, new students’ qualifications were inferior to those in the previous year, financial income declined and, as a result, the institution’s ranking deteriorated further in the following year. In the authors’ view, the considerable impact of ranking methods gives cause for concern, since the U.S. News and World Report rankings sometimes change in the course of time. As a result, a different position in the league table does not mean that there has actually been any change in the quality of education provided by the institution concerned.  

Michael Sauder, in another U.S. study, analyses how rankings have affected students’ choice of law programmes. This study is based on information about the status of various institutions in the U.S. News and World Report rankings and on statistics for applicants for the various law programmes. This study indicates that the ranking of higher education institutions has considerable impact on the number of applicants and their qualifications, and on the number of applicants who actually registered for the programme concerned. According to Sauder, in the long term this may result in a self-perpetuating process, in which rankings affect the student population, and that this, in its turn, affects future rankings.

It may also be worth noting the impact on the number of applicants if an institution boycotts rankings – as many higher education institutions have decided to do, for example in the United States. According to Jamil Salmi and Alenoush Saroyan, institutions that refused to submit data to the ranking organisers sometimes suffered negatively – although not always. Reed College is one institution that actually benefited from its refusal to participate in college ranking systems. Its ranking dropped after it refused to supply data to U.S. News and World Report but, paradoxically, this resulted in a substantial increase in the number of well-qualified applicants in subsequent years.

To summarise, we may note that there are differences in the results of the various studies, presumably due to the use of different data. University rankings prove to be least important in studies based on the subjective opinions of potential students as regards how they were affected by rankings in their choice of education programme or higher education institution. On the other hand, a greater impact may be noted in the case of studies based on objective data in the form of ranking statistics and details of the number of applicants and students registered. Since totally different types of data with specific deficiencies are involved, it is difficult to determine which conclusion is nearer.

49. Ehrenberg & Ives (1999). The results of a similar study are also presented in Monks & Ehrenberg (1999).
the truth. As a result, we must be prepared to note that the results point in different directions.

**Student choices in Great Britain**

In the case of Britain, there are a relatively limited number of empirical studies of the role played by university rankings in potential students’ choice of education programme and, here too, the results are somewhat ambiguous, although they do not differ as much as in the American surveys.

A survey carried out by the Institute for Employment Studies in 1998 indicates that ranking is not a particularly important factor in students’ selection of universities or colleges. This survey is based on a national questionnaire answered by approximately 20,000 students, a questionnaire involving 1,900 school pupils in their eleventh year of study and extended interviews with more than 200 students in their first year of university studies. The results indicate that most of the respondents had decided what they wanted to do at an early stage. Their plans were affected by the grades they anticipated, their school experiences, access to student advisory services, and the expectations of their school and home environment. In the first instance, students chose a given subject, and in a subsequent phase gave priority to the education programme, access to assistance and resources, the reputation of the institution concerned and future employment opportunities. Quality evaluations and ranking lists played a very limited part in their choice of programmes or higher education institutions.52

Another survey carried out by Charles Eccles indicates similar results. In this study, the focus was on the annual rankings published by British newspapers, based on statistics for universities and colleges. Eccles considers that potential students are more affected by the competition for places and the quality of the learning environment offered by a particular university than by the university’s rating in the various ranking lists in a given year.53

There also some surveys, however, that indicate that rankings are of some importance in student choices—the UNITE Student Experience surveys is one example. This questionnaire survey includes questions on the crucial factors in choosing a university, and has been conducted every year since 2001. Comparison over time shows that the proportion of students stating that university rankings were an important factor in their choice of university has increased from 19 percent in 2001 to 29 per cent in 2007. In the 2007 study, this proved to be the sixth most important factor, after factors related to the university’s reputation. High-achieving students from affluent families were the primary users of such rankings, however.54

---

Research on education programme choices in Australia

In the case of Australia, access for surveys of the role played by rankings in student choices is even more limited. There is, however a 1998 study that indicates that ranking was not particularly important for the selection of education programmes or institutions by potential students. This study investigated the factors that affect the choice of university among applicants for higher education, and it was based on a questionnaire distributed to 937 persons in the first instance, and a follow-up questionnaire involving 538 participants in connection with admissions. This questionnaire was also supplemented by repeated interviews with a more limited number of applicants.55

This study indicates that the applicants tended to have limited information about key issues, and that they had inadequate information about their favourite university and the programme in which they intended to participate. The overwhelming proportion of the information to which they had access came from student counsellors, admissions bodies and open-house visits to the various institutions concerned. This survey indicates that interest in a particular subject was clearly the major factor in the student’s choice of education programme. In the case of the choice of higher education institution, on the whole applicants attached considerable importance to the institution’s prestige or reputation and, in this context, admission credits were frequently applied as an indicator of quality. Other important factors included access to employment after completion of studies and opportunities to proceed to third cycle studies. Some institutional factors were also important, for example easy access to the university from the student’s home.

One third of the respondents stated that they used rankings – for example in the form of the Good Universities Guide. But such rankings appeared to play a relatively minor role in student choices – almost 75 per cent of those who had studied ranking lists said that they had only been influenced by them to a limited extent, or not at all. The research report indicates, however, that the indirect effect may be somewhat greater, since study and careers counsellors study the rankings and they, in their turn, convey such information to potential students.

What do Swedish students choose?

So far, there are no studies that investigate the way potential students in Sweden are affected by university rankings. On the other hand, there is a limited number of studies that indicate which factors tend to play a crucial role in Swedish students’ choice of education programmes and higher education institutions. The studies presented in the following show that different students are affected by different factors. In general terms, however, the selection of education programme tends, in most cases, to be based on a given subject area or profession. On the other hand, the choice of higher education institu-

tion is often governed by proximity to the student’s home or family, and the quality of the education programme or institution.

A questionnaire survey conducted by the Swedish National Agency for Higher Education in the autumn of 1996 covering 3,000 students in higher education investigated, among other things, why students opted for the programmes they had selected. The key reasons revealed by the questionnaire were that they were interested in the subjects they studied (almost 90 per cent), that the programme led to a given profession (almost 75 per cent), that this enhanced their possibilities of obtaining a job which they would be happy with (more than 80 per cent), and that this made them more attractive in the labour market (76 per cent). Almost two-thirds of those who justified their choice in terms of their interest in the subject concerned simultaneously stated that their studies would lead to a profession to which they wished to devote themselves. A considerably smaller number of students said that they were simply interested in the subject (20 per cent), or that the aim was simply to get a good job (8 per cent). In other words, it was normally a question of studies based on personal interest and acquiring a suitable future profession.\(^{56}\)

A follow-up study by the Agency for Higher Education investigated the route to higher education, in the form of in-depth interviews with 24 students. This study indicated that the choice of education was by no means always a rational decision process, but instead involved different forms of rationality, randomness and necessity. Sometimes the choice of programme took the form of a negative decision process in which impossible options were excluded until only one alternative remained. This might, for example, involve rejecting unemployment and opting for academic studies instead, or rejecting programmes for which the student had inadequate grades. In other cases, the commencement of higher education was regarded as a “non-choice”. Some students, for example, could justify their studies because they had always been interested in a particular subject, or because they had wound up in their course by chance. In some cases, higher education studies were perceived as a necessity – fate was the determining factor, the student had a vocation to pursue a particular programme, or it was generally assumed that the student should enter higher education. Most of the interviewees described their choice of education programme as a combination of rational choice, random chance and necessity.\(^{57}\)

Subsequent studies also indicate that interest in a particular subject area or profession has the greatest impact in the choice of education programme, as is confirmed, for example by Studentspegeln, a questionnaire survey directed at more than 11,000 Swedish students and conducted by the Agency for Higher Education. The majority of the respondents (95 per cent) said that they were studying because they were interested in the programme they had selected. A relatively high proportion (more than 80 per cent) said that they wanted to

---

improve their opportunities in the labour market. Roughly 20 per cent said that their studies were an alternative to unemployment and, in this case, a somewhat lower proportion said that they were studying because they were interested in the programme for which they had opted.\textsuperscript{58}

Similar results are revealed in a survey conducted by the National Swedish Federation of Teachers, in a questionnaire that covered more than 3,000 students in law, medicine and social work programmes. The responses indicate that an overwhelming majority had chosen their programme because they were interested in it. Only a limited proportion said that the choice was because it was a good programme or had high status. When asked what factors in their immediate surroundings had influenced their choice, the majority said that they had been influenced by a representative of the profession concerned, or by their parents. The media, on the other hand, were said to have exerted very little influence. A relatively high proportion also responded with an “other” entry, which indicates that they often considered their choice to be independent, or that they had always wanted to pursue the programme concerned. As regards the choice of institution, the most important factors proved to be proximity to their home and family, followed by the quality of the programme.\textsuperscript{59}

An Agency for Higher Education report presents a similar picture. Statistics from the autumn term of 2006 indicate that more than half the Swedish student population were studying in their local municipality or county, and another 20 per cent were studying in a neighbouring county. Only 26 per cent were studying in a more distant county. However, these proportions differ from one programme to another, since certain programmes are only conducted at a limited number of locations (medical and nursing education, for example).\textsuperscript{60}

As regards factors that affect the choice of education programme, reference should also be made to Statistics Sweden’s surveys of upper-secondary pupils’ study interests. These surveys clearly show that the choice of programme takes precedence over the choice of institution. During the period 1998/99 to 2005/06, 72 – 79 per cent of the respondents stated that having the right education was the most important factor, while 12 – 16 per cent said the getting into the right higher education institution was the most important.\textsuperscript{61}

International students – why they choose a Swedish education?

International students are, in fact, somewhat more aware of the importance of ranking for potential student choices in Sweden than Swedish students. A survey has been conducted in this area that indicates that international students who come to Sweden study rankings to some extent.

\textsuperscript{58} Swedish National Agency for Higher Education (2007:20 R).
\textsuperscript{59} National Swedish Federation of Teachers (2007).
\textsuperscript{60} Swedish National Agency for Higher Education (2008:33 R).
\textsuperscript{61} Statistics Sweden (2006).
United Minds undertook an assignment commissioned by the Swedish Institute and the Agency for Higher Education in the form of an Internet questionnaire covering international students in Sweden in 2007, and this was subsequently followed up by the Swedish Institute. This questionnaire was exclusively targeted at “freemovers” (students who arranged their studies independently) from countries outside the Nordic region who were studying in an English-language Master’s programme. The questionnaire was distributed to e-mail addresses provided by institutions offering this kind of Master’s programme in English. In 2007, the questionnaire was distributed to 2,600 students, and received responses from 757 persons who complied with the criteria. The 2008 questionnaire was distributed to 5,000 students, and received responses from 881 persons who complied with the criteria. Most of the respondents were men, and the major proportion came from Asia. Approximately half the respondents were studying in technology programmes, followed by economics and other social sciences.62

Answers to the questionnaires indicated that many respondents had very limited knowledge about Sweden and Swedish education before they started to look for information about studies in other countries. In the 2007 questionnaire, almost 60 per cent said that they had limited or no knowledge about Sweden before they chose their education programme. In the 2008 questionnaire, 30 per cent said that they knew almost nothing about Sweden or education in Sweden. Europeans, Asians and Africans tended to be better informed, while Latin Americans were the least well-informed. When asked which sources of information were the most important for the choice of institution, the majority replied that the major source was information on the Internet, for example in the form of the website for the institution concerned, the Swedish Institute’s website, or chat sites and Internet forums. Information from friends and acquaintances was also a common source.

Responses to these questionnaires indicate that perhaps the major reason for international students to choose Sweden for their studies was that they found a programme or course that suited their needs. Other key reasons included the existence of English-language programmes, that Sweden was regarded as a secure and safe country, that students were interested in the Swedish life style and that there was an attractive higher education institution. The absence of fees was also a relatively important factor, as revealed by a question as to whether the respondents would recommend other students to study in Sweden. Almost 90 cent replied positively but, when asked whether they would recommend others to study in Sweden if fees had to be paid, the corresponding proportion was less than 40 per cent. Studies with no fees attached were a particularly important factor for students from Asian and Latin American countries.

The key factors for the choice of education programme and institution were the quality of the education provided and its contents and focus (98 per cent stated that these factors were rather important or very important). The special competence and reputation of the institution concerned were of almost equal importance, together with the possibilities of obtaining employment on completion of studies, and programmes in the English language. Student welfare aspects were also relatively important. Between 60 and 70 per cent of the respondents said that positive student and leisure activities were rather important or very important, together with satisfactory scholarship opportunities and assistance in finding accommodation.

As regards the importance of rankings, roughly one third of the respondents in both 2007 and 2008 said that they had used ranking lists as a source of information for the selection of courses or programmes. This applied in particular to students from Asia. In the 2008 questionnaire, almost 8 per cent said that the university’s ranking position was a rather important or very important factor in their choice of education programme and institution. In 2007, almost 60 per cent stated that the institution’s ranking was the crucial factor in their choice of country of study and education programme.63 Ranking lists were particularly important for Asian and Latin American students, but less important for students from North America. Utilisation of ranking lists was also more common in the case of male students and students of medicine, economics and the social sciences.

**International trends in ranking**

The ranking phenomenon for universities and other higher education institutions has expanded rapidly since the early years of the present century, and there are currently a large number of rankings with various focuses in the higher education sphere. There has been keen criticism of most of these ranking systems but, in the current debate, there is hardly anyone who would advocate their elimination. The most common view of the ranking of universities and other higher education institutions is, instead, that “they are here to stay, and what can we do about it?” And this approach also characterises most of the trends in the development of ranking systems.

**Greater interest in quality aspects**

On the one hand, there is a trend that is in opposition to the fundamental principles on which rankings are based – the presidents of several U.S. and Canadian higher education institutions no longer supply information to influential

---

63. Hence, a higher proportion state that they were influenced by rankings than those who claim that they used rankings as a source of information. This may be because the questions have been misunderstood. But this is not necessarily the case, since it is possible to obtain information about ranking results from indirect sources. The questions are also posed in different ways, and this may explain the different responses.
ranking designers (U.S. News & World Report and Macleans, in Canada). This means that such institutions will no longer be represented in the most prominent ranking lists – and this clearly has an impact on their visibility and publicity, and may affect future student recruitment. In this case, however, irritation about the way rankings are carried out appears to have been the overriding factor. Only time will show if this means that rankings will improve.64

On the other hand, there is widespread interest in finding out more about ranking. There is considerable interest in research on, and the evaluation of, different ranking systems (this report is only one of several similar attempts). This interest exists on the part of higher education institutions, governments, public authorities, national and international organisations, researchers, students and so on. Quality issues in ranking are a hot topic. Many conferences are held in this area, and there is a growing volume of literature in this field.

One indication of the topicality of the ranking quality issue is the Unesco-Cepes observer function that has been recently established to monitor whether rankings comply with the principles laid down in Berlin in 2006 as a result of the formation of the International Ranking Group in 2004. The 16 “Berlin Principles” cover the purposes, objectives and target groups for ranking, the design and weighting of indicators, the collection and processing of data, and the presentation of ranking results. It is hoped that the formulation of these Principles will result in greater awareness of methodological ranking issues, and hence to improvement and refinement of the ranking methods applied. The formation of an observer group to monitor rankings has revived the focus on the Berlin Principles, although there are no sanctions that can be invoked against ranking designers who fail to comply with these Principles.65

The Institute for Higher Education Policy in Washington has recently established a “policy centre” to monitor ranking trends and specifically to gather information, in the form of a “Ranking Clearing House”. This is yet another sign of increased interest in ranking and research in this area.66

Much of the criticism of ranking lists involves the biased selection of indicators. The question is raised of whether simple measures for resources and results are actually capable of covering the quality aspects of higher education to any great extent. Projects are also under way which attempt to describe the outcome of higher education in a more sophisticated form than ranking. The OECD Assessment of Higher Education Learning Outcomes (AHELO) is a project of this nature, in which the possibilities of assessing the quality of higher education in the form of knowledge tests similar to those employed in


66. Alisa F. Cunningham, address to the IMHE General Conference on 8–10 September 2008. See also www.ihep.org/Research/rankingsystems/clearinghouse.cfm.
undergraduate education (e.g. PISA tests) are being investigated. This project is still in its early stages, but a pilot study may be initiated in certain areas and countries in 2009.67

A European ranking market?
A specific European trend in the ranking field may be anticipated. Valerie Pécresse, the French Minister of Education, has notified that she intends to raise the question of a European ranking or classification system at a meeting of the ministers of education of the EU countries in November of 2008. One reason is that it is assumed that the Academic Ranking of World-class Universities conducted by the Jiao-Tong University in Shanghai, which is possibly the most well-known ranking system, gives a less favourable rating to European Universities (particularly non-English speaking institutions).68

The German Centrum für Hochschuleentwicklung (CHE) has declared that it is interested in extending its system for ranking of Master’s programmes to the whole of Europe as a result of the Bologna process and the possible subsequent increase in mobility in Europe in the higher education field. In point of fact, the CHE has already extended its German rankings by including Austrian, Swiss, Belgian and Dutch institutions. The CHE has also implemented a ranking of “excellent” Master’s programmes in Europe in the natural science and mathematics areas.69 In other words, there are plans to extend this operation into other academic fields.70

Multidimensional and interactive rankings
The German CHE organisation applies a slightly different ranking approach, as described in more detail in the next chapter. This ranking system is not based on a joint weighting of indicators, but takes one indicator at a time. The CHE describes this approach as “multidimensional” ranking. The ranking results are published in Die Zeit once a year – for a number of different indicators for several different education programmes in Germany. It is, however, possible – all the year round – to refer to a website and personally select the indicators and the programmes for which information is required.71

This type of multidimensional and interactive ranking (if it is to be termed ranking) is increasing all along the line. Several British, American and Canadian newspapers are currently offering possibilities of studying other indicators, subject-classified lists, etc. via websites, in addition to their traditional ranking lists. The results often contain lists with the same methodological pro-

---

67. www.oecd.org/document/22/0,3343,en_2649_35961291_40624662_1_1_1_11,00.html
70. Berghoff, Brandenburg and Möller-Böling (2008).
blems as the main rankings, but in some instances, as in the German case, the user can personally select the indicators and weightings applied.

The purpose of multidimensional rankings is often explicitly claimed to be to give potential students information about higher education which they themselves consider to be relevant. Such rankings (to the extent that they can be regarded as rankings) are broken down in terms of subjects or education programmes, and this further enhances their relevance from a student perspective. The breaking-down of information, and the categorisation of subjects or programmes presents problems, however, and the results are not always clear and unambiguous. This type of information is even more difficult to handle in international comparisons, since each country has its own classification system, which is not always easy to compare across national frontiers.
Survey of existing ranking systems – some conclusions

Summary
Rankings of universities and higher education institutions are conducted all over the world. A few are international, in the sense that they rank higher education institutions in several different countries, and such international rankings are perhaps subject to the keenest criticism. On the other hand, they have considerable impact.

We have chosen to study in more detail a number of national rankings that are designed to provide information for students. In the United States, there is a wide range of rankings with a relatively long history, but there are also extensive protests against rankings – and this also applies to Canada. Australia has a long tradition of major student surveys that has gradually come to form the basis for both rankings and the allocation of resources. In Great Britain, higher education institutions are ranked by several national newspapers to provide information for students before they make their choices, and in Germany we find perhaps the most well-known example of multidimensional, interactive ranking. Sweden has a relatively brief and limited history of the ranking of universities and other higher education institutions.

Existing rankings – a rich flora
As has already been noted in the previous chapter, there is currently a rich profusion of rankings of universities and other higher education institutions. This has resulted in a growing volume of research publications in this area, and the Swedish National Agency for Higher Education’s assignment is not the first survey of this nature. Various typologies and analyses of rankings are presented in a number of articles dealing with ranking, Salmi and Saroyan (2007), for example, present a classification of ranking systems based on the type of entity responsible for ranking design, and the purpose of the ranking concerned. It is clear that, in the western world, commercial ranking sponsored by newspapers and periodicals are the predominant category. The aim is often to give potential students, parents and interested members of the general public information about the quality of higher education. In other parts of the world, on the other hand, rankings tend to have different goals – accreditation and accountability requirements, for example. The initiators of such ranking systems are sometimes governments, government authorities and bodies within the higher education sector itself.72

Dill and Soo (2005) present the various characteristics of a number of different ranking systems – and this has been followed up by Usher and Savino (2006), who cover a wider range of rankings. In these surveys, the focus is on the indicators, methods (e.g. weighting) and sources of data employed, rather than on the type of designer of the ranking system. In both cases, the authors have tried to determine whether the various rankings employ consistent definitions of quality but, ironically, they come to quite different conclusions. Dill and Soo perceive a clear convergence in what the various rankings try to measure, while Usher and Savino, who cover a wider spectrum of rankings, find that the definition of quality is arbitrary in this field.

In this report, we make no attempt to pursue the analysis to this extent, but confine ourselves to studying a number of different ranking systems in order to demonstrate some of the variations in the type of indicator used, how the indicators are weighed up, the data sources employed, and the strengths and weakness of the various systems. We have decided to concentrate on relatively well-known ranking systems in countries that are not too dissimilar from Sweden. And, above all, we focus on systems in the “Anglo-Saxon” countries, due unfortunately to our limited language abilities. We present three different ranking systems in the United States, in particular because ranking has long traditions in this country. Canada has an education system that is not entirely dissimilar to the Swedish equivalent, and ranking has had a rather stormy history there. Australia also has rather well-developed ranking activities, based in particular on student questionnaire surveys. Great Britain is perhaps one of the countries in which ranking is most widespread – almost every newspaper publishes its own lists. But there are also other systems for the provision of student information in Britain that are of interest, and they are also mentioned in the background information for this assignment. In addition, there is a rather extensive academic debate on ranking in Great Britain. In Germany, there is a well-developed interactive system for student information (CHE Hochschuleranking) which is also mentioned in the background to this assignment. Finally, the rankings carried out in Sweden are clearly of interest in this context, and they are described in the final section of this chapter.

Before embarking on these more detailed descriptions, we briefly present some international ranking systems, since they are perhaps the most well known on a worldwide basis. Usher and Savino’s survey of ranking systems, which is summarised in Figure 1, provides a background for our limited selection of ranking presentations.

73. Ranking exists, of course, in many more countries than those discussed here. Reference is made to Pokhollow, Chuchalin, Agranovich & Mogilnitsky (2007), Devinsky (2008) and Kalanova (2008) for accounts of rankings in Russia, Slovakia and Kazakstan, respectively.
Figure 1. Characteristics (indicators and data sources) of 17 different ranking systems.

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of indicators</th>
<th>Number of indicators</th>
<th>Entrant characteristics %</th>
<th>Personnel resources %</th>
<th>Physical resources %</th>
<th>Results %</th>
<th>Final outcomes %</th>
<th>Research %</th>
<th>Reputation %</th>
<th>From surveys, No. of indicators</th>
<th>From third parties, No. of indicators</th>
<th>From the institution concerned, No. of indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asiaweek, India</td>
<td>18</td>
<td>25.0</td>
<td>28.3</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>16.7</td>
<td>20.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>18</td>
</tr>
<tr>
<td>Daily Telegraph, GB</td>
<td>4</td>
<td>0.0</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Education 18.com, Hong Kong</td>
<td>9</td>
<td>20.0</td>
<td>15.0</td>
<td>5.0</td>
<td>0.0</td>
<td>0.0</td>
<td>20.0</td>
<td>40.0</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Excecelencia, Spain</td>
<td>71</td>
<td>0.0</td>
<td>25.0</td>
<td>25.0</td>
<td>0.0</td>
<td>25.0</td>
<td>0.0</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Financial Times, GB</td>
<td>17</td>
<td>9.0</td>
<td>90.0</td>
<td>10.0</td>
<td>0.0</td>
<td>27.0</td>
<td>20.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>17</td>
</tr>
<tr>
<td>The Guardian, UK</td>
<td>7</td>
<td>28.0</td>
<td>35.0</td>
<td>10.0</td>
<td>10.0</td>
<td>17.0</td>
<td>0.0</td>
<td>0.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>La Repubblica, Italy</td>
<td>23</td>
<td>10.0</td>
<td>44.4</td>
<td>15.6</td>
<td>10.0</td>
<td>0.0</td>
<td>20.0</td>
<td>0.0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Macleans, Canada</td>
<td>24</td>
<td>10.7</td>
<td>20.0</td>
<td>48.3</td>
<td>5.0</td>
<td>0.0</td>
<td>16.0</td>
<td>5.0</td>
<td>1</td>
<td>5</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Melbourne Institute, Australia</td>
<td>26</td>
<td>11.0</td>
<td>3.5</td>
<td>11.0</td>
<td>12.6</td>
<td>4.8</td>
<td>40.0</td>
<td>17.1</td>
<td>3</td>
<td>23</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Netbig, China</td>
<td>18</td>
<td>12.0</td>
<td>21.8</td>
<td>6.0</td>
<td>0.0</td>
<td>0.0</td>
<td>45.2</td>
<td>15.0</td>
<td>1</td>
<td>10</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>Perspektwy, Poland</td>
<td>18</td>
<td>8.0</td>
<td>20.5</td>
<td>11.5</td>
<td>0.0</td>
<td>0.0</td>
<td>50.0</td>
<td>1.0</td>
<td>2</td>
<td>15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shanghai Jiao Tong-university, Kina</td>
<td>6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>10.0</td>
<td>0.0</td>
<td>90.0</td>
<td>0.0</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>The Times, GB</td>
<td>9</td>
<td>3.3</td>
<td>53.3</td>
<td>6.7</td>
<td>3.3</td>
<td>3.3</td>
<td>30.0</td>
<td>0.0</td>
<td>-</td>
<td>9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Times Higher Education, GB</td>
<td>5</td>
<td>5.0</td>
<td>25.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>50.0</td>
<td>0.0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>U.S. News and World Report, U.S.</td>
<td>15</td>
<td>15.0</td>
<td>20.0</td>
<td>15.0</td>
<td>25.0</td>
<td>0.0</td>
<td>25.0</td>
<td>1.0</td>
<td>3</td>
<td>11</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Washington Monthly, U.S.</td>
<td>8</td>
<td>33.3</td>
<td>16.7</td>
<td>11.1</td>
<td>22.2</td>
<td>0.0</td>
<td>16.7</td>
<td>0.0</td>
<td>-</td>
<td>1</td>
<td>7</td>
<td>-</td>
</tr>
<tr>
<td>Wuhan University Centre for Science Evaluation, China</td>
<td>45</td>
<td>10.6</td>
<td>8.5</td>
<td>16.6</td>
<td>3.4</td>
<td>0.6</td>
<td>40.6</td>
<td>11.7</td>
<td>2</td>
<td>22</td>
<td>21</td>
<td>-</td>
</tr>
</tbody>
</table>

Comments: The above table is the result of the merging of two tables in Usher & Savino (2006). The grey blocks indicate rankings that are analysed later in this chapter. Our analyses do not have the same degree of precision as the above table, and our results also differ from the results in the table for other reasons – we do not comply exactly with the categories of indicators employed, and we analyse rankings of a later date than Usher and Savino (the above rankings were carried out at some time between 2001 and 2006).

**International ranking systems**

Hardly anyone who is interested in the higher education sector can have failed to be aware of the annual rankings published by the Jiao Tong-university in Shanghai or by Times Higher Education – two of the very limited number of international ranking systems in the world.

**Shanghai Jiao Tong’s Academic Ranking of World Universities**

In 2003, the Jiao Tong-university started to rank research universities around the world in accordance with its achievements in the Academic Ranking of World Universities (ARWU).74 The original aim was to study the gap between

---

Chinese universities and world-class universities. The high degree of concentration on research in the ARWU ranking (see Figure 1, above) is because it is not possible to rank the quality of higher education all over the world due to the lack of comparable data (which, in its turn, depends on the considerable differences between counties and between universities).

The ranking group at the Jiao Tong-university claims that it is easier to find internationally comparable data for the quality of research. The strong point in their ranking is that they only employ data that can be checked. No information is obtained directly from the university concerned or from subjective surveys. This ranking is based on a limited number of indicators of two types: bibliometric citation indicators and scientific prizes (the Nobel Prize and the Field medals awarded for mathematics).

The ARWU ranking has been criticised on two grounds, right from the start. On the one hand, there has been criticism of its strong focus on research – its failure to cover any aspects of quality in higher education. A more relevant criticism concerns the selection of indicators to depict research quality. In the case of scientific prizes, this type of indicator has been criticised because the variation is extremely low (quite simply there are very few Nobel and Field awards). Another question that is often raised is what the number of Nobel Prize winners from a given institution actually says about quality. In addition, this is considered to be a far too superficial attribute, which can also be manipulated by employing Nobel Prize winners who do not actually participate in the institution’s activities.75

Considerable criticism has also been directed at the bibliometrically based citation indicators, in particular because this type of indicator is a poor way of depicting research quality, especially in the humanities and the social sciences, although medical and natural science research is published in a manner that complies rather well with the measures for bibliometric studies of this nature. Another common criticism is that bibliometric studies favour “Anglo-Saxon” universities, since they are accustomed to publishing their research in English-speaking journals.76

Times Higher Education QS World Universities Ranking

The Times Higher Education QS (previously the Times Educational Supplement) World Universities ranking is more of a hotchpotch than the AWRU ranking.77 The purpose of this ranking, ever since it started in 2004, is to offer an international comparison between various universities’ success in attempting to achieve world class, while simultaneously recognising that universities are multi-faceted organisations. Times Higher Education QS (THE-QS) identi-

77. See www.topuniversities.com/worlduniversityrankings/
fies four pillars on which it considers a world-class university is based: research quality, teaching quality, student employability and international outlook.

These four pillars result in six different indicators obtained from various sources. Research quality is measured by an academic “peer-review” in the form of a questionnaire survey in which academics state which institution, in their opinion, has the best research in their own particular discipline. The number of citations per academic staff member is also used as an indicator for research quality in this context. Teaching quality is measured in terms of the number of students per teacher, and student employability is also measured by means of a questionnaire survey (“employer review”) distributed to various potential employers. Finally, international outlook is measured by two indicators: the proportion of international teachers and researchers, and the proportion of international students.

The THE-QS World Universities ranking has also been criticised for failing to delineate education and teaching quality, and the indicator that is supposed to measure research quality (academic peer review) also has a very heavy weighting in this ranking (40 per cent).

But the major criticism of the THE-QS ranking is that some of the sources of data – both the questionnaire survey distributed to academics and the questionnaire for employers – are not transparent, and because they are purely a reflection of reputation and do not manage to chart quality. This is a typical problem for this type of questionnaire survey, as clearly illustrated by the example of an American questionnaire directed at students asking which were the best law schools. Princeton was ranked as one of the 10 best, although Princeton does not have a law school. This type of questionnaire appears to merely recycle reputations rather than rewarding genuine quality and, as a result, is only another form of popularity competition.78

The internationalisation dimension in THE-QS has also been criticised – partly because it is difficult to make comparisons between countries and partly because it tends to favour “Anglo-Saxon” universities. THE-QS has also been criticised for its citation indicator, on the same grounds as ARWU.79

The THE-QS World Universities ranking has modified its indicators and methodology over time, and this has affected the positions of higher education institutions in the ranking lists. Changes in ranking positions that might be interpreted as a dramatic shift in the performance of the institutions concerned may simply reflect a change in the way measurements are made.80

There are a few additional international rankings that cover universities in many parts of the world, but their focus is rather specific. The Leiden University’s various bibliometric rankings are exclusively based on scientific

citations and publications and, as a result, do not depict any aspects of the quality of education programmes. Webometrics ranks universities and higher education institutions on the basis of the extent to which they use the Internet, but with rather dubious results.\textsuperscript{81}

**Specific problems in international rankings**

The designers of international ranking systems face two problems of a general nature that have consequences for the quality of such systems. Firstly, it is hard to find comparable measures on a worldwide basis, because higher education takes such different forms in different countries. Secondly, the quality of the gathering of information about the higher education sector varies considerably. One result is that the international rankings focus more on the quality of research than on the quality of education, since it is simpler to find comparable data for research. The attempt to find comparable criteria also means that the international rankings often apply relatively few and relatively superficial indicators.\textsuperscript{82}

Variations in the classification of institutions, research and programmes in various countries also present problems for the designers on international rankings, and it is by no means always certain that the units compared by applying the same criteria are particularly similar. Even a simple matter such as the national names of institutions can be a problem. In Brussels, for example, there are two quite separate universities whose names, in both cases, would be best translated into English as the “Free University of Brussels”. One of the, Université Libre de Bruxelles, is a French-speaking university, while the other, Vrije Universiteit Bussel, is Flemish-speaking. The extent to which the international ranking designers are aware of the functions performed by the university concerned is far from clear. This also applies to university hospitals, which are sometimes ranked as separate units and sometimes as part of a larger institution. Another example of the difficulties in interpreting national languages is the German CHE report on “excellent universities” at the second cycle level, in which the Swedish Royal Institute of Technology (KTH) is consistently referred to as “KTH – Sweden’s Largest Technical University”.\textsuperscript{83} Some mistakes (perhaps many) in data management and processing are probably the result of lack of knowledge about systems at the national level.

**The United States – a wide range of commercial rankings**

Newspapers and magazines are the primary vehicle for university rankings in the United States. The first ranking initiated by the media occurred in 1957

\textsuperscript{81} See www.cwts.nl/cwts/LeidenRankingWebSite.html and www.webometrics.info for further information.

\textsuperscript{82} Usher & Savino (2006), Kälvermark (2007), HEFCE Issues paper 2008/14

\textsuperscript{83} Berghoff, Brandenburg, Carr, Hachmeister & Müller-Böling (2007).
in the Chicago Tribune. Initially, rankings focused on research, but were extended to the education field in the early 1980s, for example in the form of the Fiske Guide to Colleges and U.S. News and World Report. University rankings did not become really widespread until after the millennium shift, however.

Currently, several entities rank the more than 4,000 education arrangers in the American higher education sector on a regular basis. The most well-known rankings are conducted by newspapers and periodicals such as the Atlantic Monthly, Business Week, the Financial Times, Money, U.S. News and World Report and the Wall Street Journal, and they often market their rankings as relevant information for potential students who plan to study at a university or college.

This section includes a survey of three well-known ranking entities in the United States: U.S. News and World Report, the Washington Monthly and Princeton Review. They have been selected because they carry out national rankings that cover virtually all education areas, although with different focuses and methods. The U.S. News and World Report magazine's rankings employ rather traditional methods, and are primarily designed to give potential students information about the "quality" of various higher education institutions. The Washington Monthly, for its part, focuses on the "benefits to society" of universities, in a broader sense. In contrast, Princeton Review differs from the other two in that it is an education company, and its rankings are mainly based on a national student questionnaire.

U.S. News and World Report

Background and aims

The U.S. News and World Report magazine is published once a week, and has a broad news focus on political, economic, health and education issues. This magazine has been conducting rankings of American universities and colleges since 1983. The “America’s Best Colleges” special issue was first published in 1987, and is now on an annual basis. A follow-up publication entitled “America’s Best Graduate Schools”, which focuses on second cycle education, appeared in 1994. This magazine has also been published on the Internet since 1993, permitting for example searches for information about university rankings and the extraction of certain ranking lists.

According to U.S. News and World Report, the purpose of university ranking is to give potential students the information they require to choose the university that suits their needs. According to U.S. News and World Report, this ranking provides an objective guide, enabling students and their parents to compare the academic quality of different higher education institutions. U.S. News and World Report considers that it is important to help them to

84. Unless otherwise stated, this description and following sections are based on the U.S. News and World Report website at www.usnews.com
make a well-informed choice, since a university education is both a crucial and expensive investment for potential students and their families.

U.S. News and World Report’s rankings have a wide distribution and are considered to be the most well-known in the United States. According to a contributor to the Chronicle of Higher Education, U.S. News and World Report had a circulation of about two million in 2006, and the ranking issue sold even more copies. In addition, several hundred thousand copies of the special issue were sold, and approximately 50,000 users each paid USD 15 for access to information on the www.usnews.com website.85

**Ranking characteristics – focus, indicators and weighting**

U.S. News and World Report ranks more than 1,400 universities and colleges at the first cycle level, and divides them into different categories based on their focus and, in some cases, the region in which they are located. The following classification system was applied for the 2008 rankings.

- **National universities (262 units):** institutions with a broad range of programmes at the first, second and third cycle levels.
- **Universities with Master’s programmes (572 units):** institutions that resemble national universities, but with few or no third cycle research programmes. They are ranked in four regions: the Middle West, Northern, Southern and Western United States.
- **“Liberal arts” colleges (265 units):** institutions that mainly offer first cycle undergraduate education in which more than half the programmes are in the humanities, mathematics and the natural sciences.
- **Colleges at the first cycle level (319 units):** institutions with a focus on undergraduate education in which less than half the programmes are in the “liberal arts” area. They are ranked in four regions: the Middle West and Northern, Southern and Western United States.

In addition, economics and engineering programmes at the first cycle level are listed separately. There is also a separate ranking for the traditional “black” colleges and universities, that is to say institutions founded prior to 1964 whose primary purpose was, and is, to provide higher education targeted at Afro-Americans. It should also be mentioned that a large number of institutions are not ranked at all, primarily institutions with a special focus on the fine arts and music, for example. Other institutions that are not included do not employ the major national SAT and ACT admissions tests, since they have too few students or predominantly “non-traditional” students.

In addition to rankings at the first cycle level, U.S. News and World Report conducts a separate ranking for more than 12,000 programmes at the third cycle level. In this case, a distinction is made between the various subject areas, such as economics, law, medicine, natural sciences, social sciences and

the humanities. A further division into specific subjects is made for the three latter subject areas. The ranking methods employed for the different subjects vary to some extent. In some subjects, for example political science, the only indicator applied is the programme’s reputation, while in other subjects, for example medicine, a combination of reputation and information about aspects such as research activity and doctoral students’ admission scores is employed. The results of these rankings are published annually in “America’s Best Graduate Schools”.

U.S. News and World Report offers rankings based both on individual indicators and a weighted index. On the magazine’s website, for example, it is possible to select separate rankings based on indicators such as the admissions ratio, degree ratio, ethnic diversity, the proportion of international students, the proportion of students who live on campus and the proportion of students over the age of 25. Even more rankings based on separate indicators are available in the printed and Internet versions of “America’s Best Colleges” (for which a charge is made).

Perhaps U.S. News and World Report is best known for its combined weighted ranking listing the “best” higher education institutions. Since the late 1980s, this weighted ranking system has been based on two main sources: statistics provided by the institutions concerned and a questionnaire answered by university presidents and other higher education staff. In recent years, this index has comprised seven broad categories of indicators to which various weightings are applied, depending on how important they are considered to be. In the 2008 ranking, the following indicators were included in the weighted ranking:

• **Academic reputation (25 per cent weighting).** This indicator is the result of a questionnaire completed by university presidents and other senior representatives of the institutions concerned who rate the programmes provided by the other institutions on a scale of 1–5, in which “1” represents “marginal” and “5” denotes “outstanding”. In the spring of 2008, this questionnaire was distributed to 4,272 persons, and the response rate was 46 per cent.

• **Throughflow (20 per cent weighting).** This indicator is based on two measures: the proportion of new students who continue their studies at the institution concerned in the following academic year, and the proportion of a class of new students that takes a degree within six years.

• **Faculty resources (20 per cent weighting).** This indicator is based on several measures, with different weightings within this category: the proportion of classes with less than 20 students, the proportion of classes with more than 50 students, the average faculty salary, the proportion of professors with the highest academic qualifications in their area, the student/teacher ratio, and the proportion of full-time teachers and researchers.
• **Students’ competence (15 per cent weighting).** This indicator includes several different measures with different weightings: new students’ results in the SAT and ACT admissions tests, the proportion of new students in the 10 or 25 highest percentage categories, respectively, in their senior high school [upper secondary] class, and the number of applicants for student admitted.

• **Financial resources (10 per cent weighting).** This indicator denotes the amount of money expended in each year for each student on education and research activities, student services and other education costs. Expenditure on leisure activities, accommodation, health care and other services is not included.

• **Degree ratio (5 per cent weighting).** Only applies to national universities and “liberal arts” colleges. Indicates the ratio between the proportion of students who are expected to take a degree and the proportion who are actually awarded a degree in a given year.

• **Financial contributions from former students (5 per cent weighting).** The average proportion of former students who make financial contributions to the institution concerned. This indicator is justified on the grounds that it measures the extent to which students were satisfied with the institution.

If we study these seven indicators more closely, it appears that one indicator (students’ competence) is a measure of the characteristics of new students admitted to the institution concerned. This is applied purely as a quality indicator, and does not check the student’s initial characteristics. Two indicators (faculty resources and financial resources) may be regarded as a measure of resources, while a further two indicators (throughflow and the degree ratio) reflect education results. There are also two indicators (academic reputation and financial contributions from former students) that may be regarded as a measure of the institution’s reputation (among other academics) or popularity (among former students). On the other hand, there are no indicators for the form or content of the education programmes provided, that is to say the input/output process. And, similarly, there is no measure covering the final outcome or the quality of research, and no indicators for student-welfare factors.

These seven indicators carry different weightings, depending on how important they are considered to be, and the components included in the indicators concerned also have different weights. The first step in computation of an institution’s ranking is the weighted total of all the points, and subsequently these figures are adjusted on a new scale in which the best university or college has a rating of 100, and the remaining universities or colleges are ranked as a proportion of this value, and finally the institutions are ranked in declining numerical order.
Strengths and weaknesses

One advantage of the U.S. News and World Report ranking of universities is that it provides a quick and clear picture of a large number of American higher education institutions. It would be difficult and time-consuming for the individual citizen to achieve such an overview independently. Another advantage is that the institutions concerned are divided into different categories, depending on their focus. This means that only institutions of the same type are compared with each other, and hence the ranking avoids some aspects of the “one size fits all” problem. U.S. News and World Report presents the indicators used and they way they are weighted in a clear and open manner. In addition, its website provides opportunities for the construction of individual ranking lists, based on the indicators considered to be particularly relevant.

There are also obvious weaknesses in the U.S. News and World Report rankings, however. In the first instance, the reason why certain indicators have been selected and others excluded is not clearly explained. In many cases, it is a question of feeding in information that is easy to obtain, rather than producing relevant “quality” indicators. As already mentioned, there are no indicators covering the education process and final outcomes, although it may be assumed that both these factors are relevant for potential students.

The indicators also suffer from certain validity problems – that is to say they do not always measure what they purport to measure. Using academic reputation as a quality indicator is problematical, for example, since the academics participating in the questionnaire may, for example, lack information about the institutions to which they allot points, and that this may not be a neutral assessment. Financial contributions from former students are another problematic indicator, since this probably measures the willingness or the ability to make contributions rather than student satisfaction. This is also an indicator that is primarily relevant for private universities and colleges.

Assuming that the combined weighted index has the greatest impact, it is also important to consider how the various indicators are weighed up. The reason why the academic reputation indicator is given a full 25 per cent weighting may be questioned, for example, while the degree ratio only accounts for 5 per cent. This is especially problematical in view of the fact that academic reputation is usually regarded as a poor indicator of an institution’s quality. A common objection to the U.S. News and World Report’s ranking and other similar rankings is specifically the subjective, unsystematic weightings applied to the indicators. In September 1996, Gerhard Casper, the former President of Stanford University wrote a letter to the chief editor of U.S. News and World Report, expressing criticism on precisely these grounds.

Another problem which the U.S. News and World Report rankings share with many other university ranking lists is that differences between institu-

---

86. Several problems concerning the use of academic reputation as an indicator of quality are discussed in Chapter 3.
tions may appear to be greater than they really are. Certain differences between universities at the higher or the lower extremities of the ranking scale are not necessarily even statistically significant. This is another point raised by Casper in his letter.

“Could there not, though, be at least a move towards greater honesty with, and service to, your readers by moving away from the false precision? Could you not do away with rank ordering and overall scores, thus admitting that the method is not nearly that precise and that the differences between #1 and #2 – indeed between #1 and #10 – may be statistically insignificant?”

Another problem is that the indicators and the manner in which they are weighted often change from one year to the next. In one study, Marguerite Clarke notes that the U.S. News and World Report made between four and six changes per year in its college rankings over a six-year period. This mainly involved changes in weightings, definitions and methods, and indicators were only replaced or added to a limited extent. Clarke concluded that the changes in the U.S. News and World Report’s ranking methods make it virtually impossible to compare the ranking of a particular institution over time.

**Washington Monthly**

**Background and aims**

The Washington Monthly is a monthly periodical that mainly covers the American political scene, and it started to present university rankings in 2005 in the form of a research report which was, in principle, a protest against the established rankings for universities, and against the U.S. News and World Report rankings in particular. Commencing in 2006, the Washington Monthly has published its rankings in its September issue each year.

The Washington Monthly’s criticism of the established rankings is mainly based on the inadequate measures of academic quality, and the fact that they have a negative impact on higher education. Higher education institutions devote all their energies to trying to improve their positions in the various rankings, rather than focusing on “higher purposes” and helping to improve society. As a result, the Washington Monthly wants to present an alternative ranking that has favourable consequences for higher education institutions and for society as a whole. It has therefore applied a broad perspective in which the aim is not merely to rank institutions for students, but also from the taxpayers’ viewpoint – so that they know whether higher education institutions are using tax funding in a satisfactory manner and training students who will

---

90. Unless otherwise stated, the description here and in subsequent sections is based on information from the Washington Monthly’s website at www.washingtonmonthly.com
91. It is not clear, however, if a ranking was also conducted in 2008. We have been unable to find any information about this.
help to make the United States competitive in a globalised world. The Washington Monthly claims that its ranking is not just a question of what universities and colleges can do for students, but also what universities and colleges can do for society.

**Ranking characteristics – focus, indicators and weighting**

Higher education institutions are divided into two categories in the Washington Monthly rankings: National universities, and colleges offering “liberal-arts” subjects. On at least one occasion, this magazine also ranked “community colleges”, that is to say local colleges that focus on first cycle education. The characteristics that the Washington Monthly aims to cover in its ranking are divided into three categories:

1. **Social service.** This involves the extent to which the institution concerned encourages the students’ desire to serve their country, and it comprises three measures: the size of the institution’s Reserve Officers’ Training Corps programme in relation to the institution’s size, the proportion of former students working in the Peace Corps in developing countries, and the proportion of Federal “work-study” programme grants (that facilitate simultaneous working and studying on the part of students) that the institution devotes to social service projects.

2. **Research activities.** This category covers the institution’s research endeavours, and it also comprises three measures: the overall investment in research at the institution concerned, the number of defendants of doctoral theses in the natural science and engineering faculties, and the number of students at the first and second cycle levels who enter third cycle programmes.

3. **Social mobility.** This covers the institution’s commitment to low-income students, and it includes two rather complex aspects. One measures the proportion of students with Federal Pell programme scholarships, which is supposed to indicate which universities and colleges try to recruit low-income students. The other measures the degree ratio for such students. The students’ average SAT score is included as a control variable in both cases.

The above description indicates that one indicator – the proportion of students with Pell scholarships – may be regarded as a measure of new-student characteristics. Other indicators measure resources – namely the amount invested in research, and possibly also the size of the Training Corps programme and the proportion of Federal grants devoted to social service projects. Other indicators, such as the number of defendants of doctoral theses, the number of students who move on to third cycle studies and the degree ratio for students with Pell scholarships, may be regarded as measuring the results of the education provided. The proportion of student volunteers working in developing countries may be considered to be a measure of final outcomes.
On the other hand, there are no indicators that cover the actual education process. In point of fact, this magazine does not include any indicator that reflects the quality of the education provided. According to the Washington Monthly, this is because it is virtually impossible to obtain reliable information about how much learning goes on in American universities and colleges. The information that exists – for example in the form of tests that measure students’ actual knowledge – is not available to the general public, and the Washington Monthly says that “Until we have good information, we’d rather stay silent than try to go down the path of U.S. News in devising oddball heuristics”.

It is also worth noting that in some cases the Washington Monthly uses absolute numbers instead of relative figures in its ranking indicators. This means that the larger institutions are favoured in the rankings, due to their size, and this is also the intention. The Washington Monthly justifies its method on the grounds that it is precisely a large number of doctoral students and doctors, in combination with major investments in research, that can help the United States to retain its competitiveness in an increasingly globalised world. Furthermore, in contrast with many other rankings, the various indicators in the Washington Monthly’s list carry an equal weighting in the final index. This is said to be because all indicators are equally important. Separate rankings based on the three main categories are also presented, in addition to the overall index.

The differences between the rankings produced by the Washington Monthly and the U.S. News and World Report mean that the results also differ considerably.

Texas A&M, the University of California in Los Angeles and the University of California at Berkeley head the Washington Monthly’s ranking list. But these institutions are ranked 62, 25 and 21, respectively in the U.S. News and World Report ranking. In the same year, only one of the national universities (Stanford University) appeared in the U.S. News and World Report top-ten list, and it was also among the 10 best national universities in the Washington Monthly’s ranking. Generally speaking, it may be noted that the State universities achieve better scores in the Washington Monthly’s rankings, while the private universities tend to dominate the U.S. News and World Report’s top-ten list.

92. It may be mentioned, however, that on one occasion the Washington Monthly made an attempt to cover the prerequisites for learning in a ranking of community colleges. In this case, data concerning degree ratios was used, and also information from the “Community College Survey of Student Engagement” which measures, among other things, learning activity and the interaction between teachers and students. This ranking was criticised, however, because this study was not considered to provide a basis for ranking. See Inside Higher Ed (20/8 2007).
Strengths and weaknesses

One advantage of the Washington Monthly’s ranking is that it questions the established university rankings in the United States, and it offers an interesting alternative. This magazine presents relevant criticism, primarily of the U.S. News and World Report’s rankings, and it maintains, quite correctly, that universities and other higher education institutions are not merely for students, and that they should also contribute to the development of society as a whole. The rankings that the Washington Monthly has conducted also clearly indicate that the choice of focus and indicators has a considerable impact on the ranking results achieved.

The question is, however, whether the Washington Monthly’s ranking actually measures how effectively universities contribute to the development of society as a whole. It is not clear why the social service, research and social mobility categories have been specifically selected to determine the social usefulness of higher education institutions, and why other relevant aspects have been omitted. Cooperation with industry, innovation research, internationalisation, ethnic diversity and gender equality are a few examples of other aspects that could be important for the impact of higher education on society.

If we consider the indicators used in the three categories, we can see that there are also validity problems. One example is in the social service category, which is supposed to measure how well the institution concerned promotes a desire to serve one’s country. The question is, however, to what extent the indicators included – the size of the Reserve Officers’ Training Corps, the proportion former students working as volunteers in developing countries and the proportion of Federal grants devoted to social service programmes – reflect this.

Furthermore, the lack of an indicator for the quality of education is a serious limitation. The Washington Monthly defends itself by claiming, quite simply, that no satisfactory measure of education quality and learning is available. But the question is whether this argument is valid. Firstly, it is not clear why the Washington Monthly places such stringent demands on the quality indicator when other indicators in its index have more or less the same weaknesses. Secondly, it may be argued that it is worth trying to measure the quality of the education provided, even in the absence of perfect data. This is clearly a key aspect in many respects, not merely for potential students but also for society as a whole. It is hard to envisage, for example, that an institution’s investment in research is of any great value if we fail to take into account the quality of the research conducted.

Princeton Review

Background and aims

The Princeton Review is an education company that is mainly concerned with advisory services and preparations for the various national admissions tests
such as SAT.\footnote{93} In other words, it is not a periodical and it has no links with the Princeton University, as its name might suggest. This company was founded in 1981, and it has conducted university rankings since the early 1990s. Its rankings are published both in the “Best Colleges” guidebook and on the company’s website.

Princeton Review states that the purpose of its rankings is to give potential students information that is as uncensored as possible concerning many different aspects of the various institutions covered. As a result, its ranking is based on information from students studying at the institutions concerned. In addition, Princeton Review makes a special point of presenting separate rankings for a number of different areas, for example academic quality, demography, politics, quality of life, leisure and social activities. According to the company, this is because no single institution is best in all respects, and that different students are interested in different factors when they choose a university or college. Academic quality is not the only aspect taken into account and, as a result, student welfare factors are also presented in the Princeton Review rankings.

**Ranking characteristics – focus, indicators and weighting**

The Princeton Review rankings are based on a national student questionnaire survey conducted on an annual basis. This questionnaire does not cover all universities and colleges – only the most prominent institutions. The 2008 questionnaire covered 368 universities and colleges that were selected because they were considered to maintain high quality standards, both in academic terms and in other respects. The selection process was based on data from more than 2,000 universities and colleges and on visits to higher education institutions, and also on discussions with a large number of administrative staff at these institutions. The underlying idea was that the selection should represent a broad spectrum of universities and colleges (i.e. both private and State institutions, universities and colleges in different parts of the country, and institutions of different sizes).

The questionnaire is addressed to all students currently studying at the institutions selected, or who have studied there at some period during the previous two years. The questionnaire may be completed directly on the Internet, but it is also distributed in paper format to the institutions concerned. During 2008, the questionnaire was answered by about 120,000 students at 368 institutions all over the United States. This represents an average of 325 students per institution. Roughly 95 per cent of the respondents completed the questionnaire via the Internet.

The questionnaire poses more than 80 questions in a number of different areas, for example academic quality, demography, politics, quality of life, leisure and social activities. Most of the questions offer several response options

\footnote{93. Unless otherwise stated the description here and in subsequent sections is based on information from the Princeton Review’s website at www.princetonreview.com}
on a 1–5 scale, but there are also open questions which give the respondents an opportunity to explain their viewpoints. The questions asked fall into seven categories:

1. **Academic qualities and administration.** This includes questions about the atmosphere in the classroom, how much the students study, how accessible the teachers are, how good the teachers are at teaching, to what extent discussion features in the lessons, how good the study and career counselling is, how good the library is, and how smoothly the administration functions.

2. **Quality of life.** In this case, the questions concern, for example, how satisfied the students are, do they find the campus area attractive, and what do they think about the accommodation and the food.

3. **Politics.** Questions under this heading include how popular activist political groups are at the institution concerned and what are the political views held.

4. **Demography.** The questions involve the extent to which the student group is perceived as ethnically and socially varied, how much students with different social and ethnic backgrounds mix with each other, and how tolerant the students are as regards homosexuals.

5. **Social life.** This includes questions about views about the town or city in which the institution is located, and relations between students and local inhabitants.

6. **Leisure.** Under this heading, questions are asked about matters such has how good the sports facilities are, and how popular various sports and cultural events at the institution are.

7. **Parties and entertainment.** In this case, the questions concern, for example, how much alcohol students consume and how popular the various student associations are.

This summary indicates that several of these factors – quality of life, social life, leisure and parties and entertainment – reflect various types of student welfare characteristics. In the case of the more specific factors, it may be noted that there is at least a certain element of new-entrant features, such as the students’ ethnic or social backgrounds, and perhaps the students’ political views and tolerance of homosexuals might also be included in this category. There are also questions that measure resources, for example teacher accessibility, library resources and administration, and possibly also student welfare aspects such as study and career counselling, accommodation, food and leisure activities. There is also an indicator for results, in the form of the question about how satisfied the students are. In addition, the ranking includes some indicators that may be considered to measure the education process, for example questions about the classroom atmosphere, teaching skills and encouragement of discussion. On the other hand, there are no indicators for the final outcome, the institution’s reputation or the quality of research.
When Princeton Review has collected its questionnaires, each multiple-choice response is weighed up to produce an average value for each institution. The 20 best institutions are subsequently ranked on the basis of each multiple-choice question, and the result is some 60 ranking lists, each containing 20 institutions. Some questions are jointly weighted to produce broader categories with their own ranking lists. Comments obtained from the responses to the open questions are also presented, in parallel with the ranking lists. Since the questions cover such different areas, different institutions are included in every list. In the ranking conducted in 2008, Stanford University headed the list for the best classroom atmosphere, for example, while Northeastern University in Boston was at the top of the list for the best career and occupational guidance, and the University of Florida was rated as the best party-school.

When the results have been compiled, they are sent to the institution concerned, which then has an opportunity to make comments and suggest amendments before the results are published. In addition, students can subsequently comment on the results, in order to get some idea of the validity of the results. The follow-up studies that have been carried out indicate that a high proportion of the students – about 80 per cent tend to agree with the impression given by their institution.

Strengths and weaknesses

The Princeton Review ranking has several advantages. Firstly, the information collected is the result of a student questionnaire, resulting in information about how students feel about their studies and the time they have spent at a specific institution. Secondly, this ranking covers several aspects that are often not included in rankings, for example various student welfare questions and questions about the education process itself. Another strength is that each aspect of the rankings is reported separately, and is not jointly weighted to produce an overall index. This means that potential students can focus on the aspect in which they are most interested. Finally, the feedback that Princeton Review gives to both institutions and students in order to assess how reliable the questionnaire results are may also be regarded as an advantage.

But asking students about their experiences can also be a disadvantage, however. Firstly, it is by no means certain that the students have enough information to answer the questions. This applies, in particular, to the questions about other students at the institution concerned. It is not clear, for example, whether the respondents can do more than guess when asked how popular activist political groups are or the extent to which people from different ethnic and social backgrounds mix. Secondly, it is by no means certain that students have any reason for being completely honest about matters that might be negative for the institution of their choice. They may, consciously or unconsciously, want to convey a positive picture of the institution concerned, for example to potential employers.
There are also obvious weaknesses in the validity of the questions posed. One may wonder, for example, if the best way of finding out about social life is by asking what students feel about the town or city in which the institution is located and what relations with local inhabitants are like. When it comes to academic quality, several relevant questions are asked, but there are no questions about key aspects such as education methods and forms of teaching.

There are also some deficiencies in the data collection and analysis methodology. One problem is the narrow range of institutions. It is not clear why Princeton Review has decided only to include prominent institutions and, more specifically, on what grounds was this decision based. Furthermore, the ranking lists do not report the results for all universities and colleges in the sample, but only the 20 best in each category. This means that the results are not particularly useful for students who want to get a picture of a wider range of institutions. Another disadvantage is that the student response rate is not clearly indicated. This makes it hard to judge the reliability of the questionnaire data. However, one common problem with student questionnaires is that the response rate is low and, at a guess, the Princeton Review survey is no exception in this respect.

**Criticism and protests against ranking in the United States.**

University rankings in the United States have been subject to severe criticism, particularly by higher education institutions. The criticism involves both the purpose of ranking and various methodological problems, particularly in connection with the use of academic reputation as an indicator. The protests have almost exclusively been directed at the U.S. News and World Report rankings, presumably because they are the most well-known and are considered to have considerable influence, but also because they have more problematical aspects.  

These protests were initiated in the early 1990s, when Reed College in Oregon started to boycott the U.S. News and World Report rankings. Reed College said that it been unhappy about this magazine's ranking methods, even though Reed was ranked among the top-ten in the country in its category. This dissatisfaction was reinforced in 1994 when the Wall Street Journal revealed that several higher education institutions had manipulated their data to improve their ranking positions. As a result, Reed College announced that it would no longer supply data to U.S. News and World Report.

Students at Stanford University demonstrated their support for Reed’s boycott by forming the “Forget U.S. News Coalition” (FUNC) in 1996. This group then expanded to include other universities and colleges in the United States. FUNC also included Gerhard Casper, Stanford’s former President.

---

94. This entire section is largely based on the following articles from the “Inside Higher Ed” Internet newspaper: More Momentum against ‘U.S. News’ (20/6 2007), Refusing to Rank (17/8 2007), ‘U.S. News’ Adds Surveys that could Alter Methodology (9/4 2008) and ‘U.S. News’ sees Drop in Participation (22/8 2008).
who played an active part in the ranking debate. In September 1996, he sent an open letter to U.S. News and World Report, in which he pointed out that the ranking results were misleading.

In the same year, the former President of Alma College in Michigan, carried out an investigation of the U.S. News and World Report ranking methods for 158 colleges. This report indicated major deficiencies in the reputation questionnaire, and that many college presidents had insufficient information about the institutions they were supposed to grade. In the following year, the Chairman of Alma College contacted 480 colleges and urged them to boycott the U.S. News and World Report ranking on the grounds that the reputation questionnaire was based on inadequate data.95

In 1997, Stanford University followed the example set by Reed College and Alma College. Stanford announced that it would supply objective information to U.S. News and World Report, but that it would not participate in the reputation survey. In the following year, Stanford started to publish an alternative database on its website under the heading “Stanford University Common Data Set” which, it was considered, provided a better basis for comparison of higher education institutions by potential students and their parents. Eventually, however, FUNC was dissolved and Stanford began to participate in the reputation questionnaire again.

Protests against the U.S. News and World Report rankings have been renewed in the past decade. In May 2007, 12 college presidents wrote to hundreds of other presidents, urging them not to respond to the reputation questionnaire, and not to use ranking results in their marketing. At the annual meeting of the “Annapolis Group” – an organisation for “liberal arts” colleges – in June of the same year, this question was discussed further. As a result, a majority of the approximately 80 presidents who attended the meeting announced that they no longer intended to participate in the reputation survey. Instead, they discussed the possibility of developing an alternative database for information about universities and colleges aimed at potential students.

Participation in the U.S. News and World Report reputation questionnaire has declined significantly in recent years, presumably due to growing dissatisfaction on the part of higher education institutions. Previously, roughly two-thirds of college presidents participated in this questionnaire, but this proportion dropped to 58 per cent in 2006, and declined to 51 per cent in 2007, and 46 per cent in 2008. The response rate has declined markedly in the case of “liberal arts” colleges, which have protested particularly vigorously in recent years. The U.S. News and World Report’s chief-editor has refuted this criticism and these protests, and defended the use of reputation questionnaires. A leading article of 22 June 2007 stresses that the reputation of higher education institutions has proved to be a key factor in students’ opportunities to obtain employment after completion of their studies, and that this method is well-

established, for example in the business world, where business leaders are often asked to rank their competitors. In order to mollify such criticism, U.S. News and World Report has also considered certain changes in its ranking methods. In the 2008 reputation questionnaire, respondents had an opportunity to suggest how ranking methods might be improved. A new question was inserted, in response to criticism that the existing ranking was too static. This question was designed to show which universities and colleges had improved their performance.

**Canada – a long and stormy tradition of university ranking**

The ranking of higher education is a well-established phenomenon in Canada. One of the most well-known rankings of Canadian universities is conducted by the Macleans magazine, and was inaugurated in 1991. As time has passed, the Macleans ranking has been extended into new areas, the most recent example being a special ranking for law schools.

Ranking activities are somewhat controversial in Canada. Relations between ranking entities and universities have been very hostile, and this has had a considerable impact on the rankings that are carried out. This criticism and its consequences are summarised under a separate heading at the end of this section.

**Macleans’ ranking of Canadian universities**

**Background and aims**

The Macleans magazine ranks Canadian higher education institutions in November each year, and this is published in a special issue in book format (on paper and digital) and on the Internet. In the autumn of 2007, this ranking covered 47 universities, but it excluded universities with less than 1,000 full-time students or which did not admit all types of students (for example on religious grounds). In recent years, Macleans has experienced difficulties in obtaining information from universities, and has been obliged to adjust both its methods and its indicators.

The target group for the Macleans ranking is pupils in their final year of upper secondary education who plan to move on to a university education. The aim is that this ranking system should be an evaluation of the academic quality of Canadian universities, in which Macleans tries to give potential students the information they require to choose the university that best meets their needs.

On various occasions during the year, Macleans publishes an issue on university studies, ranking and students’ experience of education at Canadian universities. In addition, Macleans presents extensive information concerning the ranking of programmes on its “On Campus” website;

students can also find information about the labour market and employers, the financing of studies, the advantages of higher education, and advice about how to select the right programme, based on their interests or plans for the future. This website also displays rankings produced by other parties, for example what college students in British Columbia or Ontario think about their programmes. In addition, the Macleans ranking is supplemented, for example, by information about the proportion of international students, the proportion of students from other provinces, the proportion of new students who continue their studies after their first year, the average score for admissions and the number of students awarded a degree.

The rankings conducted by Macleans in recent years are divided into three categories, based on the type of university concerned:

- Universities that primarily provide first cycle education (Primarily Undergraduate University Rankings).
- Universities that provide a broad education in programmes at both the first and second cycle levels, and which also including professional and research programmes (Comprehensive University Rankings).
- Universities with a wide variation of research areas and third cycle programmes, and which offer medical training (Medical Doctoral University Rankings).

Macleans considers that the universities in these three groups have different prerequisites and, as a result, does not compare them with each other.

A tool which Macleans calls the “personalized university ranking tool” has been available for the past two years or so on the Internet, in which a user can choose up to seven quality indicators, weight them in terms of their importance, and then select a number of Canadian universities which can be compared on the basis of the indicators chosen. This tool was developed in response to criticism from the universities, since it was considered that the rankings failed to provide scope for potential students’ individual priorities. One important difference between this tool and the Macleans university ranking is that various types of universities can be compared with each other, and that they are not divided into different categories.

**Ranking characteristics – focus, indicators and weighting**

Previously, Macleans employed 22–24 indicators to rank universities into three groups. In the latest ranking, however, this has been reduced to 13 indicators in six areas. Furthermore, the group covering research and medical programmes was assessed on the basis of an additional indicator, namely the total number of volumes in the library (1 per cent weighting).

This information provides a basis for rankings obtained from sources to which the public has access. In other words, such information is no longer obtained from the universities. The ranking’s website includes clear information about the way this information was collected and its source. Each indica-
tor is weighted (see the figure in parenthesis for each indicator), but the bases for weighting are not disclosed. The indicators covered by the ranking may be divided into six areas:

1. **Students and classes (20 per cent weighting).** This indicator provides information about the student/teacher ratio and how successful the students have been in the past five years in terms of awards and scholarships, for example Fulbright and Rhodes scholarships.

2. **Teaching staff (18 per cent weighting).** This indicator covers how successful teachers have been in the past five years in achieving national awards, for example “3M Teaching Fellowships” and “Royal Society of Canada Awards”, and in receiving research funding. Research funding is divided into the following categories: humanities and social sciences, and medicine and natural sciences.

3. **Resources (12 per cent weighting).** This indicator provides information about how much money to which the institution concerned has access per full-time student, and the total amount of research funding available in relation to the size of the teaching staff.

4. **Student support (13 per cent weighting).** This indicator is designed to evaluate the support available to students, in the form of the proportion of the university’s budget devoted to various types of student services and scholarships.

5. **Library (15 per cent weighting).** This indicator provides information about library facilities: how many books are available (only applied to the research and medicine category), the number of books per student, the library’s budget including funds for digital library resources, and the budget for purchases of new books.

6. **Reputation (22 per cent weighting).** Macleans asks upper-secondary school principals, business leaders, student counsellors and others about the university’s reputation. The respondents rank the university in three areas: highest quality, the most innovative, and “future leaders”. The results for these three categories are summarised in a “Best overall” score.

Most of the Macleans indicators in the 2007 ranking cover resources for education and research, although one indicator presents educational results (students who receive prizes and scholarships) and a further indicator reflects assessments of the reputation of the institution concerned. However, there are no indicators that deal more directly with the education process and the quality of education provided, for example indicators for teaching methods, placement opportunities and the interaction between students and teachers – in other words there are no process indicators (unless the prizes awarded to teachers for teaching skills are regarded as an indirect measure of this aspect). The Macleans ranking includes student welfare factors in the form of the proportion of the university’s budget devoted to student support and scholarships.
Many researchers, including Geoff Scott⁹⁷ and the group that developed the Australian CEQ 2001⁹⁸, maintain that successful studies are affected by many factors other than those that occur in lecture halls and laboratory sessions, and that student welfare aspects are crucial in this context.

The Macleans ranking has no indicators for new-student characteristics and final outcomes on completion of studies in pay and labour market terms, such as whether the education provided offers an all-round education, a capacity for lifelong learning or other general skills (the final outcome and results of education).

Macleans maintains that its rankings evaluate the academic quality of universities. The degree of success of teachers and students in winning prizes, scholarships and acquiring research funding are some indicators of this aspect, and also of the reputation factor. The success of teachers in the research field is a common indicator in the ranking context, and possibly the extent to which they are cited in prominent academic journals. On the other hand, only a limited number of ranking lists include an indicator for student successes in winning prizes and scholarships.

When the data for the various indicators is collected and weighted, each university has a total score that determines the university’s position in the ranking list.

**Strengths and weaknesses**

The question is does the Macleans ranking do what it claims to accomplish: in other words, does it evaluate academic excellence or quality? Four out of six indicators (students, teaching staff, resources and reputation) are more or less linked to academic quality, even if they are not particularly good indicators. Academic excellence is often associated with successful research, although this is also considered to include skills in supporting students in their learning process or transforming research into benefits for society as a measure of academic excellence. In this spirit, Macleans also counts prizes for teaching proficiency as a resource indicator. Applying measures such as research funding and scientific awards is more common. Reputation is a questionable but frequently used indicator.

The question is also whether this ranking provides the information that potential students need to choose a university. The fact that the ranking has no indicators for the quality of teaching and only has one indicator each for future prospects after completion of studies and for student welfare factors is one weakness in this context. In addition, there are no indicators for the quality and focus of research. Some students may wish to continue into third cycle studies and may therefore be interested in information on matters such as the research environment at various universities. This would require a ranking by

subject rather than an aggregated ranking at the institution level, if such a ranking were to be genuinely useful for students. An aggregated ranking does not reveal the quality of the individual subject areas.

The Internet-based tool offered by Macleans allows students to build their own rankings. It is true that this does not change the choice of indicators, but it does enable students to compare universities in terms of what they consider to be the most important indicators, since they can choose the indicators that they wish to include and apply their own weightings. This increases the relevance of this ranking for student choices. This tool enables students to rank universities based on up to seven indicators, it provides a satisfactory overview, and the tool is also easy to use.

At a more general level, the Macleans ranking method offers one advantage in comparison with many aggregated rankings of higher education institutions, since Macleans divides institutions into groups according to their education level and type of activity. As a result, institutions with quite different objectives and prerequisites are not competing with each other for ranking positions. In practice, however, comparisons are made by applying the same indicators in the three groups, and this means perhaps that these rankings may not be so fair after all. In addition, there is still the problem that different subject areas are combined and compared jointly at the institution level. Hamish Coates has shown that universities perform differently depending on the range of programmes they offer, and that this means that aggregated rankings of institutions are misleading.

In common with other ranking lists, Macleans provides an illusion that there is a definite quality difference between universities with different positions on the list. In reality, however, this difference may be very small, and it may be due to the weighting of the indicators. Furthermore, the basis for the Macleans weightings is not clear, even if the weightings applied are clearly indicated. For example, Macleans gives the “reputation” indicator a 22 per cent weighting, but does not explain how it has arrived at precisely this value. The Internet ranking tool allows users to apply their own weightings of the indicators and hence neutralises this problem to some extent, but the problem still occurs in the magazine and book versions.

Furthermore, several of the indicators used by Macleans are incomplete. The library indicator, for example, provides no information about access to periodicals or staff who can help students to make use of the library’s resources. Student support, which is based on the proportion of the budget devoted to student support, is another problematical area. This indicator does not indicate the degree of efficiency with which these funds are used, or whether the student services provided are what students need and demand. Similar criticism may be directed at other indicators used in this ranking.

Attention has also been drawn to problems concerning the statistical data employed – whether the statistics are correct, how transparent they are, and if they can be compared between institutions and over time. In its most recent ranking, Macleans only uses official statistics collected from a more limited number of sources than in the past, and not from individual institutions. It is not possible to assess the degree of quality assurance of the statistical data on the basis of the information provided by Macleans. Similarly, it is not possible to compare the results over an extended period, since the methods and the indicators have changed several times.

University rankings under fire

A couple of years ago, criticism from a number of Canadian universities switched from words to action. They refused to supply the information required to Macleans, and Indira Samarasekera, the President of the University of Alberta., conducted an open debate with Tony Keller, the editor of the Macleans rankings, on the Inside Higher Ed forum. Macleans faced united opposition on the part of 22 universities. The universities criticised Macleans’ methods, but offered expert assistance. Indira Samarasekera also claimed that providing statistics in the format required by Macleans called for considerable resources that were needed for more important activities, and that the universities considered that tax-funding should not be used to finance a commercial product. Tony Keller replied that potential students are entitled to information and that the universities had failed to report the statistics in a transparent and readily accessible form. In the event, Macleans was obliged to reduce the number of indicators in its ranking as from 2006, to obtain statistics from other sources and, in some cases, to use less up-to-date information.

Macleans was also forced to reformulate a new initiative that would have involved investigating recent graduates’ views on their experience of their programmes in a questionnaire format. The universities refused to participate in this project, so instead Macleans cited legislation requiring freedom of information and demanded that the universities release the results of the National Survey of Student Engagement (NSSE) and the Canadian Undergraduate Survey Consortium (CUSC). These two surveys are commissioned by the universities to assist in assessment of the programmes and the services provided. The NSSE poses questions to a sample of students concerning the way education programmes are conducted, with comparisons in five areas and subsequent benchmarking. The CUSC investigates the extent to which students are satisfied with their education. As a result of this dispute, the results of these surveys are now available on the Macleans website, and are marketed in parallel with the magazine’s own ranking.

102. NSSE website at http://nsse.iub.edu/index.cfm
103. CUSC website at www.cusc-ccreu.ca/hime.htm
Indira Samarasekera’s comments on the way in which students use rankings to make their choices may be of interest in this context. In her article, she states that the number of applications to the University of Alberta increased by 36 per cent between 2006 and 2007, at a time when this University was not participating in the Macleans ranking but had not publicly announced that it was boycotting it.

**Australia – from student questionnaires to ranking**

Ranking is a contentious issue in Australia and it is also contradictory. In parallel with a continuous and critical discussion of university rankings, Australian universities also use ranking results in their marketing. In addition, the government has also developed a form of ranking system to determine the allocation of funding to the higher education institutions with the best teaching performance. At the national level, there are both pure ranking systems and also other surveys that do not provide direct rankings, but which may be used to arrive at rankings for education programmes and institutions.

This section commences with a description of the Australian Graduate Survey (AGS), which was conducted for the first time 30 years ago and which makes important contributions to current Australian rankings. Subsequently, there are descriptions of two ranking systems which both utilise the AGS results but have totally different aims, namely the Good Universities Guide, which is aimed at potential students, and the Learning and Teaching Performance Fund (LTPF), which is a government initiative to reward good teaching methods. Each description is followed by a discussion of the strengths and weaknesses of the ranking concerned.

**The Australian Graduate Survey**

**Background and aims**

Australia has extensive experience of investigating the extent to which university education lives up to student expectations – in other words it is a form of quality measurement. As early as 1971, the Australian universities started to ask their students about work and continued studies, and about employment and pay after they took their degrees, in the form of a Graduate Destination Survey (GDS). Right from the start, the universities employed a questionnaire format, mainly to get in touch with their former students, but the results were obviously also of interest for potential students, their parents and study and career advisers. In addition, this survey is considered to indicate trends in the labour market.

A questionnaire was developed which was intended to supplement the labour market orientation of the GDS, and which looked at education results based on student opinions and evaluation of the programmes and courses in which they had participated in the form of a Course Experience Questionn-
aire (CEQ). A corresponding questionnaire has been developed for graduate doctoral students in the form of a Postgraduate Research Experience Questionnaire (PREQ), but this is not discussed in this context. The two surveys—establishment in the labour market and continued studies, and opinions about the education provided—are carried out jointly, and are collectively referred to as the Australian Graduate Survey (AGS). They are sent to students a few months after they have taken their degree. The AGS is backed by both the government and the Australian Vice-Chancellor’s Conference.

The AGS survey is a university initiative and is coordinated at the national level by Graduate Career Australia, an organisation in which employers, universities and the government are represented.

The universities defray the major part of the costs. It is estimated that higher education institutions each pay AUD 150,000–250,000 per survey (representing a total of AUD 6–8 million), while the government contributes around AUD 500,000. All the Australian universities participate in the AGS.

The universities use the results in their internal quality processes, in order to improve their teaching and their marketing. For its part, the Ministry of Education uses the AGS information to assess and plan the requirements for higher education. Since 2005, the Ministry also uses this information to reward universities for their learning and education performance. The responsibilities of the higher education sector to taxpayers and to students who have paid their study fees is a further aspect of the AGS survey.

Data collection and analysis

The AGS involves a certain degree of centralisation, since some questions are mandatory (three out of eleven indexes). A single organisation receives and analyses the information from all the universities, and the results are reported to the Department. In addition, this information is published in a comprehensive form by Graduate Career Australia on the organisation’s website, and in reports sold commercially. The AGS offers higher education institutions some degree of individual adaptation, however, since they can tailor their questionnaires to comply with a specific range of questions, determine the method of distribution (e-mail, Internet or paper) and decide the way in which reminders are to be sent, and how often.

In practice, the institutions distribute the questionnaires to their graduate students in May and October—roughly four months after their major examinations. All Australian and international students (with the exception of students at “offshore branches”) receive the questionnaire, making a total of about 200,000 persons per year. The response rate varies from one institution to another, and ranges from 30–80 per cent. This figure is somewhat higher

---

106. www.graduatecareers.com.au
for Australian students than for international students – on average 62.5 of
the Australian students responded to the AGS questionnaire in 2007. Stu-
dents send in their completed questionnaires to the universities, and the raw
data is then submitted to Graduate Careers Australia, where the information
is analysed and the results for each institution are reported to the Ministry of
Education and the institution concerned.

Although the results are not presented in the form of a ranking list, it is per-
factly possible to compare education areas and universities on the basis of the
indicators. The commercial Good Universities Guide108, for example, uses this
data as part of its ranking of Australian universities. The government Learning
and Teaching Performance Fund (LTPF) also uses information from some
aspects of the AGS in its inputs for the three mandatory indexes.

Survey characteristics – focus, indicators and weighting

The AGS has been developed and refined over time – often after deficiencies
and needs have been identified by the universities and in the public debate.
Development has frequently been assisted by research results and tests, and
always in a dialogue with the higher education sector.

The AGS is a questionnaire in two parts. In one part, students provide
information about their occupations after they took their degree, and their
assessments of their education, with responses on a scale ranging from “agree
completely” to “don’t agree at all”. The questions posed have been impro-
ved over time, and new questions have been added. For example, one study
developed questions about general accomplishments, while another suggested
questions to cover new areas, such as IT support, intellectual challenges and
lifelong learning.109

In the section of the AGS concerning establishment in the labour market
(GDS), the questions involve the category of studies in which students took
their degrees, work during and after the study period (details such as skills/
qualification requirements and pay), any continued post-degree studies, and
the manner in which the student applied for employment. This section con-
tains some 40 questions in which students provide factual information rather
than an assessment of their situation after taking a degree. We will not exa-
mine the GDS in depth at this stage.

In contrast with the GDS, the second part of the AGS, the CEQ is evalua-
tive, and it has been subject to keen debate in Australia. If investigates stu-
dents’ experiences of the teaching provided and their opinions, and the skills
and accomplishments they have developed in the course of their education.
The students are asked to assess various statements on a five-grade scale rang-
ing from “agree completely” to “don’t agree at all”. Three areas in the CEQ
are mandatory, and the universities must include them:

a. **Good teaching** ("Good Teaching Scale"). In this index, students assess the teaching in courses or programmes. The statements involve, for example, feedback from teachers, whether the course contents were presented in an interesting manner, and whether the teachers had tried to understand student problems in the course concerned.

b. **General skills** ("Generic Skills Scale"). In this index, students assess the extent to which the course or programme helped the student to acquire the skills and attainments anticipated after completion of the course/programme. The statements involve, for example skills such as decision making, problem solving, analysis, the ability to express ideas well in writing, planning and tackling new problems.

c. **Overall satisfaction** ("Overall Satisfaction Item"). In this case, the students state their degree of satisfaction with the course or programme concerned. The assessment scale differs from that used in the other questions.

A further eight areas may also be included in the CEQ, if the university wishes to do so:

1. **Clear goals and levels** ("Clear Goals and Standards Scale"). This index follows up whether the students consider that they have received sufficient information about the learning targets for the course or programme concerned, and about the expected workload.

2. **Appropriate assessment** ("Appropriate Assessment Scales"). In this index students assess, for example, the extent to which the course or programme’s tests and examinations involved the presentation of facts (superficial learning) or understanding ("higher order learning").

3. **Reasonable workload** ("Appropriate Workload Scale"). This index reflects student assessments of their workload during the course or programme.

4. **Student support** ("Student Support Scale"). In this index, students evaluate their degree of satisfaction with the support facilities provided by the institution concerned, for example library services, study and career counselling, advisory services of various kinds, and other learning and health resources.

5. **Learning resources** ("Learning Resources Scale"). This index also reflects student evaluations of the resources that have assisted their learning process – for example the quality of course materials – but there is also a focus on support functions such as libraries and IT.

6. **Learning environment** ("Learning Community Scale"). This index attempts to determine whether students felt that the course or programme offers an environment that encourages students to investigate new ideas and share their knowledge and information in an intellectual and stimulating atmosphere, and whether the students themselves have participated in such an environment.
7. **Lifelong learning (“Graduate Quality Scale”).** This index assesses skills and attainments at a general level, for example whether students consider that the course or programme has encouraged them to be enthusiastic about further skills development, and to evaluate various perspectives and ideas. To some extent, this index resembles the index for general skills.

8. **Intellectual motivation (“Intellectual Motivation Scale”).** In this index, students assess the extent that the course or programme has motivated them for learning, and how intellectually stimulating it was. This index overlaps with both “Good teaching” and the overall rating for the course or programme.

The information collected is not weighted. The AGS results are subsequently presented by “Graduate Career Australia”, with a focus on the questions related to education areas and the labour market. Two brief labour-market summaries\[110\] are published every year, and users can search for this information in a service that provides information about the average starting salary for a person with a degree in a given subject from a particular university. Anyone interested in the opinions of students about their education can order a report showing the CEQ results from the “Graduate Career Australia” website.

Even if this website does not provide direct access to the AGS results, the information is available on order for interested parties who are prepared to pay. In other words, it is easy for other entities to obtain information that can be assembled to form rankings, based on the survey results.

**Strengths and weaknesses**

One of the advantages of the CEQ survey is that it has such widespread support. To some extent, this may be because the Australian higher education system is relatively small and homogenous, thus making it possible to conduct surveys with considerable backing from higher education institutions. Participation was originally based on a joint commitment to quality, quality improvement and a collective interest in marketing Australian higher education at an international level.

Another advantage of this survey is that the questions and scales used have been developed on the basis of needs, and backed by research. In addition, they have been tested and adjusted as the result of a dialogue with the higher education sector, and they are widely considered to measure what they are supposed to measure, even if there is some discussion about what the CEQ fails to measure. The CEQ has been criticised, for example, because it focuses too much on teachers, is not well adapted to the needs of distance education and adult students, and because the information is not sufficiently detailed for many subjects and years.\[111\]

---


\[111\] Craig McInnis, Patrick Griffin, Richard James & Hamish Coates (2001).
The main focus in the indicators and the indexes used in the AGS (i.e. both the GDS and the CEQ) is on the final outcome (establishment on the labour market and pay), the results of education (general skills and lifelong learning), and the resources for education and research (student support, teaching resources, teachers’ support for students).

There is also an indicator for student welfare factors (health care). On the other hand, the AGS provides very little information about the characteristics of new students, although some information of this nature is collected by the universities when students register for a course or programme.

Although the CEQ is concerned with teaching aspects, the indicators largely fail to reflect the learning process. There are questions about the student’s interaction with teachers, the learning environment, superficial learning and feedback, but the CEQ fails to fully capture the elusive element which is an intrinsic part of the intervening process between education resources and education results. Student-centred teaching methods are only investigated to a very limited extent, for example, and there is no indicator that covers the course syllabus.

The CEQ has also been criticised for excessive influence on the concept of what is good and effective teaching, and the factors covered by the CEQ have become the dominant paradigm in this area. The critics consider, for example, that the CEQ survey is far too superficial and limited, and that it is unable to register crucial nuances in the teaching environment, for example the quality of problem-based teaching environments. At the same time, others regard the CEQ as being based on well-established practice about what constitutes good education, and hence that it covers factors that may be considered to be quality indicators. Another argument put forward is that the CEQ responses have given the universities an external reference point for education quality that did not exist previously that can be used by higher education institutions when making comparisons.

Another advantage of data that is collected on a regular basis is that it facilitates comparisons over time, and makes it possible to make to identify patterns and trends. As already mentioned, information about the degree of establishment in the labour market can be used to identify labour market trends. Geoff Scott’s project on student motivation for learning is another example of this. He processed 168,000 CEQ questionnaire responses in this project – a volume of data that would be hard to conceive without access to the AGS.

The fact that the universities can tailor the questionnaires to some extent presumably makes them more useful at the internal level. Information from the survey may, for example, be analysed and compared with the information which the universities collect from students when they register for a course. Harris and James, however, question whether all this data is actually used to

113. Ibid.
improve teaching methods, and they note that the patterns in the information have scarcely changed in the course of time. On the other hand, Hamish Coates’ psychometric analyses of data from the three mandatory CEQ indexes – slightly less than 100,000 answers per year – have identified positive changes at one third of the universities between 2004 and 2005.

And the fact that “Graduate Career Australia” is responsible for analysis of the data collected rather than the universities themselves guarantees the independent character of the analysis, although distribution of the AGS by the universities is more open to criticism. The universities may influence the answers since they have contact with former students in the data collection process. Perhaps the student responses tend to be more positive than if some independent entity distributed the survey and was responsible for follow-up.

Another weakness is the response rate, which varies considerably between different institutions. In comparison with other student questionnaires, for example that used by the German CHE, a response rate of 30–80 percent is not unusually low, but the fact that some institutions only have a 30 per cent response rate is problematical for reliability aspects and the generalisation of results. The high response rate for some universities may be because they contact their graduates themselves and encourage them to respond. As already mentioned, this, in it turn, may also be problematical.

Another problem that the CEQ shares with other student questionnaires is that this survey places relies to such an extent on the students’ subjective experiences. This means that it is important to be able to generalise the results, and that they are used in the right way. If a university is to use the CEQ to improve and change its teaching methods, the student’s experiences and opinions are an important part of this process. On the other hand, if the results are to provide a basis for giving teachers permanent tenure or for the allocation of resources to the universities, reassessment of such utilisation may be called for.

One disadvantage of the CEQ which suggests some caution in the interpretation of the results is the delay factor. Students completing the questionnaire are saying what they thought about their education shortly after they took their degree, and the results are published one year later. In other words, the answers cover an interval of up to five years, and there may have been considerable changes in the education provided during this period. Another uncertain factor is that the students evaluate their education as a whole, which may lead to rationalisations after the event and similar problems concerning what the student recalls most clearly – either positive or negative experiences.

Following utilisation by the government of the AGS results for funding decisions, some differences in the way higher education institutions apply the survey may be noted. Some institutions, for example, try to improve their evaluation unit and increase students’ awareness of their learning process. In

---

Andry Onsman’s view, some universities focus on improving their scores in the survey rather than on improving the prerequisites for student learning.117

The Good Universities Guide

The Good Universities Guide is a commercial product that is available both in printed form (for sale) and on the Guide’s website, free of charge.118 The publisher, Hobsons, also issues other products and other information about university studies in Australia. The Good Universities Guide is primarily aimed at Australian students, and the website refers international students to other sources of information.

Anyone interested who opens a user account may access information concerning, for example, how difficult it is to gain admission to an education programme and the salary after completion, and information about MBA programmes and scholarships. The Guide ranks both state and private universities and colleges.

The website refers to both “ranking” and “rating” (i.e. assessments and scores). The information used is obtained from five sources: the Ministry of Education, Graduate Career Australia, the admissions centre in every state or territory, surveys of the quality of teaching (e.g. the CEQ), and various national statistics.119

The website allows users to compare and rank universities and colleges at four different levels (“rating gateways”):120

• Assessment of a university or college on the basis of background information (23 criteria) such as a student breakdown, the CEQ results and information about admissions.
• Comparison between different universities or colleges at the subject level, applying 21 criteria such as student-group profiles, cultural diversity, gender distribution, teacher/student ratios, quality of teaching and post-degree establishment on the labour market.
• Assessment and ranking at the course level, including how difficult it is to gain admission, specialisations, study fees, etc.
• Assessment of a university or college based on information about the campus area, for example services and student support and costs on the campus.

Where institutions and programmes are compared and ranked, this often takes the form of a star system, in which five stars are awarded for the highest rating-group and one star for the lowest category.

118. www.gooduniguide.com.au
Strengths and weaknesses

One advantage of the Good Universities Guide is that it is a multidimensional ranking based on the available data, and surveys of Australian higher education institutions and programmes. Another advantage is that ranking in five different groups, and is not a listing of universities or programmes in first, second or third place, etc. This avoids the problem of minor statistical differences between institutions having an undue impact on positions in the ranking list, and giving the impression that the differences are greater than they really are.

One weakness of this ranking system, however, is that the reader does not have a clear picture of the way the universities and colleges in the Internet-based tool have been selected. For example, someone who wants to compare different universities and colleges in Western Australia will come up with a list of eight institutions, including both state and private arrangers. But someone who clicks on Tabor College, for example, and the Australian College of Medicine will find information about very few indicators – mainly for the structure of the student-group, and nothing about the quality of the teaching. Considerably more information is available on the website for other universities in Western Australia.

Many information items appear to be very precise, but the degree of exactitude may be questioned, and it may also be difficult to compare different institutions since some information is not quality assured. Furthermore, some of the tables and rankings are at the institution level, which leads to limited usefulness from the potential student’s point of view. On the other hand, the Guide also gives students an opportunity to search for information and make comparisons at the course and subject level – and this is a positive factor.

Overall, the Good Universities Guide is easy to use, and the information is presented in a readily accessible manner. In addition, the website provides guidance about points to consider before choosing an education programme.

The “Learning and Teaching Performance Fund”

Background and purpose

In 2003, the Australian government announced an initiative to promote outstanding education environments. The Learning and Teaching Performance Fund (LTPF), which was one aspect of this initiative, was set up in 2005. The underlying idea was to encourage the learning and teaching process at higher education institutions. In the first year, 14 universities (out of the 24 that participated) shared AUD 54 million (approximately SEK 325 million). In the next year, the number of “winners” doubled to 30 universities (of 38 participants), and they shared total funds amounting to AUD 83 million. In the following year, AUD 113 million was allotted to 23 universities. The maximum funding awarded to a single university in this year was AUD 11 million. All state universities currently compete for LTPF funding, although prior to 2009, the government has announced that a maximum of AUD 74 million is to be allo-
cated, both to institutions that demonstrate that they have the best teaching and to institutions that have made the greatest improvements.

The allocation of these funds is based on an expert panel comprising representatives of the higher education sector who assess the results of the various institutions by applying seven indicators. The panel then places the universities in four groups for each education area and subsequently submits its recommendations for which universities are to receive LTPF funds to the Ministry of Education. The panel also submits written comments on its allocation proposal and on the process, which are then published on the Department’s website.121

The Minister of Education's presentation of the results of this evaluation every autumn attracts considerable media interest in Australia, even if the LTPF is not a ranking in the accepted sense. This does not prevent the newspapers from publishing rankings based on the results in terms of the amount of LTPF funding awarded to the universities concerned, and the universities use the results in their marketing.

Characteristics – focus, indicators and weighting

The selection of the education programmes that are to be rewarded for their teaching takes place in two phases. First, the universities that wish to participate must show that they have a strong strategic commitment to the learning process and to teaching. Normally the institutions that apply meet this requirement. In recent years, all 38 state-funded universities have participated. In the second phase, seven indicators are used to evaluate the university’s quality, of which two are based on the GDS, three on the CEQ and two on official statistics for higher education.

In each successive year, the government has changed the principles on which allocations are based. In the first round, for example, higher education institutions received points that were aggregated per institution – in other words a ranking. Following criticism by the universities, this principle was modified in the next year so that funds were allocated to four education areas, and not aggregated at the institution level. And in the first round, different weightings were applied to the indicators, but this was changed in the second round.

The indicators used in 2007 were:

- Student satisfaction with the teaching (CEQ)
  - General skills
  - Good teaching
  - Overall satisfaction
- Final outcome (GDS)
  - Full-time employment
  - Full-time or part-time studies at second cycle level
- Success (statistics collected by the Ministry of Education)
  - Continued studies

– Proportion of students approved for their courses

The indicators are applied to four education areas:

• Natural sciences, computer science, engineering, architecture and agricultural sciences
• Management studies, law and economics
• Humanities, arts and teacher education
• Health sciences and medicine

In order to ensure that the student breakdown in a given university does not affect results, this information is adjusted, on the basis of factors in which the national average is the point of comparison. These factors include gender, the proportion of full-time and part-time students, the students’ socioeconomic background, unemployment figures in the students’ home area, and the size of the institution. In the first round, 17 adjustment factors were applied, but this was reduced to eight factors in the third round (2008).

The results of this assessment are presented on the Ministry of Education’s website, both in the form of raw data before adjustment and with percentage adjustment per indicator, although funds are allocated on the basis of the adjusted information. The results are also presented per education area, and the institutions are sorted into four “bands”. Within each band, the institutions are listed in alphabetical order with no indication of the total score. The number of institutions per band varies over time, and depending on the education area.

Strengths and weaknesses

One advantage of the LTPF system is that more attention has been paid to learning and teaching issues, and that higher education institutions have focused on them to a greater extent at the internal level. Previously, funding was allotted purely on the basis of successful research endeavours, but first cycle education is now also a potential source of funding. At the same time, the experts are now warning that the substantial amounts involved – the University of Melbourne, for example, has received at total of AUD 27 million in three rounds of allocations – may mean that institutions concentrate on improving their ranking positions, and that this does not necessarily imply improved quality of teaching.122

On the other hand, an assessment of the LTPF process conducted by the Ministry of Education indicates positive changes in learning and teaching methods since establishment of the Fund. There is, for example, a focus on this area which has resulted in improvements, there is a national dialogue on teaching and learning quality, and students’ education results may have also improved. For the most part, the universities have normally utilised the addi-

tional resources provided by the Fund for broad initiatives that benefit many students and teachers, and only in a few cases in the education areas that have been awarded additional resources for the university in question. Such initiatives have involved, for example, increased support for student services such as mentor programmes and assistance for students at risk, and strategic personnel initiatives such as “Learning and Teaching Fellows”.123

Although the impact of the LTPF is disputed and hard to identify with any degree of certainty, the Fund has indubitably strengthened the teaching and learning process in many universities. This is a major achievement, and indirectly reflects on the Fund’s ranking of quality within the various education areas.

One of the weaknesses of the LTPF is that it rewards those who are already successful in the teaching and learning field, and this means that those who are less successful in this area must improve their performance without any additional funding. On several occasions, however, the panel of experts has rewarded institutions that have a high level of quality, although not sufficiently satisfactory in any of the specific education areas. Furthermore, the basic principles applied by the new Labour government differ from those of the previous government – institutions struggling to improve their performance but which fail to meet the required standards, will also be entitled to LTPF funding in the 2009 round.

One of the advantages of the LTPF is a relatively high degree of openness about the process and the assessments, although uninitiated persons may find it hard to understand certain technical details, for example how the adjustments are applied. Information of this nature is available on the Ministry of Education’s website.124

The fact that the evaluation process and grounds for such evaluation have been adjusted from year to year may be regarded as positive – this demonstrates openness and a desire to rectify weak points. On the other hand, the institutions do not know what applies from one year to the next which, in its turn, makes it difficult to establish long-term quality, if the aim is to achieve good results in the LTPF process. In addition, this makes it difficult to compare results from year to year. For example, the University of Wollongong, which was the surprise winner in 2006, came in sixth place in the Fund’s allocation list in the following year.

Another disadvantage is the time lag which – as already mentioned it takes from three to five years from the point at which students provide information about their experiences of the institution’s teaching (as new students) to the reward disbursed to the university in question for achieving a satisfactory quality of teaching and learning. A great deal may have happened in the teaching performance during this period – both positive and negative.

There are also weaknesses in the indicators employed in the LTPF – primarily in education results and final outcome, but also to a lesser extent in the education and learning process. Questions may also be raised about whether the indicators applied really help to measure a good teaching environment. Continued studies – and drop-outs – are often due to several different factors, of which teaching quality is only one. Full-time employment and continued studies at second cycle level also depend to some extent on factors beyond the university’s control, for example the labour market situation. An indicator that involves a student’s approval for completion of a particular course may also be questioned, since approval also depends on requirements for the course in question.

A further weakness is that there may be a negative impact on the quality of the data on which the allocation of funds is based. Some differences in the way the various institutions apply the survey may be noted after the government started to use the AGS results to determine funding. Some institutions, for example, have focused on improving their evaluation unit and increasing student awareness of their learning. Andrys Onsman considers that certain universities primarily focus on improving their survey results, rather than on achieving better results by improving the learning process.125

Another side of the coin is that when the AGS began to be used for funding allocations, the methods and the indicators were also questioned to a greater extent – for example do the indicators really cover the quality of teaching and the forms for student support and their learning incentives? The effects of the adjustment factors were also criticised.126 The results of the student surveys are clearly more crucial when coupled with allocation of resources and, consequently, the way the results are arrived at is also more sensitive.

Great Britain – newspaper rankings for student choices

As in the United States, the ranking of universities and other higher education institutions has a relatively extensive and comprehensive background. The most well-known British international ranking is published by the Times Education Supplement, which has already been described in this chapter. But there are also several other national rankings of British universities and higher education institutions. Newspapers are the primary participants in this field in the UK, and three media-based British rankings conducted by the Times, the Sunday Times and the Guardian are briefly presented in the following.127

The Higher Education Funding Council for England (HEFCE) initiative for the provision of various types of information about higher education in the UK is also of interest. This information was previously provided via a “Teaching Quality Information” website which is currently referred to as “Unistats”. This operation is not classified as ranking, but it has certain similarities with the German Centrum für Hochschulentwicklung (CHE), which is referred to subsequently. As a result, we will also describe Unisatats in more detail at a later stage in this section.

The Times’ “Good University Guide”

Background and purpose

The Times’ Good University Guide was published for the first time in 1992, and the most recent version is from 2008. This is a consumer product targeted at readers of this newspaper, in particular the parents of potential higher education students. According to the Times, this ranking aims to present the best universities in a rather traditional sense.

Ranking characteristics – focus, indicators and weighting

The Times’ Good University Guide ranks both entire higher education institutions and different subjects. The institution ranking currently employs eight indicators, but has varied over time, with more than 14 indicators in the first version. Only three of these indicators are used in the subject rankings, of which there are slightly more than 70.

These indicators are based on two main sources: the Higher Education Statistics Agency (HESA) and the Higher Education Funding Council for England (HEFCE). HESA provides some statistical information, while HEFCE covers both the National Student Survey (NSS) and the Research Assessment Exercise (RAE), which evaluates research.

The following eight indicators were included in the Times’ Good University Guide for 2008:

1. **Student satisfaction**: average satisfaction according to the National Student Survey totals for 2006 and 2007, or one of these years if there is no data for both years.
2. **Research quality**: the overall quality of research based on the 2001 Research Assessment Exercise, which combines quality and volume measures.
3. **Entry standards**: average Universities & Colleges Admissions Service (UCAS) scores for new higher education students under the age of 21 in 2006/07.

---

128. The following section is based on www.timesonline.co.uk/tol/life_and_style/education/good_university_guide/, HEFCE Issues paper 2008/14 and Yorke & Longden (year unknown).

129. The Times’ subject rankings only include three indicators: research quality, entry standards and prospects for graduates.
4. **Student/teacher ratio**: average number of students per academic employee in 2006/07.

5. **Disbursements for services and facilities**: expenses per student for resources for library services, IT, career counselling, sport, health and advisory services, estimated for 2004/05 and 2005/06.

6. **Throughflow ("completion")**: The proportion of students who commenced higher education in 2004/05 and expected to take a degree.

7. **Highest rating ("good honours")**: the proportion of students who achieved a “first or upper second class degree” in 2006/07.

8. **Prospects for graduates**: the proportion of graduates in “graduate level employment” or in continued studies six months after taking their degree in 2006.

In accordance with the previous classification of indicators, it may be noted that the Times’ Good University Guide employs one indicator for entry standards, two or three for resources (student/teacher ratio, disbursements for services and possibly research quality), two for results (throughflow, highest rating) and one for final outcome (prospects for graduates). Student satisfaction is normally considered to be in the process indicator category (see the section on learning impact methods in Chapter 1), but this is a somewhat indirect measure – perhaps it should be located in the results category.

Research quality may clearly also be regarded as a result, but not from the primary education perspective applied by the Times in this context. The research category is obviously relevant if research is to be regarded as a separate category, but we have preferred to treat it as a resource.

The Times does not employ any value-added indicators in the form of results/input values. Input values are regarded in the Times’ ranking as a quality measure in its own right, and this indicator is used to measure results in the light of entry standards. The absence of the reputation indicators commonly employed in rankings is perhaps more surprising.

The eight indicators have different weightings in the Times’ ranking. Student satisfaction and research quality each have a weighting of 1.5, while the other six indicators have a weighting of 1.0, on the grounds that the first two are indicators of the university’s primary functions.

**Strengths and weaknesses**

The Times Good University Guide’s ranking is a relatively simple and transparent listing of British higher education institutions. The indicators employed are reported openly, and this also applies to the sources of information. The results of both the ranking of institutions and subjects is readily accessible for all interested parties. On the other hand, it is harder to understand the calculations in the form of standardisations and weighting.

The results of the ‘Times’ ranking are scarcely surprising, assuming that the aim is to show which are the best universities in a traditional sense. There is
some degree of consensus in the results with other British rankings, and we will return to this at a later stage. Naturally, this largely reflects the fact that the indicators applied are precisely the traditional indicators of the quality of a university. This is problematical, however, since although student satisfaction (a less traditional indicator) has a higher weighting than many of the other indicators, statistical analyses indicate that the indicators that best explain the overall ranking results are research quality, entry standards and degree and credit results. In other words, the results are more traditional than intended, probably due to several unmeasured factors that are not subject to control.\textsuperscript{130}

The Times has developed a relatively sophisticated classification of subjects and, of course, the large number of subject rankings is of considerable interest from a student perspective. However, the limited number of indicators measured and ranked at the subject level (research quality, entry standards and prospects for graduates) reduces their relevance significantly.

Much of the criticism directed at rankings in general and, in particular, at ranking results in the form of weighted “league tables” may, of course, also be directed at the Times’ ranking. The choice of indicators is hardly logical – it is governed more by what is available that by what really counts. In addition, every indicator has specific validity and reliability problems (see below), and the weighting and combination of the indicators is arbitrary to some extent.

The absence of indicators for reputation might normally be regarded as an advantage, but the absence of indicators measuring the actual education process (form, contents, quality) and value-added aspects or more specific learning outcomes is a weakness, if the aim is to give potential students a full picture.

**The Sunday Times “University Guide”\textsuperscript{131}**

**Background and purpose**

The Sunday Times has been publishing its University Guide since 1998, and the most recent issue appeared in 2008. The Sunday Times Guide aims to be the first point of reference for students and their parents and, according to this newspaper, the various rankings are designed to show the considerable diversity in the higher education sector and the variety of experiences offered by the institutions concerned. In addition, it is assumed that this ranking reflects the international status of these institutions.

**Ranking characteristics – focus, indicators and weighting**

The Sunday Times University Guide only contains ranking of institutions, and has no rankings at the subject or education programme level. On the other hand, there are several different rankings to choose from – rankings that are based on some specific indicator. The main ranking weighs up information from nine different indicators, based on data obtained from many different

\textsuperscript{130} HEFCE Issues paper 2008/14.

\textsuperscript{131} This section is based on www.timesonline.co.uk/tol/life_and_style/education/sunday_times_university_guide/, HEFCE Issues paper 208/14 and Yorke & Longden (year unknown).
sources: HESA, HEFCE, the Sunday Times’ own questionnaires, the Quality Assurance Agency (QAA), the Scottish Funding Council (SFC) and the Higher Education Founding Council for Wales (HEFCW).

The nine indicators included in the Sunday Times University Guide in its 2008 ranking are:
1. **Student satisfaction**: weighted results from the National Student Survey in 2008.
2. **Teaching excellence**: the proportion of the total number of subject evaluations which have been considered to be excellent in the rolling evaluation since 1995.
3. **Vice-chancellor and expert assessments**: the weighted number of mentions of high-quality first cycle programmes by vice-chancellors and, at the subject level, by academics.
4. **Research quality**: the quality and quantity of research assessed according to the most recent Research Assessment Exercise in 2001.
5. **Entry standard (“A-level/Higher points”)**: the proportion of high Universities & Colleges Admissions Service (UCAS) scores for new students.
6. **Unemployment**: the proportion of students assumed to be unemployed six months after taking their degree.
7. **Highest rating (“firsts/2:1s awarded”)**: the proportion of students receiving the best or next best degree qualifications.
8. **Student/teacher ratio**: the number of students per teacher.
9. **Drop-outs**: The number of students who terminate their studies before completing a course in relation to the number of students expected to drop out.

Hence, the Sunday Times University Guide employs one indicator for entrant characteristics (entry standard), two indicators for resources (research quality and student/teacher ratio), two indicators for results (highest rating and drop-outs), two indicators for the education process (student satisfaction and teaching excellence), one indicator for final outcome (unemployment) and one indicator for reputation (vice-chancellor and expert assessments). Placing the expert assessments in the reputation category may appear to be unfair, but in this case this does not involve in-depth assessments, but merely simple questions to vice-chancellors and academics in the various disciplines about where the best education is located. Very possibly, these persons may know more about the real situation in first cycle education than most people, but it is by no means certain that such subjective opinions are always well-founded.

One of the indicators – teaching excellence – is hard to place since what is actually being measured is not clear. This indicator is based on the quality assessments that have been carried out over the years, and they might cover both resources and results, but more likely they primarily reflect the process between resources and results.
The indicators are calculated and weighted in rather complicated procedures – points are awarded for each indicator in accordance with the position achieved in relation to the average. Additional weight is given to two indicators that are considered to be particularly important, namely teaching quality (excellence combined with student satisfaction) and the students’ entry standard. These two indicators are weighted by a factor of 250 each out of a total of 1,100, and hence they represent 45 per cent to the total weighting. Research quality also has a heavy weighting (200).

In addition to the main list, the Sunday Times’ website also publishes alternative ranking lists, both for the individual indicators above, and for other indicators not included in the main list (e.g. application pressure and scholarships).

Strengths and weaknesses

The Sunday Times University Guide includes several rather different indicators in its ranking. This makes it multi-faceted, of course, in the sense that the Guide tries to capture several different dimensions of quality. The sources on which this ranking is based are more varied than in the other British rankings we have studied, although the indicators are relatively hard to understand after the various weighting procedures, and it is difficult to understand what the final weighted list actually measures.

As in the case of the Times Good University Guide, there are several problems with the various indicators, and the Sunday Times’ weighted list also tends to be more traditional than was intended, since it tries to weight the teaching quality. Statistical analyses indicate that dramatic changes in positions are almost always due to entry standards and drop-outs – two rather unstable indicators, one of which has a heavy weighting.

Nonetheless, perhaps the major criticism of the Sunday Times ranking is that it does not fulfil its purpose – to inform students and their parents about quality in higher education. As far as students are concerned, the ranking by subjects is probably of greater interest than the ranking of institutions, and the Sunday Times offers no information at the subject level.

The Guardian “University Guide”

Background and purpose

The Guardian University Guide was first published in 1999. The latest edition (“University Guide 2009”) contains information targeted at potential students for the 2009/10 academic year, and the Guardian ranking exclusively focuses on information to future students prior to their selection of higher education options. As a result, the emphasis is on rankings at the subject level, although there is also a general ranking list for higher education institutions. But the

132. The following section is based on http://education.guardian.co.uk/university

guide2009/0,,2276673,00.html; HEPCE Issue paper 2008/14 and Yorke & Longden (year

unknown).
Guardian University Guide is primarily interested in the quality of education, rather than research outcomes or prestige factors. The ranking of institutions is explained on the grounds that there may be students who have not yet decided what subject they wish to study.

Ranking characteristics – focus, indicators and weighting

The Guardian University Guide’s rankings are based on seven different indicators, based on data from HESA and HEFCE (“National Student Survey”). Higher education institutions are ranked for about 45 different subjects – one ranking for each subject offered by the institution concerned. The comprehensive ranking of institutions is the average of results in the subject-specific rankings.

The seven indicators used in the Guardian University Guide are:

1. **Teaching**: teaching quality according to the National Student Survey (average satisfaction in four questions).
2. **Assessment and feedback**: the quality of assessment and feedback according to the National Student survey (average opinion in five questions).
3. **Added-value results**: students’ entry qualifications in comparison with final grades. Results are based on the probability of the achievement of high student grades in comparison with the actual outcome.
4. **Student/teacher ratio**: the number of teachers for a given subject in comparison with the number of students in this subject.
5. **Expenditure per student**: the amount of money expended per student in the subject concerned, excluding expenses for academic staff.
6. **Entry qualifications**: the average UCAS scores for new students for the subject concerned.
7. **Career prospects**: the proportion of graduates employed in a job with “the right” qualification requirements within six months of taking a degree.

Hence, The Guardian University Guide uses one indicator for entrant characteristics (entry qualifications), two indicators for resources (student/teacher ratio and expenditure per student), two indicators for process between resources and results – even if measured indirectly (teaching and assessment and feedback), and also – rather unusually – an indicator for added-value (added-value results). It is interesting to note that the Guardian’s ranking contains no indicators for results, apart from the added-value heading, and that there are no indicators for reputation or research quality.

Both the institution ranking and the separate subject rankings are based on these seven indicators. Higher education institutions with data for at least five of these indicators and subjects, and with at least 35 full-time students, are included in the Guardian tables.

The two indicators based on the National Student Survey have a lower weighting than the others – teaching has a 10 per cent weighting and feedback
5 per cent. The remaining five indicators have a 17 per cent weighting, although this does not apply for medical, dental surgery and veterinary programmes, where there is no added-value or career indicator. In this case, teaching has a 15 per cent weighting, assessment and feedback 10 per cent, and the remaining three indicators 25 per cent.

The ranking for higher education institutions is based on the subject rankings, although the number of students at the various institutions and the number of institutions with teaching in various subjects are taken into account.

Strengths and weaknesses

Perhaps the Guardian University Guide is the ranking model that most clearly achieves the objective of providing information for potential students, in comparison with the other British ranking systems described above. This can be seen in the subject-based rankings, and also in the strong focus on indicators that cover teaching – but not research.

The Guardian ranking also differs from the other British rankings that contain an added-value indicator, which is often considered to be one of the strongest measures of education quality. Unfortunately, both entry qualifications and final grades suffer from several measurement problems, with the result that the added-value cannot be defended from a statistical viewpoint in this case. Furthermore, the fact that entry qualifications reappear as an indicator in its own right means that entry qualifications are counted twice and, as a result, carry a very heavy weighting in the Guardian’s ranking.

As is the case for the other rankings, the weighting of the indicators to produce an overall result cannot be fully understood on an intuitive basis. Although the underlying methods are explained in some detail, it is not possible to independently remodel some of these calculations.

The Guardian has deliberately chosen to exclude indicators for research quality from its ranking, but this has been questioned, particularly by representatives of the institutions themselves. Good research is generally considered to be a prerequisite for good education. The Guardian considers, however, that it is more sensible to measure student and teaching factors rather than the research conducted by professors who may not be involved in teaching, and it also considers that if research really does have an impact on teaching, this should show up in the teaching quality indicator, and not merely as a separate measure of research quality.

Evaluation of British ranking systems

In April 2008, the Higher Education Funding Council for England (HEFCE) published a report on ranking and “league tables”, entitled “Counting what is measured or measuring what counts?” This report is based on a thorough review and analysis of the three British newspaper-initiated rankings described above. This section is based on the HEFCE Issues paper 2008/14, particularly Appendix B. Some of the argumentation is also presented in Yorke & Longden (year unknown).
bed above, and two international rankings by THES-QS and the Shanghia Jiao Tong-University (also briefly described in this chapter). The authors of the report have focused on a detailed analysis of the methodological quality of these rankings, involving analysis of the validity and reliability of the various indicators and of the mutual correlations between indicators, correlations between indicators and ranking results, factor analyses of the dimensionality of the rankings, and the correlation between the various rankings. In other words, this has involved an evaluation of the rankings on a technical/statistical basis, taking into account the purpose of each ranking. The conclusions are interesting from several viewpoints.

Many of the conclusions are of such a general nature that they can be applied universally and, as a result, there is reason to return to them in the next chapter. But since we have presented examples of three newspaper-based British rankings in the above, which are also evaluated in the HEFCE report, reference to these methodological problems in the present context is a natural step.

All three rankings – the Times Good University Guide, the Sunday Times University Guide and the Guardian University Guide – primarily focus on resource and result indicators (entrant characteristics are regarded as resources in this context). Few of the indicators used measure the actual process between resources and results which, it might be claimed, actually reflects genuine teaching quality. The aspects of the National Student Survey that cover the way students have assessed teaching are a relatively indirect manner of measuring this process, and general satisfaction might be the result of the outcome of the students’ studies (degree, employment, etc.), rather than a reflection of the actual quality of their teaching. The Sunday Times includes teaching excellence based on quality assessments of various education programmes, but even if this is a much more direct measure, it is hardly transparent. There is also a risk that it becomes out-of-date for many subjects (the assessments have continued since 1995).

Above all, these resource indicators are criticised from a validity angle. According to the authors of the report, resources do not say much about the quality of the teaching about which, it is alleged, students are to be informed. Instead, resources tend to reflect the prerequisites of the various higher education institutions, without measuring how effectively the resources are used. One may even wonder whether there is any correlation between resources, student learning and the results achieved.134 The fact that resource indicators are nonetheless employed to a much greater extent than process indicators is often justified on the methodologically dubious grounds that they are easier to measure.

134. The report refers to a study that has examined research on this matter, which concludes that “inputs such as expenditure per student, student/staff ratios, research quality, admission selectivity or reputation have only an inconsistent or trivial relationship with student learning and achievement.” HEFCE Issues paper 2008/14, p. 18. See also Pascarella & Terenzini (2005).
According to the authors of the report, result and outcome indicators are more valid measures of the quality of a higher education institution. But the problem with these indicators is, instead, their reliability and consistency, since they vary more than most types of indicators and there appears to be little consensus among ranking designers as to which indicators should be applied. In addition, it is difficult to interpret them as direct indicators since, for example, grade results and employment depend on so many different factors. Ideally, they should be checked against resource factors, the size of the institution, its geographical location, its subject profiling and labour market trends. The danger is that result and outcome indicators say more about an institution’s recruitment and reputation than they do about the actual quality of teaching. As a result, they should also be checked for the relevant characteristics of new students (gender, qualifications, social and ethnic background) if they are to constitute valid measures of quality.

As regards the weighting of indicators in the British rankings to form a final list, the report criticises both the weightings, standardisation and normalisation, and also the transparency of the process. Similarly, there is no consensus in these procedures, and different ranking designers clearly mean quite different things when they refer to “normalisation” for example – and yet no one appears to use this term in the sense applied in statistical theory. The weightings are a reflection of the fact that ranking designers regard different indicators as having different degrees of importance, but the exact value of each weighting is rather arbitrarily chosen in most cases.

The methodology applied is often described in relative detail, but it does not cover everything, and in several cases it is not possible to replicate the calculations in the final ranking, notwithstanding access to the input indicators. In other words, there are serious deficiencies in terms of openness and transparency.

Statistical analyses of ranking results in relation to the various indicators often reveal some intractable problems. One problem is that the lists are not always so one-dimensional as the statistical techniques imply and, in point of fact, two dimensions are being measured, and they are sometimes not correlated. Another problem is that there is sometimes an underlying dimension that explains so many of the positions on the list that the weightings applied to many of the indicators have little impact on the final outcome.

An analysis of the relationship between the various rankings indicates that there is a considerable degree of agreement in the positions achieved by the various institutions in the three rankings. The Times’ and the Sunday Times’ rankings are more similar, however, while the Guardian’s ranking differs to a greater extent (which may be partly because the Guardian’s ranking clearly covers two dimensions).

The report concludes that that the indicators employed are “poor operationalisations” of what we really want to know, and this leads to inadequate validity – the rankings do not cover what they are supposed to cover. The methods
employed also present problems and, as a result, the weightings applied do not always have the desired effect and, as a result are only capable of depicting research excellence or reputation.

**Unistats.com**

Student information may take many forms and, the case of Great Britain, there is another interesting alternative to rankings in the provision of information about higher education for potential and other categories of students. This initiative is close to some aspects of the German ranking system presented in the following, although Unistats does not refer to its operations as ranking.

Unistats.com is an interactive website that collects official information for purposes of comparison for academic subjects at various universities and university colleges in Britain. This information is obtained from the Higher Education Statistics Agency (HESA), the Learning and Skills Council (LSC) and the Higher Education Funding Council for England (HEFCE).

The Unists website is totally owned by HEFCE in behalf of the Funding Council, and it also includes the Higher Education Funding Council for Wales (HEFCW), the Department for Employment and Learning in Northern Ireland (DEL), and the Scottish Funding Council (FC). The website is managed, however, by UCAS (the UK Centralised Admissions Service).

Unistats.com involves a number of different indicators at the subject level, and has obtained information from the HEFCE National Student Survey, which is a major annual questionnaire for students in their last year of first cycle studies (final year graduates). This questionnaire asks students how they experienced various aspects of their studies.

HESA provides information from official statistics as regards: student data as regards entry qualifications, their education and results, data concerning people who terminated their higher education, and context data about students’ backgrounds. LSE also presents similar information for students who are directly financed by Further Education Colleges in England (if there is no destination data).

This information is updated on an annual basis, and takes into account the most recently available sources. At the moment, this means that NSS data is taken from the 2008 survey, while other information is primarily for the 2006/07 academic year. The major proportion of this information covers subject areas, rather than specific courses or programmes, although the ambition for the future is to also collect information at the course level.

A search may be conducted for information, both for the subject and the university concerned. Registered users may also save a “shortlist” that can be used on a subsequent occasion. There are also opportunities to download special types of data, both for a specific institution and for the entire sector. This

---

135. The following section is based on information from Unistats.com.
downloading function also includes a detailed description of all the variables, definitions data-collection methods and references to the original sources.

Unistats.com offers information for students about higher education in Britain in a relatively straightforward and flexible manner. It is possible to compare various institutions with each other in the subject areas concerned, even if this does not involve a full-scale ranking procedure (if there is such a thing). This information is based on official data, quality assured as far as possible, and collected by government authorities. Problems concerning the response rate, definitions, and comparability are clearly presented on the website.

Nonetheless, the information has been selected, of course, and for many students this probably means that several aspects are not covered. There are no student welfare indicators, for example, and the information on teaching quality in the education process is entirely based on subjective responses in questionnaires with a relatively low response rate. If fuller information was required, Unistats might be supplemented by more information about education programmes, higher education institutions and university towns. Naturally, the limitations are largely determined by the quality-assured information that is available. In other words, this involves a compromise between the relevance and the reliability of the information concerned – and Unistats focuses on reliability.

**Germany – multidimensional and interactive rankings**

Higher education in Germany is ranked by an independent organisation, “Centrum für Hochschulentwicklung” (CHE), in cooperation with the Die Zeit newspaper. We will be studying this ranking in more detail in the following section.

**The CHE (and Die Zeit) “Hochschulranking”**

**Background and purpose**

The “Centrum für Hochschulentwicklung” (CHE) (the Centre for Higher Education Development) is a private, non-profit-making organisation which was founded in 1994 by the German vice-chancellors’ conference and the Bertelsmann Foundation. Its primary aim when it was founded was to promote reforms in German higher education. Since 1998, the CHE has also conducted rankings of German higher education and, since 2005, this has involved collaboration with the Die Zeit newspaper. In contrast with many other newspaper-based rankings, Die Zeit is purely responsible for the publication and distribution of this ranking, while responsibility for the concepts, indicators and data lies exclusively with the CHE.

---

The CHE has gradually extended its ranking operations, both geographically and in terms of the subject covered. It has been ranking Austrian universities since 2004 and Swiss universities since 2005 (this activity has now been terminated), and in 2006/2007 it was also involved in an experimental project covering universities in the Netherlands. In addition to the ranking of first cycle programmes it also covers research and what are termed “excellent programmes” at the second cycle level.

The primary aim of the CHE ranking of higher education institutions is to provide information for potential students about the quality of, and surrounding, higher education. However, the CHE also aims to supply information and data to institutions for purposes of comparison. This double purpose presents some problems – potential students often require less complex information since they form part of a group with limited experience of the higher education sector, while the institutions concerned tend to require detailed and sophisticated information. The CHE rankings try to establish a balance between reaching their target group and gaining acceptance among the institutions. The latter factor is particularly important since the CHE relies on information from the institutions.

**Ranking characteristics – focus, indicators and weighting**

Die Zeit publishes the CHE higher education ranking on an annual basis in a special student guide (“Studienführer”), in which 30 different subjects are ranked on the basis of five indicators, although they are not the same for all subjects. The indicators include (for example in the case of political science):

1. **General study situation**: average student assessments in a major student survey of the teacher and student situation in the specific programme concerned.
2. **Advice**: average student assessment as regards support, advice and counselling.
3. **Equipment**: assessment by CHE, students and professors of the facilities available in the form of library and data/IT resources.
4. **External research funds**: total amount of external research funding per research staff-member.
5. **Research reputation**: average assessment by professors of research in the discipline concerned.

The CHE and Die Zeit ranking differs from the majority of rankings in a number of crucial aspects. Firstly, this ranking applies to academic subject-disciplines – even in the newspaper version – rather than entire higher education institutions. As a result, the Die Zeit “Studentführer” devotes a couple of pages to rankings within each of the subjects covered. It is hard to make comparisons across subject boundaries, since different indicators are employed for different subjects.
Secondly, the CHE ranking is multidimensional (applying its own terminology). This is intended to indicate that there is no weighting of indicators to arrive at an overall result. Hence, it is possible to read off results for each of the five indicators in the Die Zeit student guide – in other words there are five separate rankings for each subject covered.

Thirdly, the results are presented in a manner that contrasts with most of the other rankings. Die Zeit does not give the institutions any exact value for the indicators for the various subjects. Instead, they are ranked in one of three different groups: the top group (above average), a middle group (average) and a lower group (below average). The results achieved for a particular indicator are denoted by a green, yellow or red dot, and the trend for each indicator in comparison with the previous year is symbolised by an arrow – pointing upwards to indicate better results or downwards to indicate a deterioration.

The CHE also offers opportunities for interactive rankings via two different websites – one in German and the other in English. These websites offer the user a choice of routes in the search process – university, university town/city, or a subject/discipline. These two websites provide information concerning the town/city, universities and their special profiles, and the various subjects offered. There are also several links to the institutions’ own websites and other information pages.

In addition to access to the ranking overview presented in the Die Zeit “Studentführer”, these websites also provide an opportunity to compose personal “My rankings”, in which up to five rankings can be selected on each occasion. In a way, it is also possible to weight the results, since the user can select green, green and yellow, or all results for each of the indicators. In other words, if one indicator is considered to be particularly important, the user may decide to concentrate on institutions with the specific education programme concerned that are above average on this indicator.

There is a broad spectrum of indicators: towns/cities and universities, students, outcomes, internationalisation, teaching, resources, research, labour market and employability, and overall assessments by students and professors. Most of the indicators are covered by the CHE list of indicators. Information is provided about new-student characteristics, resources (both physical and staff), results, final outcomes, reputation and research. In addition, many of the indicators cover the process between resources and results to some extent, by measuring contacts with teachers and other students, the quality of counselling, etc. There are also some student welfare indicators, for example accommodation and opportunities for sports activities. There are no value-added indicators, however.

These indicators are based on several different types of sources and data. Some of the indicators are based on various types of statistical information, while others are obtained directly from the institutions concerned. There are indicators that are based on bibliometric analyses, although a high proportion are based on various forms of questionnaire surveys involving students.
and professors. Each indicator is marked with a letter that denotes its source – “F” for factual information, “S” for student opinions, and “P” for professional views.

**Strengths and weaknesses**

Seen in the light of the aim of giving potential students information about higher education in Germany, the CHE ranking is relatively comprehensive. It contains a great deal of information about higher education in different subject, which is presumably of considerable interest for students. The CHE covers a broad spectrum of indicators, particularly via the interactive websites, and the possibilities of developing a personal ranking (or information lists) are an attractive feature. The CHE also provides information about aspects that are less common in rankings but which have considerable potential interest for students, for example student welfare factors and information about the actual teaching situation.

Naturally, the CHE ranking avoids many of the methodological problems of the weighting of indicators that tends to hide information, rather than make it available to potential students. The grouping of results in three broad categories – above, at, or below average – also means that the differences between higher education institutions are not exaggerated (except between the groups).

But what makes the CHE ranking most attractive from the methodological viewpoint also makes it less viable in a media context, and less direct. The Die Zeit “Studentführer” is a small book which is purchased separately, and it would be hard to imagine the results of the CHE ranking in the headlines on the first page of a newspaper. Quite simply, there is no simple list running from best to worst. As a result, this ranking does not provide any simple answer as to which is the best university for studies in Germany. This ranking calls for a greater degree of commitment on the part of potential students than most other rankings. However, assuming that the potential student is interested in investigating how higher education functions in Germany in more detail, and sufficiently involved in this question, the websites provide a great deal of information, and what the information really measures is relatively transparent.

Nonetheless, the CHE ranking also suffers from a number of methodological problems. One, of course, is that the classification of subjects is not entirely self-evident (this applies to all rankings of subjects/disciplines or programmes), and a considerable proportion of higher education in Germany is probably excluded.

Furthermore, it is difficult to obtain official statistics for several aspects which may be of potential interest, partly due to the disaggregated level of such information, but also due to Germany’s federal nature in which the various constituent states have different routines. As a result, the CHE relies on good relationships with the institutions in order to obtain some of what is termed “factual” information. Information of this kind is not quality assured to the same extent as official statistics.
But the greatest problem with the CHE ranking is that so much of the information (particularly information that is perhaps of greatest interest to potential students) is based on questionnaire surveys addressed to students or professors. Questionnaires generate subjective opinions, which are not necessarily uninteresting, per se, but which call for an awareness of what they are actually saying. The way students evaluate their study situation is not merely a question of what the actual study situation is like. It may also involve the level of ambition, personal experiences and the way opinions are expressed in a questionnaire. The way professors assess research or programmes in their area may, of course, be based on their professional evaluation of whether the research or programmes concerned are of high quality, but personal networks and general opinions about other institutions probably also affect their opinions.

However, the greatest problem with data derived from questionnaires is that the reliability of the results is closely linked to the number of respondents and who actually responded. Naturally, this is also connected with the sample that was deliberately selected, although this is a minor problem in this case. The 2008 CHE ranking is based on replies from 92,378 students, but this is only 21.3 per cent of the total. But the number of respondents in absolute terms is of less importance than the persons who replied (dissatisfied, satisfied or other special groups), and this is precisely what the sponsors of the questionnaire cannot control. As a result, all indicators based on questionnaires (a rather high proportion) must be interpreted with considerable caution. And hence the CHE ranking may be said to suffer from quite a different problem compared with the Unistat website: the information has a high degree of relevance, but low reliability.

**Sweden – limited range of university and higher education rankings**

In comparison with the Anglo-Saxon countries, the history of rankings in Sweden is relatively brief. In the late 1990s and early years of the present century, the Moderna Tider magazine (now discontinued) published rankings of Swedish higher education institutions for various subjects. This started with rankings largely based on various forms of statistical information such as "application pressure", the number of teachers who had defended a doctoral thesis, the number of professors, and throughflow, but also on status and reputation information from questionnaires completed by academics. These rankings were developed, however, into an attempt to depict the quality of the education provided and certain student welfare aspects, via questionnaires address-
sed to students. As a result, this ranking was confined to a limited number of subjects per year.\textsuperscript{138}

The closure of Moderna Tider did not lead to the total disappearance of the ranking of universities and other higher education institutions in Sweden, however. In recent years, rankings – for different purposes and for different target groups – have become increasingly relevant, even if this is often a matter of the Swedish positions in international listings.\textsuperscript{139} We will be looking at four relatively current rankings of Swedish universities and higher education institutions in this section.

The first of these rankings, issued by Sydsvenska Industri- och Handelskammaren, [the Southern Swedish Chamber of Industry and Commerce] is a rather traditional ranking which closely resembles the international and Anglo-Saxon rankings that have already been described.

The next two rankings are published by Fokus, but designed by the Urank association. From the point of view of the three designers, these rankings should be primarily regarded as methodological experiments, and a way of stimulating discussion of these issues, and as a result, they should not be considered to be ranking operations of the normal commercial type. On the other hand, publication in Fokus probably means that they are widely regarded as typical traditional rankings. As a result, there are grounds for looking at the Fokus/Urank rankings in more detail in this context. They involve two different rankings, one of which is a rather traditional “academic” ranking, while the other is an alternative “broad recruitment” ranking.

The final ranking which we discuss is the Sensate Näringsliv coordination ranking, which we will only describe briefly since it has a much narrower focus than the others, although its interactive Internet application, which is reminiscent of the German CHE ranking, is of some interest.

\textbf{Sydsvenska Industri- och Handelskammare ranking [Southern Swedish Chamber of Industry and Commerce]} \textsuperscript{140}

\textbf{Background and purpose}

The Sydsvenska Industri- och Handelskammare has published rankings of Swedish higher education institutions since 2006. The primary function is said to be to provide a basis for a discussion of quality at higher education institutions and universities, and the prerequisites for higher education and research. But this ranking is also conducted in the light of the fact that it has been noted that it is essential for companies to be able to recruit qualified and


\textsuperscript{140} The following section is based on Högskolerankningen 2008, Handelskammarens rapport No. 5, 2008.
well-educated personnel and that they have access to research. In this sense, the target group for this ranking is primarily industry and commerce.

**Ranking characteristics – focus, indicators and weighting**

The Sydsvenska Industri- och Handelskammare ranks higher education institutions, and not subjects/disciplines or education programmes. This ranking weighs up information in eight different areas, based on eight different indicators, most of which are derived from official statistics published by the National Agency for Higher Education (the NU statistics database), although bibliometric compilations from the Danish Technical University “Promotion Centre” are also employed, which in their turn are based on data from the “Thomson Web of Science”.

The first five of the eight areas identified as of interest corresponds to the private sector’s need to recruit and train qualified personnel, while the three subsequent indicators are based on the private sector’s need to obtain access to research at higher education institutions. The Sydsvenska Industri- och Handelskammare bases its ranking on the following eight areas and their associated indicators:

1. **Efficiency of first cycle education**: weighted value for the number of graduates per number of FTE students and the average credits after three years,
2. **Student attractiveness**: the number of first-hand applicants for programmes, trend over the past two years.
3. **Teaching resources**: number of FTE (full-time equivalent) students per teacher.
4. **Teacher qualifications**: number of teachers who have defended a doctoral thesis.
5. **International exchanges**: total number of students from other countries or studying on other countries in relation to the number of FTE students.
6. **Attractive research for companies**: proportion of external research.
7. **Efficiency of research programmes**: degree throughflow.
8. **Research productivity**: number of scientific articles published in relation to research revenues.

The Sydsvenska Industri- och Handelskammare’s report examines the arguments in favour of each of these indicators – or parameters as they are termed in this context. This discussion also provides guidance as to the most appropriate categories in which the various indicators may be placed. Translating these Swedish indicators into an internationally inspired (Anglo-Saxon) typology is by no means a simple matter, however. Some traditional entrant characteristics are not included in the above ranking, but perhaps student attractiveness may nonetheless be regarded as an indirect measure of the type of students the institution concerned has managed to recruit: “More students also mean hig-
her allocations and, at the same time, the level of the students admitted rises, and this makes the institution more competitive.” It is also possible to regard the application pressures as an indicator of reputation: “A higher education institution that is appreciated by students will be able to attract more students.”

The international exchanges parameter also depends to some extent on the type of students that the institution has managed to attract – the extent to which students from other countries are attracted (and students who have a propensity to study abroad). The arguments for including this parameter should, however, be regarded more as an indicator of the quality of education (“Studies in other countries are enriching…”), and in other words this might put this indicator in the process category – between resources and results.

It is easier to classify the other indicators. Resources are measured in terms of teaching resources and teacher qualifications, and results are measured in terms of the efficiency of first cycle and third cycle education. Assuming that Sydvenska Industri- och Handelskammare’s aims for its ranking are to measure the extent to which the private sector can have access to research at the institution concerned, perhaps research that is attractive from a corporate and research productivity viewpoint may also be regarded in terms of results, although they might also be placed in a separate research category.

An equal weighting is applied to the indicators in the Sydvenska Industri- och Handelskammare ranking, although a relative value is computed for each institution. A value of 100 for the entire country for each indicator has been allotted, and the value of the various indicators have been linked to this for each institution. Subsequently, the value for the eight areas has been combined to produce a total average figure for each institution. Since the average national index value is 100, institutions with a score of more than 100 are better than the national average, while the reverse applies for institutions with a score of less than 100.

An appendix to the main report also contains diagrams for all higher education institutions, indicating each institution’s relative value for each of the indicators from 2003–2007.

Strengths and weaknesses
In many ways, the Sydvenska Industri- och Handelskammare ranking is rather traditional, in the sense that it ranks entire institutions, based on indicators that are also commonly employed in international rankings. The Sydvenska Industri- och Handelskammare ranking report is admirable as regards the clear way in which it describes the methods employed – the indicators used, the weighting applied, the arguments on which the ranking model is based – and to some extent also the problems encountered in this approach.

The aim is primarily to meet the private sector’s need for information about higher education and research at Swedish institutions, rather than to inform students about quality differences. The report also discusses the way each indicator is assumed to measure some relevant aspect in this context. But, nonethe-
less, the choice of indicators is not completely obvious. One problem discussed by the authors of this report is that several indicators are ambiguous. What is actually conveyed by the indicator for the proportion of graduates and average grades after three years – efficiency or the level of requirements? And is understanding of and information about cultural differences really covered by the indicator for the proportion of students in exchange programmes from other countries, or who study in other countries (if “free-movers” are excluded)?

This type of discussion can be applied to all the indicators, and Sydsvenska Industri- och Handelskammare is well aware of this. As stated in the report: “For this reason, the indicators and the evaluation should be regarded in a single context, both together with other indicators and with other information about the institution concerned”. One of the problems, however, with ranking lists of the combined weighting type involved in this case is that the results are seldom considered in a single context in combination with other information about the institution in question.

Combined weighting of several ambiguous indicators does not make the results more reliable, but instead less transparent. Based on the overall list, it is hard to understand the part played by the various indicators in assessment of the quality of the various institutions. The diagrams describing the indicators separately at each institution are somewhat more helpful, but information of this nature seldom reaches a wider audience.

The ambiguity and validity problems for the specific indicators still apply, however, even if combined weightings are not employed, and they do not disappear, irrespective of the extent to which the indicators are used. According to Sydsvenska Industri- och Handelskammare, “The indicators and the model have been quality assured, since we have repeated measurements for six years”. But this may be seriously questioned. The fact that indicators and models have been used for several years and have produced similar results may possibly indicate that the measurements are consistent (i.e. they measure the same things every year), but it says nothing about the quality of what is supposed to be measured – in other words validity.¹⁴¹

As is the case for most of the other rankings, the choice of indicators may be questioned, not only on the grounds of the quality and relevance of the indicators actually selected, but also on the basis of the absence of relevant indicators. The report states that there were no indicators for “company-oriented statistics”, for example the number of patents granted or the number of spun-off companies. The Sydsvenska Industri- och Handelskammare ranking, in

---

¹⁴¹. Perhaps it is considered in this context that measurement has been changed and improved during these six years, which might increase the validity, but on the other hand this makes comparisons over time more difficult.
common with most other rankings, bases its results on what is available, rather than on the indicators that are really relevant.\textsuperscript{142}

The “one-size-fits-all” phenomenon, which is also a common factor in the Sydsvenska Industri- och Handelskammare and many other rankings, constitutes a further methodological problem. The question is whether comparisons between higher education institutions are really fair, in the light of their differences. The Sydsvenska Industri- och Handelskammare’s argument, which is also the view of many other ranking designers, is that the model treats all institutions on an equal basis. On a strictly formal basis, all institutions have the same prerequisites, but in practice, of course, their circumstances differ considerably. This applies, in particular, to the way students and research activities are allocated between different faculties, since this determines the outcome of the indicators. And the Sydsvenska Industri- och Handelskammare report notes that the larger institutions are at a disadvantage because they are treated as a single unit, and that measurement at the faculty level would be a better option, but that comparable data is unobtainable at this level. This is yet another example of the determination of what is to be measured in terms of access to data, rather than the purpose of the ranking concerned.

A further problem is that causal conclusions are drawn from an inadequate ranking procedure. The main conclusion appears to be that specialisation is a key factor for a successful institution, since the five institutions with the highest ranking have subject specialisations in medicine, agriculture, economic or technology. If it had been possible to check up on this subject (or possible faculty), it is by no means clear that specialisation is the reason for such success. Would an institution that specialised in languages actually be placed in the top five in such a ranking?

\textit{Fokus/Uranks rankings}\textsuperscript{143}

\textbf{Background and purpose}

Fokus has published a ranking of Swedish universities and higher education institutions for two years in a row, designed by the Urank association. The first ranking appeared in 2007, and was recently followed up in October 2008. Both rankings were published in the Fokus magazine, but are also available (together with the relevant articles) on the URank website.

The underlying purpose of Urank’s rankings was to test methods and initiate discussion. From this viewpoint, these rankings do not comply very well with the types of rankings described hitherto (and subsequently). One crucial reason for nonetheless including the Urank rankings is that they were publis-

\textsuperscript{142} Or – to put it more stylishly – “counting what is measured rather than measuring what counts”, as the tile of HEFCE Issues paper 2008/14 puts it. (“Counting what is measured or measuring what counts? League tables and their impact on higher education institutions in England”).

\textsuperscript{143} The following section is based on information from the www.urank.se website, particularly Forneng, Lind & Nyblom \textit{En svensk universitetsranking – 2008} and Fokus No. 40, 2008.
hed in Fokus, and hence have the same impact as other rankings covered in this chapter.

Urank’s designers have conducted studies of other national and international rankings in parallel with their own rankings and, as a result have noted a growing interest in rankings in Sweden. According to the website, Urank also wants to “disseminate information about the limitations of rankings and the manner in which they should be studied and interpreted”. According to a comment on the discussion page of the “Universitetsläraren” periodical, Urank’s founders also say that their aim has “primarily been to test the way in which Swedish higher education institutions “react” to a set of the criteria that appear to predominate in an international context and the possible relevance and validity of an operation of this nature in assessment of Swedish higher education”.

In this respect, the Fokus/Urank rankings differ substantially from those reported previously in this chapter. The purpose of the Urank rankings is not to inform a specific target group about the actual quality of Swedish higher education. In the articles on the Urank website covering ranking in general and Urank’s ranking in particular, Urank’s founders also adopt a more considerably critical attitude. Nonetheless, these rankings are published in Fokus and have attracted the attention of several parties – as if they were rankings of the quality of Swedish Higher education.

Ranking characteristics – focus, indicators and weighting

In other words, the Fokus/Urank rankings are designed to test the way in which Swedish universities and higher education institutions react to a set of quality criteria on international lines. Fokus/Urank also produces an alternative ranking, however, that covers the width of student recruitment. In both cases, the entire institution is primarily ranked with a focus on the student education function, and not on research activities.

Traditional rankings involve six criteria, as Urank chooses to term them. Each of these criteria covers a number of indicators (the figure in parenthesis after each indicator denotes the indicator’s weighting within each criterion):

1. **Students**: the number of “first-hand” applicants per student admitted (30 per cent), the proportion with higher education aptitude test results in excess of 1.1 for all test results (10 per cent), the proportion of new students not resident in the same county as the institution (10 per cent), the number of students taking a degree who have studied in another country for at least one semester (10 per cent), the number of students still studying at the institutions in the second year (20 per cent), the proportion of new students who have taken a degree (or achieved 180 higher education credits) within six years (20 per cent).

---

144. See also Universitetsläraren 16/2008.

145. A similar ranking experiment has also been conducted by the University College of Halmstad, see Enmark & Lundkvist (2008).
2. **Teachers**: the proportion of teachers who have defended a doctoral thesis (50 per cent), the ratio of professors to teachers (50 per cent).

3. **First cycle education**: the number of students per teacher, weighted for the correlation between research and education (25 per cent), the level of achievement, weighted for the mix of education areas (25 per cent), the mobility factor (25 per cent), the degree of establishment in the labour market, weighted for the mix of education areas (25 per cent).

4. **Research and third cycle education**: research grants per teacher defending a doctoral thesis (20 per cent), the ratio of research to the institution's total costs (20 per cent), the proportion of research funding subject to competition in comparison with total research funding (20 per cent), the proportion of students who transfer to third cycle studies, weighted (20 per cent) and the number of defences of doctoral theses per professor (20 per cent).

5. **Libraries**: the library’s allocation from the parent organisation as a proportion of the institution’s total assets (50 per cent), acquisitions divided by the number of students (50 per cent).

6. **Student assessments**: student ranking according to Studentspegeln [Student Mirror] (50 per cent), student satisfaction with the institution according to Studentspegeln (50 per cent).

The first four criteria have a 20 per cent weighting, while libraries and student assessments each have a 10 per cent weighting. The average for each of the indicators has been computed, and the deviation from the average is noted for each institution, with the distribution’s standard deviation as a unit. Hence, a value of 30 per cent for an institution indicates that the institution is better than the average by a factor of 0.3 multiplied by the distribution’s standard deviation. A weighted average for each criterion is subsequently computed per institution, based on the values of the indicators and the above coefficients. The corresponding procedure has been applied when the weightings for the criteria are combined to produce the final ranking result.

Only three criteria or indicators have been employed in the alternative ranking describing the width of student recruitment:

1. **The proportion of first-generation students in higher education** (33 per cent)
2. **The proportion of students with a non-Swedish background** (33 per cent)
3. **The proportion of students in courses with an even gender distribution** (33 per cent)

The first two indicators involve indicators for the inflow of new students and, in addition, they are standardised in the sense that they take into account the distribution of characteristics concerning the municipalities and age-groups from which students are recruited. The proportion of courses with an even
gender distribution is defined as courses in which 40–60 per cent of the students are either men or women.

In other words, a broad spectrum of indicators is employed, particularly as regards both these rankings. In a traditional, academic ranking, there are several groups of indicators that are also common in international rankings. Entrain characteristics are primarily measured in an academic ranking under the student criterion for higher education test results and new students who come from another county, and possibly indirectly as a result of the number of applicants per student admitted. Resources are measured by indicators based on teacher, first cycle and library criteria: the proportion of teachers who defend a doctoral thesis, the proportion of professors, the student/teacher ratio and the library indicators. Results are measured in terms of the proportion of new students who take a degree or similar qualification, achievements, and possibly the proportion who have studied in other countries and mobility factors. The final outcome is measured in terms of establishment in the labour market and the proportion of students that continues to third cycle studies. The research criterion contains indicators that might be classed as resources, but which could also be regarded as a completely separate category.

Few indicators describe on the process between resources and results, although the proportion of students still studying at the institution in their second year and Studentspegeln might be included under this heading. According to Urank, some consideration was given to the possibility of using measures from the Agency of Higher Education’s quality evaluations of various subjects and programmes, but it was decided that they could not produce sufficient valid data in their present format. There are no value-added indicators in the Urank ranking (although it would be possible to establish such an indicator, especially if indicators from the two other rankings were combined. There are two indicators for new-student characteristics in the alternative width of recruitment ranking: the proportion of first-generation students in higher education, and the proportion of students with a non-Swedish background. The third indicator – the number of students in courses with an even gender distribution – is hard to place, but would perhaps be best classified in the process category between resources and results, although it might also be regarded as a new-student characteristic (an indirect measure of the degree of gender equality in recruitment).

For the most part, all the indicators in both rankings are based on data from the National Agency for Higher Education’s official statistics, with the exception of the library information, which is obtained from Statistics Sweden and student assessments, which are obtained from the Agency for Higher Education’s Studentspegeln survey.

**Strengths and weaknesses**

It is not particularly easy to assess the advantages and disadvantages of the various Fokus/Urank rankings, since the aims are primarily methodological.
In addition, the ranking designers themselves point out many of the major objections.

As in the case of the Sydsvenska Industri- och Handelskammare ranking, one serious criticism is that the Fokus/Urank rankings apply a one-size-fits-all approach. And once again, there is the question of to what extent the various institutions allow themselves to be compared in accordance with the same criteria. This issue is discussed in some detail by the ranking designers themselves, and they consider that the differences in ranking outcomes may be due to – and explained by – the fact that the institutions concerned have different prerequisites (in terms of age, traditions, and the government grant allocation system, according to Urank).

The designers of the Fokus/Urank ranking have considered dividing higher education institutions into various divisions on the same lines as the Canadian Macleans ranking, but have concluded that this is not feasible, since Swedish legislation provides no guidance or grounds for such a step.146

The Fokus/Urank academic ranking involves a fairly broad spectrum of indicators, although the selection may be criticised (as always). The arguments supporting this selection are that the aim has been to formulate some 20 indicators, primarily in numerical terms, which comply with certain academic criteria employed in US/Canadian ranking models. The Macleans’ ranking is cited as the primary model, both in terms of its design, but also due to certain similarities in university systems. Nonetheless, the choice of Macleans as a model is surprising in view of the substantial criticism of this ranking in Canada (see previous references in this chapter).

It has not been possible, however, to fully match all the criteria with the situation in Sweden, partly due to system and cultural differences, and partly due to differences in access to data. In other words, once again, we have a ranking that to some extent selects indicators on the basis of the data available and, if the Fokus/Urank ranking had an explicit target group, one might expect considerably better arguments for the choice of criteria than a claim that “they constitute a reasonable and relevant sample”.

The designers of the Fokus/Urank ranking present fewer arguments to support their weightings – they have simply been inspired by existing models. Some tests have been carried out in the form of variations in the weighting system, but the reasons for different weightings for the various indicators are, at best, implicit.

As regards the alternative ranking for width of recruitment, the aims are more than experimental, of course. In this case, in the total absence of academic criteria, the ranking process is based on more or less political or societal objectives, according to the designers, and it is scarcely surprising that the

146. Divisionalisation has also been attempted on the basis of the degree of specialisation, but this has proved to be difficult since “more than 70% of all Swedish universities and higher education institutions comply with the picture of an average university in every respect, with more than 40% in the humanities and social sciences, slightly less than 30% for natural sciences and technology, and a similar proportion for other disciplines.”
results of these two rankings have almost nothing in common. This is, per se, a good example of what academic rankings are not able to achieve.

Urank considers that there is clearly scope for discussing whether the width criteria are linked to any great extent with quality in the current sense. But perhaps potential students and other interested parties are not purely concerned with the current definitions of quality.

**The Svenskt Näringsliv [Confederation of Swedish Enterprise] cooperation ranking**

The Svenskt Näringsliv “Akademi eller verklighet?” [Academy or reality?] report (2008) presents rankings of higher education institutions based on cooperation with the private sector in 12 different subject areas. In other words, this is a ranking of fairly limited scope, but it might perhaps be extended. This year’s version is an extension of a study of six different education programmes in the previous year.

According to Svenskt Näringsliv, this ranking is carried out on the basis that cooperation is a neglected quality measure, often with no tools for evaluation and concrete development objectives.

Cooperation activities are operationalised in the form of training, mentor activities, projects and degree projects linked to working life, and similar forms of cooperation. The volume, in terms of the number of weeks for each of the cooperation activities is multiplied by the proportion of students participating in the activity concerned. Hence, an absolute value is established for each education area which is ranked in comparison with the absolute value for all other programmes. The indicators are based on questionnaire responses (in some cases telephone interviews) from those responsible for programmes or other senior personnel at institutions offering first or second cycle programmes in the principal subjects concerned. Responses have been received from 170 out of a total of 207 education programmes.

In other words, this ranking has limited scope, both in terms of the programmes covered and also the indicators included. There are several methodological problems, both concerning the reliability and comparability of the questionnaire responses, and the way in which classification of the programmes has been carried out.

This ranking is nonetheless worth mentioning in this context in view of its interactive website (www.hogskolekvalitet.se), in which the user can search for information about both cooperation and establishment in the labour market for 12 different subject areas, and can choose information about programmes at a given institution or within a specific main subject. There is information about the cooperation on which the above ranking is based, and also information about students’ establishment on the labour market obtained by means

---

147. The following section is based on “Akademi eller verklighet?”, Svenskt Näringsliv (2008), and information on www.hogskolekvalitet.se
of telephone interviews with a random sample of students from the various programmes and institutions.

**Figure 2. Summary of the survey of existing ranking systems.**

<table>
<thead>
<tr>
<th></th>
<th>Entran. characteristics</th>
<th>Resources</th>
<th>Results</th>
<th>Process</th>
<th>Final outcome</th>
<th>Reputation</th>
<th>Added-value</th>
<th>Student welfare factors</th>
<th>Institution ranking</th>
<th>Subject ranking</th>
<th>From questionnaires</th>
<th>From third parties</th>
<th>From institutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNITED STATES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. News and World Report</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>?</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Washington Monthly</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Princeton Review</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CANADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macleans</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUSTRALIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian Graduate Survey</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>x</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GREAT BRITAIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Times</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Sunday Times</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Guardian</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GERMANY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHE/Die Zeit</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWEDEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sydsvenska Industri- och Handelkammaren</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fokus/Urank academic ranking</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fokus/Urank alternative ranking</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svenskt Näringsliv cooperation ranking</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comment: This table is an attempt to summarise some of the results in the section above, and it provides a rough summary of our analyses, in which “X” indicates this alternative (type of indicator, focus or data source) applies and “0” indicates that it does not apply. A lower-case “x” is used if classification is particularly problematical, for example if certain indicators could be counted in two different categories. A “?” indicates that we have been unable to find any information under this heading.
# Rankings as student information

## Summary

Students invest both time and money in their education. As a result, it is important for them to have access to comprehensive and relevant information about higher education before they make their choices. Based on our survey of ranking in general, and a few ranking systems in particular, we conclude that most rankings offer neither comprehensive, relevant, nor reliable information about the quality of higher education.

The combined weightings applied to entire higher education institutions presents a particular problem, basically because they contain too little information. They also assume that all the potential target groups have the same highly precise definition of quality – a definition that may be seriously questioned in terms of indicators, methods, reliability and – in particular – relevance. Multi-dimensional rankings, or rather systems for student information, in which students and other interested parties are themselves allowed to define quality, are more attractive. The quality and the selection of indicators are also limitations in this context, however.

There is a great deal of information about higher education in Sweden, and much of it is of good quality. The quality of the information supplied to students depends on whether the information is used for the intended purpose, however, and not merely on access to quality-assured data and the reliability of the information. Very little quality-assured information is suitable for rankings of the type we have studied.

Information that is readily accessible and transparent for students who are interested in comparing different higher education institutions or programmes is clearly an admirable ambition, and there is no doubt some development potential in this area. Further investigation of a number of questions is required, however, and the question of what the actual target group – students – wants is of at least equal importance. This calls for an awareness that the simplification of information involves an information loss, and that information that is simplified to the extent provided in ranking lists based in combined weightings hardly provides any information at all about the quality of higher education.

## Information to students

In the previous chapter, we tried to survey the ranking of universities and other higher education institutions – both in broad, general terms and also in analyses of specific ranking systems. In this assignment, we have been asked, based on our survey, to assess the prerequisites and the advantages and disadvantages of ranking for Swedish higher education institutions and programmes.
No survey is required, however, to note that there are prerequisites for the ranking of Swedish universities and university colleges – at least if we interpret the existence of Swedish ranking as proof of the prerequisites. Moderna Tider, Fokus, Sydsvenska Industri- och Handelskammaren and Svenskt Näringsliv have all published rankings of Swedish universities, institutions and programmes on repeated occasions.

However, based on the reasons stated by the Government for its decision to give the Swedish National Agency for Higher Education this assignment, we have interpreted this as an indication that the Government is interested in the prerequisites for a particular type of ranking – a ranking designed to inform potential students and other students about quality differences:

Studying at a higher education institution involves a considerable investment for the individual concerned, as regards both time and money. As a result, applicants must be able to make high demands in terms of relevant and comprehensive information concerning the education programme in which they are interested. In order to facilitate the student’s choice, the quality differences that exist between programmes at different institutions must be indicated more clearly.148

Hence, our discussion in the following as to whether ranking can improve students’ access to relevant information regarding differences in quality between different education programmes and higher education institutions in general and, in particular, in Sweden, is based on this perspective.

As a result, this chapter focuses exclusively on ranking at the university and university college level that is designed to provide students with information about higher education. Ranking for other purposes, such as the allocation of resources, accountability requirements, accreditation, etc. is not touched on in this context. Ranking for purposes of this nature calls for quite different and more extensive investigation of the quality of higher education – what it is or what it should be – than can be covered within the framework of this assignment.

This chapter starts with a general discussion of ranking as an information tool for students. What types of rankings are to be preferred for this purpose, according to our previous survey? What indicators are important from a student perspective, and what data sources can provide information?

In this context, we also discuss the prerequisites for a ranking system of this nature for student information in Sweden. What are the available indicators and sources of information from a Swedish point of view? What opportunities do we have to use indicators and sources of information to this end, and what are the problems?

148. Government decision U2008/3731/UH.
Subsequently, we consider the general advantages and disadvantages of ranking for student information purposes in Sweden, and finally we discuss a few key factors that call for further investigation.

**Ranking as information for students?**

In many ways, the previous survey of both ranking in general and various existing ranking systems is depressing reading for someone who believes in ranking as a source of information for students. Criticism of ranking tends to be strongly negative, and many of the existing rankings have serious deficiencies from both the validity and reliability angles. In most cases, these rankings give students information that is neither comprehensive, relevant nor reliable.

There are some exceptions, however. The German CHE ranking, the British Unistats.com website and the Australian Good Universities Guide all manage to avoid many of the methodological problems. Some of the commercial rankings, for example the Guardian and Macleans, also provide specific rankings in parallel with their main product, and they are probably more relevant from a student perspective.

The question of the best way in which students are to be provided with satisfactory – in the sense of relevant and comprehensive – information about higher education is, of course, crucial, even if the rankings we have studied are not very impressive. We should therefore like to discuss a couple of key issues concerning ranking as student information – or simply student information – that appear to be particularly important, based on our survey.

**Ranking of entire institutions or ranking of subjects?**

The vast majority of existing rankings, and many of the rankings we have studied in more detail, primarily rank institutions, with little attention paid to subjects or education programmes. It would be hardly controversial to claim that such rankings only fulfil students’ information needs to a very limited extent.

Although it is perfectly possible that certain students or potential students are in fact interested in what is the “best” higher education institution overall, it is doubtful if the information provided is either relevant or comprehensive in general terms. A student who intends to be a doctor is hardly interested in discovering that the Stockholm School of Economics provides the “best” education in Sweden. In most cases, the choice of education programme is more important than the choice of institution. And even if this is not the case, is there a sensible answer to the question of which institution is best? Best for whom, and best at what?

The academic literature in this area overwhelmingly supports rankings at a less aggregated level than an entire university. The quality in different areas
and subjects may vary significantly within a specific university. A ranking that does not take this into account hides more information than it provides.\textsuperscript{149}

Furthermore, ranking entire institutions without taking their profile into account poses considerable methodological problems. One example of an obvious problem is that rankings in which bibliometric indicators have a heavy weighting (particularly international rankings) favour universities with large medical and natural science faculties, since bibliometric data reflects publication procedures for these disciplines in a positive light.

Regarding all higher education institutions as equal ("one-size-fits-all") is not defensible on methodological grounds and, in the worst case, results in an undesirable standardisation within the higher education sector. This criticism is supported by the "Berlin Principles", established by Unesco-Cepes in Berlin in 2006 to enhance awareness of ranking and to improve and refine the use of ranking methods. The third Principle states specifically that it is essential that rankings are capable of depicting the diversity of organisations in higher education, with different assignment and different targets for their operations.\textsuperscript{150}

But this is not to say that indicators at the institution level cannot be of considerable interest for students (and others). Some aspects of higher education are simply organised or controlled at the central level, and many student welfare issues may fall into this category. Some aspects of potential interest for students are also at even more aggregated levels – for example the characteristics of the town/city in which the institution is located, or even the characteristics of the country concerned.

\textbf{Multidimensional or weighted rankings?}

From a methodological viewpoint, combining the weightings of various indicators to achieve a total result is the most frequently questioned and problematic feature of ranking activities. There are several different rather complex technical problems involved in weighing various indicators together to produce a single scale, and this is discussed on more or less technical grounds in much of the current literature on ranking.\textsuperscript{151}

These technical and statistical problems are fully described in the HEFCE report referred to previously in connection with the British rankings. In an appendix to this report, the authors thoroughly analysed the impact of “normalisation”, standardisation, weighting and scoring. Without wishing to examine the technical aspects in any great detail, we may nonetheless summarise a couple of problem areas in this context.

\begin{itemize}
\item \textsuperscript{149} Westerheijden, Federkeil, Cremonini, Kaiser & Soo (2008), Marginson (20089, van Dyke (2005) and Dill & Soo (2005).
\item \textsuperscript{150} www.ihep.org/Organization/Press/berlin_Principles_Release.pdf
\item \textsuperscript{151} However, perhaps the most amusing illustration of this problem is on the "Ranking Game" website (http://monoborg.law.indiana.edu/LawRank/) which offers an opportunity to choose and weight indicators for American law schools. Playing this ranking game provides a quick and illuminating lesson in the problems involved in combined weighting methods.
\end{itemize}
Firstly, there is no real consensus in the ranking field as regards the way in which various methodological aspects of the combined weighting of the indicators are implemented. In some cases, normalisation means that the indicators are computed in relation to the size of the institution concerned, while in other cases the mix of subject areas at the institution is taken into account. Standardisation may take the form of taking into account the distribution of the various indicators and then standardising on this basis, but sometimes much simpler principles are applied. The weights selected for the various indicators are completely in the hands of the specific ranking designers, and there is little consensus as regards the weightings applied to different aspects. And when the results for the various institutions are finally correlated, this is accomplished in many different ways – often by allocating a value to the institution with the “best” results and then relating the other institutions to this – although there are variations on this theme. The greatest problem is, perhaps, that the rankings seldom apply these procedures in accordance with statistical theory\textsuperscript{152} and, in addition, they are not mutually consistent.

Furthermore, the methods applied are not always transparent. As a result, many ranking designers fail to comply with another Unesco-Cepes Berlin Principle – the sixth Principle stressing the importance of transparency in the methodology on which the rankings are based.\textsuperscript{153}

The methods used to weight indicators into a cohesive result are complex – and this is inevitable since the indicators are so different. But the processes applied introduce new potential sources of error and new uncertainties at every stage. For example, they make it hard to understand what the final list is really measuring, and statistical analyses have shown that what the list ultimately indicates is not always what the ranking designer intended. This does not refer to the choice of indicators – which is a validity problem – but rather to the fact that the weightings applied sometimes distort the results. Even if indicators for teaching and the quality of education are included, and even if they are given a relatively heavy weighting, many rankings ultimately tend to result in a list that tends to reflect prestige or research excellence.\textsuperscript{154}

There are also technical problems, in addition to the theoretical problems of the weighting of indicators to produce an overall result. The selection of specific indicators and their weightings is a reflection of an assumption about what defines quality – in a rather precise manner. This, in its turn, raises further problems.

Firstly, the definition of quality is seldom fully thought through or empirically based. In the previous chapter, we have noted that many rankings are based on existing indicators and, as a result, are readily available, rather than


\textsuperscript{153} www.ihep.org/Organization/Press/berlin_Principles_Release.pdf

on indicators that the constructors have endeavoured to define and tried to capture. In other words, quality is defined in terms of something that is easy to measure. This also results in the inclusion of indicators that, according to empirical research, have very weak links – if any – with the quality of teaching (e.g., resources at the institution concerned). Once again, this contravenes the Unesco-Cepes Berlin Principles, since the seventh Principle states that ranking designers must select indicators on the basis of their relevance and validity.

Secondly, it is by no means certain that students share the ranking designers’ definition of quality – despite the fact that potential students are the main target group for most rankings. The aggregated weighted ranking result implies that all students have an identical opinion as regards what quality means, and that, as a result, this is something that benefits all students. This assumption does not hold water, of course. On the contrary, empirical research indicates that students are anything but identical as regards their choice of higher education. Some students give a low rating to a research focus, for others the size of the institution and the subject area are the key factors, while some students give priority to good study counselling or an international focus.

In other words, the information available to students in the form of weighted, aggregated rankings is neither comprehensive nor relevant. In this case, simplification of a complex reality is in direct contrast to the quality of the information provided. As a result, multidimensional rankings of the German CHE type, for example, are to be preferred from a student perspective. If the aim is to give each individual student the possibility of personally defining what quality involves, it is better to provide information for as many indicators as possible, without weighing them up to produce a one-dimensional result.

Multidimensional rankings have their problems too, however. They are also based on a selection of indicators that restricts the individual student’s possibilities of defining quality. Fewer indicators mean greater restrictions on the possibility of depicting various aspects of quality. The type of indicators that are available in multidimensional systems is also important – and, here too, the ranking is clearly less relevant for students if the indicators included are confined to those that are easy to measure.

Another problem with multidimensional rankings is their scope. Basically, they are not so viable in media terms as the weighted rankings. If one is unable to answer the question of what is the “best” university in the world – or in the nation concerned – because “this depends…”, it is not possible to achieve headlines on the front page of a newspaper. This may be less of a problem for non-commercial entities (such as CHE or Unistats), but it also has an

---

155 In this context, reference is often made to the story of an inebriated man who drops his keys in the middle of the road but searches for them under a streetlight because the light is better there.


impact, of course, on their efforts to reach their target groups in a rapid and simple manner.

**Ranking or student information systems?**

Many of the existing (weighted) rankings list institutions, subjects or programmes in very precise terms on the basis of the results for a number of indicators. But this can hardly be defended on statistical grounds since, in many cases, the differences are too small to be statistically significant – that is to say that it is not possible to conclude that there is a genuine difference in quality. Hence, many rankings appear to provide a degree of exactitude in the quality ratings for higher education institutions that is not justified. Once again, simplification compromises the quality of the information.\(^{160}\)

For this reason, the German CHE and the Australian Good Universities Guide employ a different presentation technique. Instead of rankings on a continuous scale, they apply discrete rankings, and instead of allotting an exact value to institutions (in various subjects), the CHE ranking places the institution concerned in one of three groups – good, average or poor. This classification is, of course, ultimately based on the exact results, but presentation in broader categories means that minor differences are not exaggerated or, more precisely: small differences within the groups are not exaggerated, although this procedure may clearly exaggerate the differences between groups. In addition, this broader classification may be criticised on the grounds that it hides information – particularly since the threshold is frequently not based on the boundary line for statistically significant information.

The British Unistats.com website, which we have also examined in more detail in the previous chapter, quite correctly does not refer to its operations as ranking, although it is nonetheless of interest from a student information viewpoint. Unistats only presents information concerning all the indicators registered for the subject selected at the institution selected, rather than a ranking list for such institutions. Failure to present rankings, however, means that students do not have access to information about quality differences between institutions – users must make such comparisons themselves. The user must also personally determine whether or not such differences are statistically significant.

The question is whether student information must be in the form of “ranking”, and whether this term should be employed. If we start with the terminological question, it is in many ways more appropriate to refer to rankings in the context of multidimensional presentations – in point of fact this involves an indefinite and varying number of rankings. In view of the way this term is used in an international context, ranking is perhaps a somewhat inadequate designation. A more suitable and clearer description of multidimensional ran-

---

\(^{160}\) Van Dyke (2005).
kings would probably be “student information systems” or “a system for student information”.\textsuperscript{161}

The question of whether there is some advantage in calling this form of student information “ranking” also comes under the terminological heading. Ranking is a concept which indubitably attracts media attention, although it is widely regarded with some scepticism – at least in the academic debate. If there is a danger that the ranking concept repels some interested parties (in particular potential students), a more neutral concept might be preferred.

A more down-to-earth question concerning whether ranking is a necessary feature in giving students information about quality differences is related to the degree of openness in the approach to individual definitions and interpretations of quality. All ranking systems – even those in the cruder categories – provide an evaluation for each indicator. In other words there is an answer to the question of what is good about each separate indicator. In multidimensional rankings, the answer to the question of what is good in general terms – what is overall quality – is open to definition by each interested party, at the individual level. But the criteria for the various indicators are defined and the institution concerned either has high or low values, if it is to be classed in the highest group for a particular indicator. There may, however, be several different individual variations in the definition of what is “good”, even for specific indicators. For example, some students may be interested in a programme with as many international students as possible, while other may prefer the exact opposite, and others may consider a mixture of international and Swedish students to be the best option. Similarly, there may be considerable differences in individual preferences in terms of gender distribution, qualifications for new students, the proportion of traditional teacher-led education, the student/teacher ratio, and so on.

Ranking (and quality assessment in general) for purposes other than student information must, of course, be based on a carefully considered and precise definition of quality, at least if it is to be used for resource allocation or accreditation purposes. In such cases, the choice of indicators, and their evaluation and weighting are critical – there is no scope for individual variations in the definition of quality in this context. But in the case of students’ selection of programmes, the reasons for not permitting every possible, feasible variation in quality definitions are less pressing. In other words, both the selection and evaluation of the indicators can be left to the students themselves.

\textbf{Student indicators}

Even if the task of defining quality is transferred to the students as suggested above, one does not avoid the problem of presenting valid, reliable and transparent indicators. Since the purpose of multidimensional, interactive rankings

\textsuperscript{161} Westerheijden, Federkeil, Cremonini, Kaiser & Soo (2008).
(or information systems) may vary for each potential user, and the aims are not uniform or clear, it is not possible to determine what are the valid indicators. On the contrary, it may be claimed that if a ranking or information system is to cater for valid quality assessment for many different people, it is essential to adopt a broad approach, this permitting a wide variation of individual goals and preferences. This also increases the possibility that the students will regard the information as relevant.

In the following, we discuss various types of indicators which may be envisaged in the context of information to students about higher education. First we describe indicators that are commonly employed in the rankings we have examined previously in this report, with their advantages and disadvantages. Subsequently, we discuss the possibilities of obtaining information in the various areas concerned in Sweden. And finally, we also discuss information which is less common, but which may be of potential interest to students.

**Entrant characteristics**

Several rankings presented previously include one or more indicators for the characteristics of students who are commencing higher education for the first time. In most cases, this involves measures for some kind of qualifications – grades from previous studies, the results of various types of tests, or a combination of both. The British rankings apply slightly different indicators for qualifications, but in principle they involve some form of merit-value obtained from UCAS, the central admissions service. In the United States, SAT results are normally used, while the Australian surveys are sometimes based on knowledge tests for new students.

Such indicators for the qualification of new students are often used as a measure of the institution’s attractiveness (see, for example, the Times, the Sunday Times and Macleans). The better the qualifications of new entrants, the better the institution is considered to be. In an overall ranking, assessment of this indicator is essential – a decision must be made as to what is a good value when the weighted indicator is combined with other indicators.

In other cases, this indicator is used to assist in the computation of the value-added provided by the education programme. Entrant qualifications are regarded in this context as a control variable – what is measured is the outcome with regard given to (controlled for) how well-qualified the students were at the starting point (e.g. the Guardian). In a few cases, entrant characteristics are measured with the assistance of indicators depicting the students’ socioeconomic or ethnic backgrounds. In some cases, a high proportion of students from homes with an academic background also functions as a measure of the institution’s attractiveness, but entrant characteristics of this nature are primarily used in more “alternative”
rankings (Washington Monthly and the 2008 Black Enterprise Top 50 Colleges for African Americans\textsuperscript{162}).

Both these types of indicators have, in fact, been employed in a Swedish context in rankings of universities and university colleges. In the Fokus/Urank academic ranking, which has already been described, the proportion of students with higher education aptitude test scores of more than 1.1 above all the test results is used as an indicator.\textsuperscript{163} In the Fokus/Urank alternative ranking for width of recruitment, there are two other indicators for entry characteristics, but they are based on the socioeconomic and ethnic backgrounds of new students: the proportion of first generation students in higher education and the proportion of students with a non-Swedish background.\textsuperscript{164}

In other words, it is possible to cover at least some of these aspects in Sweden, although this is more difficult at the programme or subject level that at the institution level. The Swedish Agency for Higher Education Services (VHS) publishes statistics for admissions to higher education that provide information concerning minimum admission scores for the various selection groups for all programmes at all institutions for each round of admissions. Unfortunately, different institutions use different bases for classification, and this makes it difficult to compare institutions.\textsuperscript{165}

There is also a considerable volume of information in Sweden about the socioeconomic and ethnic backgrounds of students. Apart from the key numbers at the institution level mentioned previously, the Swedish National Agency for Higher Education (HSV) has also studied this question in more detail, in cooperation with Statistics Sweden, prior to publication in statistical reports.\textsuperscript{166} There is also a relatively rich documentation of various aspects of gender equality (i.e. the proportion of women and men) in the education sector, and also as regards regional recruitment (what part of the country do new students come from).\textsuperscript{167}

There are, however, also new entrant characteristics that are of considerable interest – particularly from a value-added viewpoint – but which are very hard to measure. If the concept of new entrant characteristics actually includes all the characteristics and abilities that students take with them when they enter higher education, factors such as critical thinking, analytical reasoning, communication skills, learning strategies, experience and other aspects should

\textsuperscript{162} www.blackenterprise.com/losts/college.asp
\textsuperscript{163} Forne, Lind & Nyblom (2008).
\textsuperscript{164} Forne, Lind & Nyblom (2008).
\textsuperscript{165} See www.vhs.se/statistik
\textsuperscript{166} UF19 SM0801 Utländsk bakgrund för studerande i grundutbildning och forskarutbildning 2006/07 (2008) and UF20 SM0802 Universitet och högskolor. Föräldrarnas utbildningsnivå för högskoleavbörjare 2007/08 och doktorandsavbörjare 2006/07 (to be published in December 2008) and “Nyckeltal” at www.hsv.se/nu-statistikdatabasen/.
\textsuperscript{167} See ”Jämställdhet” and ”Regional rekrytering” at www.hsv.se/nu-statistikdatabasen/. In the case of gender equality information, see also Swedish National Agency for Higher Education report 2008:20 R.
also be included. Unfortunately, it is hard to define and detect such characteristics.168

**Indicators for education resources**

Resource indicators are perhaps the commonest type of indicators employed in rankings all over the world, but their usefulness is strongly questioned. Resource indicators measure the resources devoted to education (in a subject or at an institution) in a broad sense. This may be a question of purely financial resources (e.g. how much money is spent on education), physical resources (e.g. access to libraries, laboratories and IT equipment), or personnel resources (e.g. access to teachers, number of teachers with doctoral degrees, or professors).

The most frequent criticism of resource indicators is that they do not necessarily have anything to do with quality. There is little empirical evidence that the quality of education is in any way dependent on resources of this nature. The question of whether teachers with doctoral degrees are necessarily better than other teachers is often raised, for example.269

This type of criticism is less problematical, however, in the case of interactive, multidimensional rankings, where the resource indicators only represent a limited proportion of the large amount of information available. And, of course, there is nothing to say that certain students may be very interested in the resources of various types available in the education programme that they are considering.

A more significant problem is that the resource indicators are nonetheless often expressed in simple terms – this is probably also one of the reasons why they are so frequently used in rankings. In most cases, resource measures convey a considerable degree of exactitude – the number of dollars per student or the number of students per teacher. But exactitude of this nature effectively hides the measurement problems associated with virtually all kinds of indicators. What is the definition of a “student”, how are students linked to a specific subject, and – even more difficult – what is meant by a “teacher”?2

This type of indicators is nonetheless often applied in the Swedish rankings. The Sydsvenska Industri- och Handelskammare ranking, for example, includes indicators for the number of FTE (full-time-equivalent) students per teacher and the percentage of teachers holding a doctoral degree. Such key numbers are readily available from official statistics in this area.170

From a student perspective, however, one may wonder how relevant such data is, since it covers an entire institution and provides little indication of the situation for a specific programme or subject area. Furthermore, characteristics

---

168. Finnie & Usher (2005). Experiments have been conducted in Australia, however, in which the Australian Council for Educational Research (ACER) tried to develop tests to capture generic skills for both new and departing students, see ACER Occasional Paper Series 01/E.

169. See, for example, Usher & Savino (2006), HEFCE Issues paper 2008/14 and Yorke & Longden (year unknown).

170. See also www.hsv.se/nu-statistikdatabasen/ or Swedish National Agency for Higher Education report 2008:19 R.
of this type cannot be readily disaggregated, and this is discussed further in
our analysis of the official statistics.

The resource indicator category can be extended in various ways. The bound-
dary line between what we have termed process indicators is by no means clear.
It is possible to include measures for the way resources are organised in the
resource indicator category – for example contact hours with teachers, the edu-
cational format for degrees, etc. might also be included under this heading.171
However, since this type of measure is so unusual in a ranking context, we
have chosen to regard indicators that refer to the way in which resources are
actually used as process indicators – and place them in a separate category.

On the other hand, where the focus is on the ranking of education pro-
grammes – particularly ranking in the form of information to students – it is
reasonable to regard the various indicators for research and research quality
as resource indicators. In this case, it is assumed that the quantity, focus and
quality of research may benefit students, especially third cycle students. It is
obviously possible to raise the same objections against these indicators as for
the traditional indicators described above – they have little relevance for the
quality of education. This is an empirical question, however. Furthermore, we
cannot exclude the possibility that, in a multidimensional, interactive ranking
system, many students might decide to focus on research indicators, among
other aspects.

The most commonly applied measures of research quality in rankings are
various types of bibliometric indicators, such as citation rates and external
research funding. And indicators of research quality and quantity are also
employed in the existing Swedish rankings. The Fokus/Urank ranking includes
measures for the research allocation per teacher holding a doctoral degree,
the proportion of the institution’s total expenditure devoted to research, the
proportion of research funding subject to competition, the proportion of stu-
dents who transfer to third cycle studies, and the number of doctoral theses
per professor. Sydsvenska Industri- och Handelskammare includes measures
for the proportion of external research and the number of scientific articles
published in relation to research revenues.

The official statistics contain some information about research allocations,
but once again it is difficult to break them down at a level that is really relevant
for the individual student.172 International databases are required for bibliome-
tric studies that take into account citations and relate them to the global aver-
age. One instance is the commercial Thomson Reuters/ISI database, employed
to some extent in the Swedish Research Council’s publications database.173
This database’s focus on English-speaking publications, and discrepancies in

---

172. See “Ekonomi” at www.hsv.se/nu-statistikdatabasen/
173. See, for example, Karlsson (2006) for a more detailed description of this database and ways
in which it can be used.
its coverage of various subject areas gives rise to problems concerning relevance and reliability in some cases, however.\footnote{174}

Other types of information about research are also probably of considerable interest for students, particularly the focus of research (what aspects of a subject have special priority at a given institution, what types of research have a strong basis in terms of scientific theory and methodology, the way in which student counselling is organised, etc). This type of information is probably most clearly expressed in the form of peer review evaluations, conducted both by institutions and research funding organisations. One complication in this context regarding the use of such evaluations as indicators in an information system for students is that they often have different aims and, as a result, are implemented in different ways. This makes it difficult to use them to compare different institutions in a research context, both as regards the manner in which they are conducted and the results achieved.

\section*{Indicators of education results}

Result indicators are another common feature of existing rankings. Criticism in this context has not involved the validity of such indicators to the same extent as in the case of resource indicators, and their relevance has also been questioned to a relatively limited extent. The Unesco-Cepec Berlin Principles explicitly state that ranking designers are to select result and outcome indicators rather than resources indicators wherever possible.\footnote{175}

On the other hand, the relevance of the indicators has been questioned on the grounds that there does not appear to be any consensus as to which indicators should be used, and because the indicators that are most commonly employed are crude measures of results.\footnote{176}

The commonest measures are the proportion of graduates, the number of credits within a given time period, throughflow, or drop-outs. The Swedish rankings, Sydsvenska Industri- och Handelskammare, for example, uses degree ratings in third cycle studies and an overall weighted value for the number of graduates per number of FTE students and average credits after three years.

These are, of course, relatively crude result indicators, but once again they are easily measured and they are available in the official statistics.\footnote{177} In the British rankings, for example, a somewhat more sophisticated measure is employed – the proportion of graduates in the highest or next-highest class. This would be difficult to apply in a Swedish context in the current situation, since there


\footnote{175. www.ihep.org/Organization/press/berlin_Principles_Realease.pdf}

\footnote{176. See, For example, HEFCE Issues paper 2008/14, Malandra (2008) and Yorke & Longden (year unknown).}

\footnote{177. See www.hsv.se/nu-statistikdatabasen/ and National Agency for Higher Education report 2008:19 R.}
is no clear, nationally defined grading system for higher education.\textsuperscript{178} In addition, it is considerably more difficult to measure such grades at a subject or programme level – this also applies to result indicators more generally – than at an institution level, despite the existence of statistical reports published by the Swedish National Agency for Higher Education and Statistics Sweden.\textsuperscript{179}

Result indicators should also involve the actual knowledge and skills that students acquire as a result of their education, and not merely the number of graduates and throughput. This involves everything from subject-specific knowledge to more general accomplishments, such as critical and analytical thinking and a scientific approach. The ability to work with others and to communicate may also be relevant result indicators.\textsuperscript{180}

The fact that indicators that attempt to measure students’ actual knowledge and skills are employed to such a limited extent is, of course, because they are hard to measure. The OECD Assessment of Higher Education Learning Outcomes (AHELO) project is one attempt to develop knowledge tests in higher education that might manage to measure such outcome indicators more successfully.\textsuperscript{181}

**Indicators for final outcome**

Indicators of final education outcomes measure the more indirect outflow from higher education. The commonest measure is the degree of establishment in the labour market, unemployment, the degree of qualifications in working life and, in some cases, income. The degree of establishment is included in the Fokus/Urank ranking, for example.

Indicators for establishment in the labour market are almost certainly of considerable interest to students, but measurement presents methodological challenges. Establishment in the labour market, the degree of job qualifications and incomes depend on several different factors, one of which may be higher education, although the links are not self-evident. In other words there are a great many variables that would need to be controlled for before any correlation between the quality of education and establishment in the labour market can be concluded. Labour market trends, the geographical location of the institution concerned and the business cycle are example of societal factors that play an important part in this process and, in many cases, they vary considerably over time. In addition, of course, there are a number of individual

\textsuperscript{178} Higher Education Ordinance (1993:100), Chapter 6, Section 18. See also Moderna Tider, No. 102, year 10 (1999) which refers to this problem 10 years ago – the situation has become increasingly complex since certain institutions are now grading in accordance with the ECTS seven-grade scale.

\textsuperscript{179} The most recent publication is UF20 SM0801 Universitet och högskolor, Studenter och exa-mina i grundutbildningen 2006/07.

\textsuperscript{180} Finnie & Usher (2005).

\textsuperscript{181} See www.oecd.org/document/22/0.3343.en_2649_35961291_40624662.1_1_1_1.00.html
factors that may be assumed to have an impact – and they are probably more important than the institution concerned or the education programme.\textsuperscript{182}

The Swedish National Agency for Higher Education describes aspects of establishment in the labour market for students in various education areas and institutions on an annual basis, and it also covers the methodological problems involved in drawing conclusions about the part played by the education programme or the institution in the degree of establishment.\textsuperscript{183}

Particularly in the economics area, there has been a focus on income differences between students from different Swedish higher education institutions. But the main conclusion in most of the studies that have been implemented is that no such differences could be noted.\textsuperscript{184}

\section*{Indicators for the process between resources and results}

The term “process” does not really make it clear that this category is designed to depict much of what might reasonably be considered to be of importance from a student information viewpoint, but which is seldom picked out in the rankings. In effect, this involves trying to open up the Pandora’s box of the education process itself (and students’ experience of their education).

To the extent that indicators for the process between inflow and outflow actually exist, they are based on student questionnaires (the National Student Survey (NSS) in the British rankings, the Australian Graduate Survey (AGS) in Australia, and the Fokus/Urank Studentspegekl [Student Mirror] in Sweden). In one case (the Sunday Times ranking in Britain), indicators are also taken from quality assessments of education programmes (the Teaching Quality Assessment (TQA) conducted by the Quality Assurance Agency (QAA)). Both education evaluations and student questionnaires are described in greater detail in the following, under the “Data sources” heading.

The Svenskt Näringsliv cooperation ranking has only one indicator (although in a somewhat different format), but this indicator might be included in the process category – cooperation between higher education and the private sector. The possibility of developing relevant and reliable key numbers for cooperation between education programmes and working life would obviously be of interest for students. The Swedish National Agency for Higher Education’s study of various aspects of such cooperation and its current status in Swedish universities and other higher education institutions provides valuable information in this context.\textsuperscript{185}

The international dimensions of education, opportunities for distance studies, teaching methods and subject profiling are some of the other possible

\begin{flushleft}
\textsuperscript{182} HEFCE Issues paper 2008/14 and Yorke & Londen (year unknown). \\
\textsuperscript{183} The most recently published report is Swedish National Agency for Higher Education report 2007:52 R, but a new report is due to be published in the near future. Agency for Higher Education reports 2006:7 R and 2007:56 R describe the establishment of third cycle graduates in the labour market. \\
\textsuperscript{184} SOU 2008:69, Långtidsutredningen Appendix 8, 2008. \\
\textsuperscript{185} Swedish National Agency for Higher Education report 2008:10 R.
\end{flushleft}
indicators that come under this heading. Education evaluations provide some information on certain aspects, although in other cases the position is less satisfactory.\footnote{Swedish National Agency for Higher Education reports 2005:1 R and 2008:15 R cover internationalisation, report 2003:31 R deals with gender equality, student influence and social and ethnic diversity, while report 2008:20 R discusses equality at all levels of higher education including the staff level. Report 2008:1 R describes particularly excellent education environments.}

**Indicators for student welfare factors**

Various student welfare factors – that is to say factors that affect students’ lives outside their actual education – may also be of interest to students. Indicators of this nature include various types of student services, accommodation, leisure activities, entertainment, health care, study counselling and the characteristics of the town or city in which studies are located. Sometimes, student welfare factors are regarded as one aspect of resources for education and research. Ross Finnie and Alex Usher, for example, include a measure for expenditure per student for student services under the resources heading.\footnote{Finnie & Usher (2005).}

We have chosen to treat student welfare factors as a separate category, however, since most of them cannot be regarded as education or research resources. Nonetheless, they may clearly play an important part in the students’ prerequisites for managing their studies in a satisfactory manner.

Although it may be assumed that student welfare factors are of interest for many students, they are not particularly often covered in university rankings. One exception is the German CHE ranking, which includes accommodation and opportunities for sports activities, among other things. The Princeton Review ranking of American higher education institutions also covers a number of student welfare factors. The national student questionnaire asks questions, for example, about study and career counselling, student experiences in the campus area, accommodation and meals, opinions about the local town or city, sports facilities, the popularity of sports and cultural events, and how many parties the students have.

There are not many surveys that cover student welfare aspects in Sweden, although one example is the Swedish National Agency for Higher Education’s evaluation of various support functions for students in terms of study guidance, career counselling and student health care services.\footnote{Swedish National Agency for Higher Education (2007:24 R).} According to the Agency for Higher Education, these support functions are important, since they are prerequisites for the completion of studies.

In the case of study guidance, the Agency evaluated services for students admitted to higher education, including study guidance at the central or local level, or “early warning systems” designed to identify students with study problems. As regards career guidance, the following aspects were analysed: labour market information, workplace training programmes, projects and degree...
projects at workplaces, study visits, feedback on job applications, interview training and personal guidance. Various forms of preventive and social/welfare measures aimed at students were surveyed under the student health care heading. This covered, for example, access to welfare officers, nurses, psychologists, doctors and physiotherapists.

The Swedish National Agency for Higher Education applied a wide range of criteria for these three support functions: committed management, student influence, documentation of goals and strategies, evaluation and follow-up, a methodical development process, cooperation with external parties, satisfactory introduction for new students, satisfactory dissemination of information about operations to students, “early warning systems”, and opportunities for education and continued professional development for personnel. The Agency also investigated, for example, whether there were special solutions for international students, students with disabilities, doctoral students and distance students, and whether a gender equality perspective was applied.

This evaluation revealed that, in the main, student support was satisfactory at higher education institutions. Improvements were required in some areas, however, for example management commitment and responsibility, the monitoring of activities, information to students and student influence. The report also assessed which institutions were most successful as regards the provision of student support functions.

**Indicators for academic reputation**

“Academic reputation” is a relatively common indicator in university rankings which is often used to measure the standing or prestige of an institution or an education programme. Sometimes, however, the aims may be more ambitious, extending to the institution’s or programmes’ “quality”. Reputation is usually measured by means of questionnaires in which, for example, academics, students and employers state their subjective opinions concerning various institutions or programmes.

Reputation indicators were particularly widely used in the earliest university rankings, for example in the United States, at a time when there were still very few “objective” measures of quality. Since the 1970s, however, their application has declined as other data sources have become increasingly important. The reputation indicator is still used in the Times Education Supplement ranking, however, (40 per cent weighting) and in national rankings such as the U.S. News and World Report (40 per cent weighting) and Macleans in Canada (22 per cent weighting).

This indicator may be criticised on various grounds, depending on what it is supposed to cover. Sometimes, the object is simply to measure what certain parties consider to be an institution’s or programme’s reputation, and in this case it is not necessarily a particularly problematical indicator. It is, however,

---

important to consider why it is interesting to know what various parties think, and whose opinions are relevant. Hence, it is a matter of some concern that many reputation surveys fail to raise this question. Instead, a specific group is selected – often academics – without making it clear why the opinions of such a group are interesting. Depending on the purpose of the ranking concerned, there may be grounds for questioning a wider group, for example employers, the private sector, sources of research funding or maybe even the general public.

However, reputation surveys usually aim to measure more than opinions, per se – for example the quality of an institution or programme in one sense or another. In this case, this indicator becomes more problematical. In the first place, one fundamental problem is that there is rarely some common point of reference for what is meant by “quality”. And as a result, one may wonder whether quality is being measured at all. In the U.S. News and World Report ranking, for example, respondents – primarily presidents of higher education institutions – are asked to assess the quality of other institutions on a 1–5 scale, in which “1” represents “marginal and “5” denotes “outstanding”. There is no indication of what is meant by “quality”, however, or what criteria apply for the various scores on this scale.

Another closely-related problem is that the respondents probably do not have sufficient information to have an opinion about all the institutions they assess. The lack of information about other institutions is demonstrated, for example, in a study carried out by the former Chairman of Alma College in Michigan in the context of the U.S. News and World Report ranking methods. He distributed a questionnaire to 138 of his colleagues, including questions about the U.S. News and World Report reputation survey. More than 80 per cent of the respondents had no knowledge of some of the institutions they were supposed to rank, and more than 40 per cent tended not to answer questions about institutions for which they did not have sufficient information.

A further problem is related to the ability of respondents to make a neutral assessment of different institutions or programmes. Even if there were a common reference framework for what is meant by quality, and sufficient information about all the institutions concerned, the question of whether the respondents are necessarily objective may be raised. The respondents may have closer links with some institutions than others, or they may have preconceived opinions. This is particularly problematical in cases in which respondents are not only expected to assess other institutions, but also their own. An Australian study indicates that there is a clear tendency for academics to favour their own institutions in such rankings.

---

190. See, for example, Stella & Woodhouse (2006) and Salmi & Saroyan (2007) for a discussion of the problems involved in reputation surveys.


Yet another problem with the academic reputation indicator is that there is a risk of “circularity”, that is to say that those who assess institutions are influenced by previous ranking results.\textsuperscript{193} In this case, ranking has a self-amplifying effect that favours well-established and prestigious institutions or programmes. As Simon Marginson put it: “All rankings generate reputation but reputational rankings recycle reputation without any necessary connection to absolute or relative performance.”\textsuperscript{194}

In other words, the use of reputation as an indicator of quality in education or research is highly problematical. Possibly, reputation might be used if it is considered to be interest, per se. Some students may, for example, be interested in the reputation or status of a given programme in the eyes of a potential employer or other important entities.

\textbf{Value-added indicators}

It is widely claimed that genuine quality assessments at the higher education level call for value-added indicators that can relate student outcomes to the student’s starting point (entry characteristics).\textsuperscript{195} Most rankings do not include such indicators, however. This is due, at least in part, to the difficulty of measuring or computing value-added factors. In addition, considerable criticism has been directed at the inclusion of value-added indicators (if they are used) that are calculated on the basis of rather unreliable entrant or result indicators.\textsuperscript{196} The production of value-added indicators is a dubious undertaking if there are no good measures for what is to be included in such calculations, even if measuring quality in terms of value-added is an attractive prospect from a logical viewpoint.

In the case of multidimensional, interactive information or ranking systems, value-added indicators are superfluous to some extent – such calculations may be performed by the user, if access to the information is satisfactory.

There are, of course, several possible indicators for ranking in general, and for information to students. One option that has been discussed recently is a greater focus on the “third” task performed by higher education institutions, namely cooperation with other aspects of society, and external information about their activities. This may be a question of measuring innovations that will benefit society as a whole, or measuring the university’s social utility in some way (to some extent the Washington Monthly ranking tries to achieve this by including Pell scholarships). In part, this discussion is derived from criticism of the tendency of rankings to merely reflect private rather than public benefits.\textsuperscript{197}

\begin{flushleft}
\textsuperscript{193} Guarino (2005).
\textsuperscript{194} Marginson (2006).
\textsuperscript{195} Finnie & Usher (2005).
\textsuperscript{196} Yorke & Longden (year unknown).
\textsuperscript{197} Marginson (2006) and Montesinos, Carot, Martinez & Mora (2008).
\end{flushleft}
Data sources

Data from third parties: Official statistics in Sweden

Official statistics are used in many rankings, especially in Sweden. They are often sought after and respected since they are quality assured by some official examiner, in contrast with information that the ranking designers have obtained directly from higher education institutions. Five of the 16 Unesco-Cepes Berlin Principles concern the collection of data, and at least three of them (Principles 11–13) stress the importance of employing quality-assured, independent, scientifically collected and reliable data.¹⁹⁸

Sweden has extensive official statistics in the higher education area for information and follow-up purposes. The following is a description of Swedish official statistics in the higher education sector – how quality assurance is achieved, what information of special student interest is available, and some of the limitations that apply to the use of such information as ranking indicators.

Quality assurance

Special rules apply for information that is collected and is to be included in Sweden’s official statistics.¹⁹⁹ The official statistics for the higher education area for which the Swedish National Agency for Higher Education is responsible and on which many of our key numbers is based, are collected by Statistics Sweden (SCB). SCB is responsible for a quality declaration based on a quality concept comprising five principal components:

- **Contents**: mainly statistical target magnitudes.
- **Reliability**: uncertainty factors and how they affect the statistics.
- **Topicality**: time aspects that have an impact on the extent to which the statistics describe the current situation.
- **Comparability and joint use**: possibilities to make comparisons over time and between groups, and to use the statistics together with other statistics.
- **Accessibility and comprehensibility**: physical access to the statistics and their comprehensibility.

The SCB MIS 2001:1 publication Meddelanden i samordningsfrågor: Kvalitetsbegrepp och riklinjer för kvalitetsdeklaration av officiell statistik [Quality concepts and guidelines for quality declaration of official statistics] provides general descriptions and definitions of SCB’s quality concepts and the various quality components.²⁰⁰

Official statistics are based on various needs. The subsequent process involves the development and clarification, in cooperation with universities and university colleges, of uniform definitions for the information requested. Local

²⁰⁰. Meddelanden i samordningsfrågor för Sveriges officiella statistik 2001:1, SCB.
registration facilities are then developed, for example in Ladok, the national study-administration system for higher education. SCB compiles the information from Ladok and maintains this information in the university and university college register. On each occasion when data is collected, there is a dialogue with the supplier of the information concerned as regards possible clarifications, updatings and corrections of the data. This assures the quality and comparability of the information in a totally different way than for information collected from higher education institutions for direct utilisation.

Information in the official statistics – key numbers, etc.

Currently, information or key numbers based on the official statistics available, for example, in the Swedish National Agency for Higher Education’s annual report and in the NU-statistik database are not specifically produced and developed to give students relevant information about quality differences between different programmes and institutions in Sweden. For the most part, such information is aggregated at a level that does not make it possible to precisely convey a quality indicator for a specific programme and a given institution. Although some information concerning basic facts in a numeric form may be of interest to a particular student, it can hardly provide information about the quality of education as a basis for selection of the “best” education programme.

The high aggregation level of the official statistics is partly due to the fact that it is based on population data from existing registers, and is not a sample. As a result, it is a rather “blunt” instrument from a student perspective, and students’ special interests are not specifically involved. The official statistics have not been developed with any special idea of quality aspects in the actual teaching or education in mind, particularly since there is no clear definition of what might constitute quality in this case.

Nonetheless, the official statistics contain a rich source of information concerning both higher education institutions and specific education programmes. A wide range of information is presented in the Swedish National Agency for Higher Education’s NU-statistik database. Key numbers are reported in the Agency for Higher Education’s annual report, both for institutions and for certain education programmes.

Key numbers based on the official statistics are produced because it is thought that they should be relevant and capable of describing developments in an extended time horizon. As a result, it is important that there are no significant changes in the key numbers from year to year. Key numbers often comprise purely quantitative information, but in some cases may also involve ratios. The key numbers do not contain components that specifically measure quality, and are primarily intended for general follow-up purposes. In terms of completeness and reliability, the key numbers are of high quality –

201. See www.hsv.se/nu-statistikdatabasen/
although, to a considerable extent, they reflect aggregates at a high level and are not intended for, or suitable for, describing trends that are subdivided into detailed categories.

There are considerably more key numbers items at the institution level than there are for programmes. The Agency for Higher Education’s annual report and the NU-statistik database present information in the form of key numbers describing the size of the institutions and the scope of their operations. A number of quantitative measures are employed, including result measures such as degree statistics and student credits. Information in the Agency’s various publications describes developments over time, and also similarities and differences between institutions. Key numbers cover a range of quantitative information in the higher education sector at the first, second and third cycle levels – both personnel and financial information. In all, some eighty different key numbers are reported at the institution level.

Key numbers for education programmes

From a student perspective, however, key numbers for various education programmes are probably of greater interest. On the other hand, the Swedish education system offers opportunities for recurrent periods of education, for example studies in the form of one or more independent courses or programmes, and this leads to special problems for the official statistics as a whole and, in particular, for key numbers at a disaggregated level. Certain key numbers for the major professional or vocational degree programmes and a couple of other programmes have been developed, however:

- **Number of students** (total number of students in the programme)
- **Number of new students** (number of entrants in the programme)
- **Median age** (among all students studying in the programme)
- **Admissions pressure** (number of first-choice applicants in relation to the number of new students)
- **Proportion of new students with well-educated parents**
- **Proportion of new students with a working-class background**
- **Proportion of new students with a non-Swedish background**
- **Number of degrees**
- **Number of degrees based on studies in other countries**
- **Proportion of men and women graduates, respectively**
- **Study period for a degree** (median of overall study period in terms of years)

Comparison problems

It is perfectly possible to weight, standardise and combine key numbers, but the problem of what information this actually provides still applies. The key numbers may appear to be precise in simple terms, but each of them involves a number of measurement problems.
At a general level, it is difficult to compare key numbers for specialised institutions and major universities with a broad range of education programmes. Comparisons of this nature result in misleading information about quality differences – or no information at all – and instead they tend to reflect differences in the prerequisites of different institutions. Higher education in Sweden is conducted at some fifty universities and university colleges of very varied sizes and with different educational focuses. The 10 largest universities are responsible for three-quarters of all education and research, and the remaining 40 institutions jointly account for one quarter. This is a clear indication that institutions and their students have different prerequisites for their operations. The differences in the research area are even greater – in this context the 10 largest institutions are responsible for 90 per cent of such activities.

Furthermore, Sweden is a small country. Many programmes at the various institutions have so few students registered that certain types of information cannot be disclosed separately – if the integrity of individual students is to be respected. No key numbers can be computed in such cases, although this does not imply that they do not maintain high quality standards.

Official statistical information is often regarded as “state of the art” indicators in a ranking context. To the extent that such information is available, it is freely employed in the various rankings, often because it is readily accessible rather than because it provides valid indicators of quality. The fact that the exactitude and simplicity of indicators may sometimes be misleading may be demonstrated by two examples of the indicators most frequently used in rankings, namely the number of full-time-equivalent (FTE) students per teacher, which is a resource indicator, and the number of degrees and throughflow, which are result indicators. The methodological challenges involved in measuring and comparing these indicators (based on Swedish official statistics of satisfactory quality) should indicate the difficulties of making quality comparisons of this nature.

**Number of FTE-time students per teacher:** At the overall level, this information indicates the number of full-time-equivalent students per teacher and provides a measure of teacher resources (teaching and research personnel) at the institution concerned in relation to the student population. The student/teacher ratio varies considerably between programmes in different education areas. In addition, there are also significant differences between institutions as regards the proportion of the teachers’ time devoted to research rather than student contacts.

The number of FTE students in the official statistics is based on the number of students registered for different courses in the various education areas. Studies in a particular course may be included as part of a programme. But this course may also be simultaneously registered as an independent student course which is not assumed to be part of a future degree (i.e. it is not part of a full education programme). The FTE measure is a measure of volume that may entail considerable variation among the individuals concerned. The number of
FTE students is based on the scope of the course and the study tempo. One FTE student may, for example, mean five students with a study-rate of 20 per cent, or one student with a study rate of 100 per cent. There are also options in the reporting of FTE students to include or exclude students involved in contract education, international exchanges or incoming/outgoing students. These various definitions for various purposes presumably lead to totally different results. The volume measure is normally used to describe the total education volume and long-term trends, and the FTE volume measure also provides a basis for the allocation of resources.

The information in the official statistics concerning the number of teachers is obtained from the central government authorities’ salary payments system, which reports the number of teachers registered for salary payments in a given week in October. Teaching activity is not linked to specific courses or programmes, and there is no connection between teaching hours and specific groups of students.

In other words, the propensity of ranking designers to employ such key numbers as a direct indicator of quality is problematical. Key numbers are often used as a general measure of volume when describing the activities of universities and university colleges and are, in this context, of considerable value for information purposes since they are based on quality-assured and comparable official statistics.

**Number of degrees and throughput**: the measurement problems are similar for result indicators, such as the number of degrees and throughput. Strictly speaking, the official statistics reflect the number of degree certificates issued on a formal basis. But this does not include students who have met the degree requirements but have not applied for a degree certificate. In other words, the degree information is linked to the formal process, and does not fully reflect the real education volume.

The Swedish education system provides considerable opportunities for individual students to adapt their education process in line with their personal circumstances. There is, for example, no information as regards dropouts from education programmes. If a student has been registered for a programme, it is both possible and permitted to interrupt such studies for a brief or indefinite period. Hence, the throughput for a programme is often affected by individual study objectives which are not necessarily related to the quality of education.

The Swedish education system also provides individual opportunities to participate in independent courses. In some cases, the courses that are completed may be subsequently combined to provide a basis for a general degree, although in such cases it may be difficult to determine an anticipated throughput rate.

What conclusions about quality may be drawn from such information about the number of degrees or throughput? Presumably it is difficult to determine whether the key number, if it is to be used as an indicator, actually
conveys any information about the efficiency or the quality of the education provided. Perhaps this merely reflects different requirement levels, although the key number concerned does provide useful information about the purposes for which it is intended.

Quality assurance standards for official statistics differ from those that apply for information obtained directly from higher education institutions and, on the whole, high standards are maintained, in the sense that the information is reliable. Sweden also has unusually good access to official statistics, with a long tradition of both the collection and quality assurance of data at the national level, including the higher education sector.

However, reliability is not the same thing as validity. Validity must be assessed on the basis of the purpose for which the information is to be used. We have already indicated problems in the uncritical application of official statistical information, both in rankings and in student information addressed to the general public. Both the aggregated national monitoring character of such information and the problems involved in defining, operationalising and measuring different aspects call for a high degree of awareness when they are used to compare quality. When employed for their original purpose – to describe the activities of universities and university colleges – the official Swedish statistics offer a rich source of information about higher education in Sweden, both for students and other interested parties.

**Data from third parties: Swedish National Agency for Higher Education evaluations**

Few rankings use indicators obtained from other quality assessment or accreditation activities. The Sunday Times “University Guide” is the only ranking of those discussed in previous chapters that employs information obtained from official government quality evaluations. One reason why such evaluations are used to such a limited extent is presumably due to the transparency factor, as discussed in chapter 1. It is not always possible to summarise evaluations based on extensive “peer review” activities in terms of the simple indicators preferred by ranking designers.

Nonetheless, quality evaluations do contain information that may be of considerable interest for students. The Swedish National Agency for Higher Education’s evaluation activities are described in more detail in the following. We also present, in particular, special evaluations of programmes in the medical, health science and health care fields, in which quality is ranked on a four-grade scale – something that may make information about the quality of education programmes more readily accessible for students.

**Evaluation and ranking of quality in education programmes**

For the most part, the evaluation methodology applied by the Swedish National Agency for Higher Education is based on international research and on many years of experience of quality assurance in the higher education sector,
particularly in Europe. Swedish evaluation research has also played an active role in the development of explanation-oriented methods, which have been applied in all evaluations since 2001. But, above all, quality assessments have acquired a unique status in Sweden in the form of the frameworks embodied in the Higher Education Act and the Higher Education Ordinance. The fundamental prerequisites for this evaluation methodology include uniformity in the national system of governance, the responsibility of public authorities and transparency based on the principle of public access to information. This enhances both the reliability and the validity of the public education system’s methods of assessment.

The Swedish National Agency for Higher Education ranked the outcomes of quality assessments on a differentiated scale in 1996 and 2007 for two extensive evaluations of first cycle programmes leading to qualifications in medicine, health science and health care. The first evaluation was carried out on a special assignment from the Government, while the second was one aspect of a government assignment to assess the quality of all higher education for general and professional qualifications in recurrent six-year cycles. We will return to these evaluations at the end of this section.

Combination of evaluation and accreditation

The Agency for Higher Education’s assessments represent a combination of quality evaluation and accreditation, in the sense that the Agency is an independent authority that decides to issue or withdraw the entitlement of a given institution to award degrees at the first and second cycle levels in the public higher education sector. In other words, an entitlement to award degrees is not an inalienable right, but can be rescinded if it is considered that the education provided does not meet acceptable quality standards. There is no appeal against the Agency’s decision, under the current rules.

The common factor for all education evaluations and, where applicable, also reviews of the entitlement to award degrees is partly the method employed in making such assessments, and partly the quality aspects and criteria on which external assessments and judgements are based.

---

208. Decisions concerning entitlement to grant degrees in the third cycle area and by private course providers are taken by the government, normally following a review by the Agency for Higher Education.
Methods for assessment and judgement

In an international comparison, the model for quality assessments in Sweden may be said to resemble the “minimum standards” method, to the extent that the state, as a source of funding, insists that the institutions provide education programmes that maintain high quality standards. However, in practice this model has not proved an obstacle to the ranking of a large number of programmes in the medical, health science and health care fields on a graded scale.

As in the case of the minimum standards method, the Agency for Higher Education’s assessment of education programmes is conducted by external experts, academics and professionals in the relevant central area for general qualifications or the professional programme qualification that is to be evaluated. The Agency’s quality assessments are based on a peer review, with the result that students, and in some cases representatives of the professional group concerned, are also included in such assessment groups. In order to reinforce the legitimacy of the assessment groups, experts and specialists are appointed on the basis of proposals by the institutions and student bodies concerned. The evaluation process covers the following stages: self-evaluation, visit to the location concerned, report, feedback and follow-up.

Quality aspects and criteria

The quality aspects and criteria on which the Agency for Higher Education’s assessments are based have been developed in cooperation with the institutions concerned, in the light of an interpretation of the goals for higher education as prescribed by the relevant legislation. The emphasis given to the various aspects depends on the programme’s specific targets and profile, and these change over time. As a result, the experts formulate their starting points for quality assessment in a special reference framework.

The aspects and criteria for the evaluation of first, second and third cycle programmes are classified in accordance with the explanation-oriented evaluation model.

• The programme’s prerequisites: teacher qualifications and skills, education environment, infrastructure
• The programme’s structure: guideline documents, teaching, course materials, and the degree/qualification involved
• The programme’s results: assurance of degree targets, assurance of education quality

One of the key measures of education results covered by this methodology is a summarised assessment of the correlation between the programme’s goals, teaching and degree qualification (“Constructive Alignment”).

In 2007, the profile specifically applied for the evaluation of programmes in medicine and health science consisted of a special assessment of quality in the practical/clinical aspects of these programmes. The views of representatives of the institution concerned and “customers” at the local and regional levels were obtained in the course of “on-site” visits.

Ranking of quality in first and second cycle medical, health science and health care programmes in accordance with a four-grade scale

Although quality evaluations contain a great deal of valuable information for students, this may be hard to penetrate. In some cases, however, a more differentiated quality assessment has been undertaken, and this may also result in information that is more accessible for students.

Both the 1996 and the 2007 evaluations presented the assessments in four groups of indicators on a four-grade scale, in which the lowest level indicated unsatisfactory quality. The summarised conclusions concerning the education results subsequently provided a basis for the Agency for Higher Education’s approval or non-approval of the programme concerned.

In a national evaluation of medical education [läkarutbildningen] at six universities in 2007, experts conducted systematic analyses of the various programmes, based on the Agency for Higher Education’s general quality aspects. In addition, the programmes were assessed on the basis of a number of well-established quality aspects that were also applied in the Agency’s evaluation of medical programmes 10 years ago.

The programmes were compared as regards the degree of fulfilment of six themes selected by the experts and considered to be essential for the planning and implementation of a scientifically based modern medical education programme.

- Decision process and financial control
- Willingness and ability to pursue a quality process
- Study environment
- Contents and structure of the programme
- Quality of clinical education
- Prerequisites for teaching and learning development

The same results were achieved when this analysis was performed in parallel with the quality criteria applied at the international level for assessment of medical education in accordance with the WFME’s global standards\textsuperscript{212} or the SPICES model.\textsuperscript{213}

Although the final conclusion was that Swedish medical education maintains a high standard in international terms, succinct quality differences were noted between the programmes. This resulted in a classification of program-

\begin{footnotesize}  
\begin{enumerate}
  \item WFME, the World Federation of Medical Education: Global Standards for Quality Improvement of Basic Medical Education, Copenhagen (2003) www.WFME.org
  \item SPICES model, Medical Education 1984;18:284–297.
\end{enumerate}
\end{footnotesize}
mes into three quality levels, in which two universities were placed at each level.

The considerable progress made in improving the quality of health science programmes is stressed, both in the follow-up three years after the 1996 evaluation and in the conclusions for the evaluation in 2007. In its summary of the 2007 evaluation, the Agency for Higher Education noted that: “The evaluation indicates that health science programmes have experienced impressive development during the past 10 years”. In other words, the Agency has not found that peer-group evaluations in which assessments from different aspects are reported on a differentiated scale have any negative effects on the development of the higher education sector. On the contrary, experience shows that this evaluation method may provide effective quality incentives.

The purpose of quality evaluations of this type is not primarily geared to giving students information about quality differences, although, quality evaluations may nonetheless contain several items of information that are relevant for many students. However, just as in the case of official statistics, such information may call for certain insights in order to reach satisfactory conclusions, and the data is not always straightforward or easy to understand. Furthermore, as with the official statistics, current information is not always available. Furthermore, many students would probably not accept the Agency for Higher Education’s definition of quality, although the quality evaluations have a valuable role to play, as one of many sources of student information.

**Information obtained directly from higher education institutions**

Many rankings mainly rely on information collected directly from the institutions themselves. This is often because official statistics or other quality-assured information from third parties is of lower quality, is less relevant, or is simply not available.\(^{214}\) Hence, obtaining information directly from higher education institutions is often a second choice.

It may be claimed, however, that almost no one understands the situation better than the universities or the university departments themselves. And, of course, hardly anyone else can produce such up-to-date, relevant and “close to the ground” information. Furthermore, certain types of information – such as student welfare factors, the programmes offered and their form and content – cannot be supplied by other parties. As a result, if information to students is to meet the relevance requirement in a system or a ranking, there are grounds for developing such systems in cooperation with the institutions concerned.

But information obtained directly from higher education institutions also involves problems. If the institutions supply information to ranking designers, there is of course a risk that the information will be formulated in a manner...

\(^{214}\) In Germany, for example, there is very little useful official statistical data about higher education, partly because Germany is a federal republic, with different procedures in the various regions. As a result, the German CHE ranking is heavily dependent on information from the higher education institutions themselves, and hence it relies on satisfactory cooperation with such institutions.
that guarantees a good position on the ranking list. We have already touched on this type of problem in our discussion of rankings in the United States.

Even if such information has not been deliberately manipulated, data provided directly by institutions has not been quality assured in the same way as the official statistics described above. This may have a serious impact on how useful such information is for comparisons between different institutions.

A considerable amount of information that is obtained directly from the institutions is, of course, indispensable in the case of the programmes offered and the institution’s focus and activities. Furthermore, in recent years some institutions have also initiated their own extensive quality evaluations of their research in various fields. In Sweden, both Uppsala University and Lund University have recently published reports containing the conclusions from the expert assessments on which the evaluations were based. This is, of course, a rich source of information for anyone interested in the quality of research in various fields at these universities, but it is not easy to compare.215

**Questionnaire surveys**

Questionnaires of various kinds are used as sources of information in several rankings. They may be of the “peer review” type, in an attempt to obtain the opinions of professional academics about institutions, subjects or programmes. This kind of questionnaire has already been discussed under “Reputation indicators” heading.

The other major type of questionnaire survey is the student questionnaire, widely used in the Australian, German and British rankings. In many cases, student questionnaires are an attempt to depict aspects that are difficult to cover in other ways – for example the process between resources and results. This may involve the quality of teachers, forms of teaching, contacts between teachers and students and overall satisfaction with the education programme, for example.

It is hard to deny that this type of information is probably highly relevant for potential students. The problems with questionnaire surveys tend to rather be of a methodological nature. In the first place, it is difficult to draw conclusions about de facto circumstances on the basis of subjective information. There may be many reasons for the way students feel about their education – and they are not all related to what the education is really like. A general sense of well-being during the period of studies is part of the picture, but this also applies to different levels of expectations and personal experiences.

A more serious methodological problem, however, is that questionnaire surveys of this type seldom provide any reliable information. A sloppy definition of the population (which is not always so easy to define) is one problem, but the chief difficulty is the poor response rate, which seriously threatens the reliability of the information. This is not primarily because dropouts mean too

---

few responses – if the sample is sufficiently large this is not usually a problem. But the real problem is that a low response rate means that the answers are probably not representative. No matter how random the sample, the respondents are not randomly selected. This, in its turn, means that considerable bias may be suspected in the answers – although nobody knows how much or what form this bias takes. Perhaps only the satisfied students completed the questionnaire, or only those who were dissatisfied, or – perhaps even worse in this context – different groups of students answered the questionnaire at different institutions. Information based on such shaky foundations should not be regarded as any kind of information at all.

The problems involved in the various questionnaire surveys carried out vary in scope. In the Australian Graduate Survey, the response rate varied from 30–80 per cent – the response rate is very satisfactory is some areas, but worse in others – with an average of 62.5 per cent. In Britain, the National Student Survey conducted by HEFCE, and used by several British ranking designers, had a response rate of 51 per cent in 2008 (above 50 per cent for the first time). The German CHE questionnaire had a response rate of 21 per cent in 2008.

There are no major annual student questionnaire surveys in Sweden that correspond to those conducted in Great Britain, Germany, Canada and Australia. On the other hand, the Swedish National Agency for Higher Education has, on a number of occasions, conducted “Spegel” [Mirror] surveys with students, and also with doctoral students. Such surveys for these two categories of students may be regarded as part of a larger quality assurance system for Swedish higher education. Student Spegel surveys were carried out in 2002 and 2007, and the corresponding doctoral student surveys were implemented in 2003 and 2008.

Both survey categories – particularly the student survey – have been inspired by the National Survey of Student Engagement (NSSE) in the United States, and they cover information about different dimensions of learning and personal development. The response rates range from 57–67 per cent, and these surveys are carried out in cooperation with Statistics Sweden and various groups of experts.216

The Spegel survey population samples are too small to permit a breakdown for specific subjects or programmes. Nonetheless, to some extent they provide interesting information about the views of students and doctoral students for various dimensions at the national, institution and subject-area levels.

The dimensions covered by the student Spegel surveys are primarily education and values, analytical thinking, student cooperation, fruitful discussions, reading, writing, reporting, teacher support and internationalisation. The doctoral student surveys include dimensions with special relevance for third cycle studies: the introduction to third cycle studies, professional development, dialogue with supervisors. supervision functions, courses, reflections, evaluation

and the study environment. In addition, such aspects as post-doctoral career paths, negative discrimination and sexual harassment are also covered in the Spegel surveys for doctoral students.

Questionnaire surveys aimed at students (and doctoral students) can be an important and valuable source of information for students and, if they are properly carried out – with statistically viable methods – they may provide information with a considerable degree of reliability. However, if information at the detailed level that is relevant for students is required, questionnaire surveys of satisfactory quality call for considerable resources for the collection of data.

**Student information in Sweden – concluding remarks**

Improving students’ access to relevant and comprehensive information about quality differences between different higher education programmes and institutions in Sweden is a desirable objective. This chapter indicates that considerable information of this nature is available, although there are certain gaps in the relevance and reliability of this data that give cause for concern. How such information is to be made readily accessible and transparent from a student viewpoint is a complex question, however, and this calls for further reflection. As a result, we conclude this chapter by raising a few aspects which we consider require further study and consideration.

**Who ranks?**

This assignment does not state who the prerequisites for ranking of Swedish higher education institutions and programmes is to be assessed for. As already mentioned, there is already some ranking of Swedish universities and university colleges and, at a general level, this means that there are prerequisites for ranking in Sweden. The entities that lie behind such rankings in Sweden are, however, either commercial interests or have very specific interests – often with the aim of creating some sort of debate.

If a ranking system designed by – or at least sanctioned by – central government bodies is envisaged, there is reason to recall the first Berlin Principle, formulated by a Unesco-Cepes expert group:

> Purposes and Goals of Rankings: Be one of a number of diverse approaches to the assessment of higher education inputs, processes, and outputs. Rankings can provide comparative information and improved understanding of higher education, but should not be the main method for assessing what higher education is and does. Rankings provide a market-based perspective that can complement the work of government, accrediting authorities and independent review agencies.  

In other words, ranking may be a source of information about higher education, but it should not be the only source. And ranking is a market-based phenomenon that can supplement information from public sector sources.

Presumably, this is a largely a terminological issue – depending on how the ranking concept is defined. But, in an international context, “ranking” – as covered by the Berlin Principles and, to a large extent, in this report – tends to be denoted as an almost exclusively commercial activity that depicts the quality of higher education in superficial and frequently inaccurate terms. This is not to say that quality assessments of a more substantial nature – in which quality is properly defined and operationalised in the form of valid and reliable indicators – cannot be employed for the comparison and final ranking of higher education institutions.

However, ranking in the form of information for students about the quality of higher education, should preferably be multidimensional and interactive, in order to allow for many different individual definitions of quality. In this case, the ranking heading is even more dubious, for the reasons discussed in the above. Providing students with comprehensive and relevant information to enable them to compare different programmes and different institutions in some sort of information system, rather than in a ranking format, may also be undertaken by central government authorities – without contravening the Berlin Principles.

**Target group preferences?**

Higher education studies involve a considerable investment for the individual, in terms of both money and time, and this is one of the reasons for the present assignment. According to the Swedish Government: “As a result, the applicant must be able to make high demands as regards relevant and comprehensive information concerning the education that is of interest.”

What is regarded as relevant and comprehensive information varies from case to case, and it is almost certainly impossible to meet everyone’s needs. Before some new system is introduced, however, there may be strong grounds for investigating in greater detail the extent to which students experience a lack of information and – in particular – what type of information they tend to require. To the extent that there is a demand for such information, this applies to both indicators and sources.

Presentation techniques are also a crucial factor in this context. Existing interactive websites, such as the Unitstats in Britain, the CHE in Germany and

---

218. Government decision U2008/3731/UH.
219. In the course of a meeting with representatives of SFS [National Swedish Union of Students] and SVEA [Swedish Student-pupil Council], it emerged that ranking – at least of the traditional type – was not regarded as comprehensive and relevant information. The SFS representative said that SFS was, in principle, opposed to “simple answers to difficult questions”. Nonetheless, representatives of both organisations could see the advantages of more information of different types about higher education, particularly information about student influence in different programmes.
the Good Universities Guide in Australia, vary in terms of their idiom, user-friendliness, transparency and simplicity. Compromises have to be made, for example, between simplicity and a rich supply of information, and the potential target group may have views about this.

Possibly the target groups should be extended somewhat, when investigating information needs. Other persons with close contacts with potential students and other student categories might be included – for example study advisers at various levels in the education system, and also parents.

**Information quality versus simplification**

Rankings clearly fulfil a function in that they simplify information about a complex real world. But it is also clear that simplification is sometimes pursued to such an extent that the information provided no longer has any value. In the context of a possible continued process of producing simple and readily accessible information about higher education, it is important to be aware that simplification and quality involve some contradictions in terms of the information provided. The most relevant, comprehensive and reliable information about the quality of higher education is not always simple. Even less complex objects, such as cars, cannot be described in the manner that is assumed in a ranking system. A “best car” rating cannot be readily established, and this applies equally to “best university” lists.

Information may be simplified to a certain extent, of course, and this makes it considerably more accessible and transparent, without losing too much quality. Major simplifications of various types of data that can be implemented without excessive information loss call for further investigation and consideration, however.

Simplification also means greater impact. The simpler the message, the greater impact it has in the media. And, in its turn, what appears in the media tends to rapidly become “the truth”. There is widespread concern about the fact that inadequate rankings have such a great impact. Hopes are sometimes expressed about finding a “sensible” ranking system that can take the edge off the international rankings, for example Shanghai Jiao-Tong and the Times Higher Education QS. But this is hardly likely. A “sensible” ranking cannot simplify information about higher education and research to the extent that it becomes as viable in media terms as the international rankings.

One clear example in this context is the Guardian in Britain, which was mentioned in the previous chapter. The Guardian ranking’s sole aim is to provide students with information before they make their choices and, as a result, it emphasises that subject-specific rankings are particularly important and useful. But when the ranking results are published, university rankings

---

nonetheless take pride of place – even though they are the most simplified, least informative and methodologically weakest alternative.\textsuperscript{222}

Media logic determines this pattern to a great extent, of course, but other factors also encourage such tendencies. A report presented to the Swedish Globalisation Council draws some conclusions about the position of Swedish higher education institutions in the light of the Shanghai ranking, and claims that the Swedish institutions perform satisfactorily if GDP factors are taken into account, and that this should “enhance their ability to attract international students.”\textsuperscript{223} The report reaches this conclusion immediately after noting that such ranking lists are highly controversial nowadays, and are strongly criticised in many quarters.

The higher education institutions themselves also contribute to the continued significance of rankings, despite massive criticism of their validity, reliability and relevance. The websites of several Swedish institutions include information and comments about their position in the international ranking lists and, in a press release dated 9 October 2008, the Vice-chancellor of Lund University comments on the improvement in the University’s position in the Times Higher Education ranking in the following terms:

This reinforces the picture that Lund University is highly successful and is transforming its operations. This is based on all our talented teachers and researchers. Such ranking lists help to form a picture of our universities, even though we, in the university world, are sceptical of the way they are drawn up in many cases.

Presumably, the aim of reducing the impact of rankings that have proved to be so viable in the media as the result of a ranking system based on sound indicators and a sound methodology is no more than a pious hope.

**Time perspective**

Generally speaking, ranking designers want to have information that is a fresh as possible. On the other hand, this involves some conflict with a preference for information from third parties, since data of this nature is quality assured to a greater extent than other sources, and quality assurance takes time. The section in the above concerning official statistics and the National Agency for Higher Education’s quality evaluations indicates that information about higher education activities does not depict the current situation. Quality assurance of the official statistics take time, and information for a given academic year is published with a time lag of approximately six months. Quality evaluations also involve protracted processes involving expert assessments after visits to the institutions concerned, and publication long after the evaluation has been carried out. Furthermore, evaluations of a specific subject are performed

\textsuperscript{222} HEFCE Issues paper 2008/14.
\textsuperscript{223} Basic report no.10 to the Globalisation Council (2008).
at intervals of 4–6 years, and this means that some information may refer to circumstances that applied quite some time ago.

Naturally, the implementation and analysis of questionnaire surveys also takes time. In the case of student questionnaires in which students answer questions about the first year of their studies in a three or four year degree programme (as in the Australian case in the preceding chapter), the topicality and reliability of the information may be strongly questioned, of course. And information collected directly from the institutions concerned also needs to be processed. From a student perspective, fresh information is probably valuable since it describes the situation as close to the student’s own study period as possible. Information of this nature is of doubtful quality, however. In other words, information that is relevant from a time viewpoint is not always reliable.

Changes in the quality of higher education may occur rapidly, and this leads to further problems from a time perspective in an information context. Even information on the current situation is not always capable of depicting the situation that applies when students actually commence their studies. Uncertain employment factors at universities and other higher education institutions may result in rapid changes. Some teachers may have resigned or retired, while others may have received research grants, which results in no teaching duties in the semester concerned. Other factors may also shift quickly as a result of changes in the personnel situation – for example the range of programmes offered, the reading list, the size of student groups, teaching methods and exam forms. The flow of, or lack of, resources for research or education may also have a considerable effects on quality, although the impact is not always so rapid.

Changes may be of a short-term nature, which means that measurements at one particular point in time do not always give a full picture. There is always a danger in drawing conclusions from surveys conducted on a particular date, and this also applies to assessments of the quality of higher education. Developments over time are a crucial aspect that gives a fairer picture of the quality of a programme or an institution at any given time point than a measurement at the actual time point concerned. As a result, it is reasonable to expect comprehensive and relevant student information to include information about various quality aspects that cover an extended period.

How can a flexible system be measured?

In our discussion of official statistics, we return to the Swedish education system several times. This system is special in many ways, and this has repercussions on what can be measured in Swedish higher education.

From an international viewpoint, the Swedish education system is unusually flexible. There are considerable opportunities for students to take up their studies again after interruptions of shorter or longer duration. Higher educa-
tion students can also choose to design their own education programme in the form of studies in independent courses. And a general degree qualification allows for considerable variation in the way it is achieved. Many general programmes also offer considerable opportunities for a free choice of focus, as the result of completion of various independent courses. The professional degree programmes have a more precise structure, but there are a number of variations in terms of higher education requirements and qualifications.

This means that comparisons of higher education institutions in Sweden are subject to certain special problems. In point of fact, different institutions have different combinations of programme, and this has a considerable impact on many of the indicators commonly employed in rankings.

But it also has more direct consequences, since we cannot describe much of the higher education conducted in Sweden in terms of indicators that have a summarising function. How should independent courses be assessed in terms of resources and results? Quite simply, it is difficult to provide students with simple and categorical information about a flexible system, even though many aspects of this flexible system are probably of considerable interest for students.

Perhaps the new degree structure that has applied since 1 July 2007 will affect the flexibility of the system. There is good reason to monitor future developments, both from the viewpoint of student information and information about higher education as a whole.

Costs

Major student surveys, the collection and processing of official statistics and quality evaluations with expert assistance all call for resources in their own right. As pointed out above, the quality assurance of information – to guarantee its relevance, reliability and comprehensiveness – is also a process that makes demands on resources.

Ensuring that such information is accessible – via an interactive website or in some other way – also involves considerable costs. Although much of the information about Swedish higher education already exists, this information must be developed, made easy to understand and be updated continuously if it is to provide relevant and comprehensive information for students. The exact costs involved to achieve this also call for further investigation.

National or international information?

To a large extent, this report discusses the difficulties of producing relevant and reliable information within the boundaries of a uniform, national higher education system. Obviously, an extension of this aim so as to supply comparable information that also applies outside Sweden’s national borders is not easily achieved. This applies, in particular, to the official statistics, which have been developed with the aim of communicating and describing developments in higher education in terms of variables that are relevant for the situation and circumstances in Sweden.
Nonetheless, information of an international, or at least, multinational character that compares education programmes in different countries – particularly the “Bologna countries” – might prove to be precisely the kind of information that is most attractive for students, and for which there is the greatest demand.

We have already described aspects of the “European ranking market trend. The French Minister of Education was expected to raise the question of a European ranking or classification of higher education at a meeting of EU ministers of education in November 2008. Furthermore, the German CHE has extended its ranking activities to include Austrian, Belgian and Swiss universities, and has also carried out test operations in the Netherlands. The CHE has also implemented rankings of “excellent” Master’s programmes in Europe in the natural science and medical areas which, it is hoped, will also be extended into other higher education fields. There are also strong indications that the European Commission’s Directorate-General for Education and Culture (DG EAC) is preparing for the announcement of a tender procedure for the development of a multidimensional ranking framework with global coverage.225

There are good grounds for monitoring developments in Europe in other areas too. Valuable information about the supply of, and to some extent also the quality of, European higher education may also be developed as a result of several other ongoing projects, involving, for example the classification of higher education and ECTS accreditation – both these projects are conducted under the auspices of the European Commission.226 The Directorate-General for Research and Development (DG RTD), in cooperation with Eurostat and DG EAC, has also published a report which aims to establish a statistical information system covering extensive data on European universities. A pilot study is to investigate the possibilities of collecting comparable data.227

Final conclusions

To the extent that students actually want to have greater access to information about higher education in Sweden, the aim of finding methods for meeting this demand in the form of comprehensive and relevant information concerning quality differences in Swedish higher education must be regarded as admirable. The question of whether this is best achieved by means of ranking may be raised, however – at least as regards the type of rankings that we have surveyed in this report. On the other hand, multidimensional and interactive student information systems, not necessarily in the form of conventional rankings, are an attractive possibility from many viewpoints, since they permit individual definitions of quality, and also comparisons between the institu-

tions and programmes that are relevant for each individual user. However, the selection of the available indicators clearly constitutes a restriction on the scope of individual choices.

Furthermore, not all aspects of higher education can be described in terms that can simply be boiled down to a single indicator. As a result, the question of the extent to which students actually require simplified information may be raised. There is a considerable volume of information about higher education in Sweden – the Swedish National Agency for Higher Education issues a considerable number of publications in the form of reports, analyses, evaluations and official statistics that provide relevant information for students. There is also a rich supply of information supplied by the institutions themselves, research councils, student bodies and other entities in the higher education sector. Doubts may be expressed, of course, about the accessibility and transparency of some of this information, but these issues should be raised in a discussion with the potential target group – that is to say the students.

There is a considerably greater lack of comprehensive, relevant, reliable, transparent and comparable information concerning international – and also European – higher education. The lack of information at a European level – based on the Bologna process’ goal of mobility within Europe in higher education – would appear to be a greater problem for students than any possible lack of information in a Swedish context.
References


Akademi eller verklighet?, (2008), Svenskt Näringsliv.


Askling, Berit, (2007) Om sakkunskap och expertis i nationella utvärderingar och kvalitetsbedömningar, Göteborgs universitet.


Clarke, M. (2002), Some Guidelines for Academic Quality Rankings, in *Higher Education in Europe*, vol. 2.4, no. 4.


Fokus (2008), nr. 40.


Harris, Kerri-Lee and James, Richard (2006). The Course Experience Questionnaire, Graduate Destinations Survey and the Learning and Teaching Performance Fund in Australian higher education. Public Policy for Academic Quality.


Högskolerankingen 2008, Handelskammarens rapport nr. 5 2008, Sydsvenska Industri- och Handelskammaren.

Högskoleverkets internationella nyhetsbrev 2006: nr 22.


Kälvemark, T., (2008), International Comparability in the Age of Globalisation – the abuse of rankings and citation statistics in higher education, paper presented at The Higher Education Research Seminar at Linköping University, 13 June.


Marginson, S., (2006) Rethinking and Re-imagining Rankings: Multiple models for world class universities, Association for the Study of Higher
Education Annual Meeting International Forum, Anaheim, California, 1 November 2006.


Marginson, S. & van der Wende, M., (2006), To Rank Or To Be Ranked: The Impact of Global Rankings in Higher Education, Centre for Higher Education Policy Studies, University of Twente.

Meddelanden i samordningsfrågor för Sveriges officiella statistik 2001:1, Örebro: SCB.

Medical Education 1984;18:284–297


Moderna Tider, (1999), nr 102, årgång 10.

Moderna Tider (2000), nr 113, årgång 11.

Moderna Tider (2001), nr 125, årgång 12.

Moderna Tider (2002), nr 137, årgång 13.


Proposition 2008/09:50 Ett lyft för forskning och innovation.


Quality for the Future RQ08, Lunds universitet 2008.


UF19 SM0801 *Utländsk bakgrund för studerande i grundutbildning och forskarutbildning 2006/07*, SCB och Högskoleverket: Sveriges officiella statistik, statistiska meddelanden.

UF20 SM0801 *Universitet och högskolor. Studenter och examina i grundutbildning 2006/07*, SCB och Högskoleverket: Sveriges officiella statistik, statistiska meddelanden.

Underlagsrapport nr 10 till Globaliseringsrådet, (2008), *Globaliseringen och den högre utbildningen*.


Universitetsläraren (2008), nr.16.

*University World News*, 1/6 2008.


van Dyke, N. (2009), Twenty Years of University Report Cards, i *Higher Education*, 49 (4).


