Chapter 9
Measuring Teaching Performance

Keith Trigwell

9.1 Global University Ranking Systems

This chapter is concerned with the place of teaching and learning in university rankings – both global and national. As a key component of core university business, teaching and learning should be a key component of any ranking exercise, but this is mostly not so. For example, of the two principal global university rankings, only the QS Times Higher Education 'World University Rankings' includes an attempt (and a mediocre one at that) to measure aspects of teaching and learning. The 'Academic Ranking of World Universities' compiled by the Shanghai Jiao Tong University Institute of Higher Education (SJTI) group focuses mainly on research performance. At the national level, the inclusion of indicators of teaching and learning in ranking calculations are more common, but the focus is still heavily on research indicators.

Such a focus is understandable, for two reasons, as Liu and Cheng (2005) suggest. First, international research indicators are well-established and are commonly accepted. Second, agreeing on, and implementing indicators of teaching performance tends to be difficult 'owing to the huge differences between universities and the large variety of countries, and because of the technical difficulties inherent in obtaining internationally comparable data' (Liu and Cheng 2005: 133). Similar conclusions are reached by other commentators. A repeated difficulty is that no ranking or quality-assessment system has been able to generate data based on measures of the 'value added' during the educational process; so, few focus on teaching and learning at all (Dill and Soo 2005: 503–505). As Altbach (2006) states, 'There are, in fact, no widely accepted methods for measuring teaching quality, and assessing the impact of education on students is so far an unexplored area as well' (Marginson and van der Wende 2007: 2; Guarino et al. 2005: 149).

In the QS Times Higher Education ranking, a high value is placed on institutional reputation and on the level of 'internationalisation' of Higher Education

K. Trigwell (✉)
Institute for Teaching and Learning, University of Sydney, Sydney, Australia
e-mail: Keith.trigwell@sydney.edu.au

J.C. Shin et al. (eds.), *University Rankings*, The Changing Academy – The Changing Academic Profession in International Comparative Perspective 3,
DOI 10.1007/978-94-007-1116-7_9, © Springer Science+Business Media B.V. 2011

Institutions (HEIs), and in the outcome, the rankings tend to favour HEIs with a strong presence in the degree market. A total of 40% of the Times index is composed of an opinion survey of academics around the world, and another 10% of a survey of global employers. There are two internationalisation indicators: the proportion of students who are international (5%) and the proportion of staff who are international (5%). Another 20% is determined by the student–staff ratio which is treated as a proxy for teaching 'quality'. The remaining 20% of the Times index is composed of research citations per staff member using the Thomson and ISI database (The Times Higher Education 2010).

The ratio of students to teaching staff (student–staff ratio) is probably the most accessible and widely used measure of teaching in the world. Research into student learning shows that there is a significant negative correlation between class size and quality of learning, and as noted in the QS Times HES literature, it does at least provide an indication that the institution in question has sufficient staff to teach its students (Sowter 2009). But there would be few people, if any, who would consider that the quality of university teaching (and learning) can be assessed using this measure alone. Of the remaining QS Times HES indicators, few have any clear connection to teaching quality. Reputation, as assessed by academic peers, is likely to be heavily weighted towards research, and international student demand is also mostly about research reputation.

Given that teaching and learning is not a key element of the global rankings of either international system, it is difficult to understand the claim made by both that potential students and parents are their key users (Stromquist 2007).

9.2 National Ranking Systems

In a recent analysis of the impact of ranking on institutional decision making, the Institute for Higher Education Policy in the USA noted that:

> At least 40 different nations now have some form of rankings that are regularly published, with more added each year. The criteria vary significantly; the best known and most influential ranking schemes evaluate institutions primarily on the basis of academic factors, but some ranking systems use the number of new patents acquired, climate for female students, environmentally friendly practices, and a host of other characteristics. (Institute for Higher Education Policy 2009: 5)

In the USA, the journal *US News and World Reports* started the annual publication of 'America's Best Colleges' in 1983, and other countries also quickly established their own national rankings using their own measures. For example, national rankings are now found in universities in Germany (Federkeil 2002); Australia and New Zealand (Clarke 2005); China (Liu and Liu 2005); Japan (Yonezawa et al. 2002); Spain (De Miguel et al. 2005); Latin-America (Ledesma 2002); Britain (Eccles 2002); Poland (Perspektywy 2002); and Russia (Filinov and Ruchkina 2002).

Among the academic factors other than research that feature in these systems are, rating of teachers by students, students' study times, accessibility of teachers,

availability of class discussion, library, study abroad, best classroom experience (The Princeton Review 2010), and students' perceptions of teaching quality, attainment of generic skills and overall satisfaction with the quality of their programmes, as well as academic progress, retention and employment/study following graduation (Department of Education, Science and Training, Australian Government 2004). The seven categories and the corresponding percentage weightings used by the US News and World Report are: peer assessment (25%); retention (20% in national universities and liberal arts colleges and 25% in master's and baccalaureate colleges); faculty resources (20%); student selectivity (15%); financial resources (10%); graduation rate performance (5%; only in national universities and liberal arts colleges) and alumni giving rate (5%) (Morse and Flemingan 2009).

While these examples illustrate the range of indicators that are being used nationally, the question remains as to whether they are indicators of what is now recognised as good teaching (or learning) and whether they capture variation in the quality of education provided or achieved. This question is addressed later in the chapter. The focus of the next few sections is on the research that has led to the practices described above.

9.3 Is Research a Proxy for Teaching Quality?

It has been argued by some (Benowski 1991; Taylor 2003; Yair 2008) that because there is a strong link between university research and teaching, that universities that are strong in research will be strong in teaching, and those indicators of research strength are also indicators of strength in teaching. From this perspective, the teaching–research relationship is bidirectional, and one can influence and enhance the other. It is underpinned by the belief that by researching the subject matters they teach, academic staff can gain and maintain passion, which is a prerequisite for excellence in teaching. In addition, the practice of teaching can foster academics to gain a deeper understanding of the field, and absorb novel ideas from their students, rendering them better researchers (Taylor 2003).

Contrary to the view above, other scholars contend that teaching excellence and research excellence are 'in direct conflict, compete for academic attention and only one of them is rewarded' (Gibbs 1995: 149). Gibbs (1995) quotes some empirical studies on teaching and research relations to corroborate his claims. For instance, a survey of Higher Education Institutions in the UK conducted by the Oxford Centre for Staff Development found that only 12% of promotion decisions were made solely on the basis of teaching excellence; and 38% of institutions made promotion decisions on grounds that did not include quality of teaching. The conflicting relation between teaching and research is also reported in other countries (e.g. Astin and Chang 1995) and in specific fields of study (e.g. Patrick and Stanley 1998). Astin and Chang (1995), using a survey method, discovered that no institutions in the USA (in their sample) have a balanced orientation towards both students (learning) and research.

Patrick and Stanley (1998) failed to show any connections between high quality of teaching and excellence of research in business fields of study. Based on a meta-analysis of 58 articles, Hattie and Marsh (1996) had earlier reported a near zero (0.06) correlation between indicators of teaching and research, using various measures of teaching quality (e.g. student evaluation, peer ratings) and of research output (e.g. citations, productivity, quality). And in a later paper, Marsh and Hattie (2002) provide evidence confirming the earlier results, noting a correlation of only 0.02 between global teaching and global research factors. These and other studies have led Dill and Soo (2005: 507) to conclude that 'empirical research … suggests that the correlation between research productivity and undergraduate instruction is very small and teaching and research appear to be more or less independent activities'.

This brief review suggests that there is certainly insufficient evidence available to support the idea that research quality can be used as a proxy for teaching quality. So, what approaches are available to measure teaching quality?

9.4 Measurements of Quality Teaching

Although assessing quality teaching remains a contentious issue among educational researchers, it is agreed upon that the choice of indicators to measure teaching quality is crucial, and it is essential to select valid indicators rather than just practical ones.

Brooks (2005) has proposed the following choice of indicators for teaching:

1. Using student questionnaires
   Employing student evaluation of teaching seems logical since teaching quality is essentially student-centred and aims at improving student learning. However, many teachers perceive student evaluation is more or less biased, thus they give little credence to the responses obtained from students.
2. Peer in-class evaluations
   Peer in-class evaluations has an obvious advantage in that it focuses more on teaching process rather than outcome, yet it may hinder teaching innovations under the influence of conservative judgement held by peer evaluators. It is also only ever an observation of a small part of teaching.
3. Using teaching portfolios
   Using teaching portfolios to indicate quality teaching seems fair, yet it is not easy to operationalise as each field has their own practice, making comparison across fields effectively impossible.

Since teaching is a means to the student learning end, an alternate approach is to focus on the students' programmes, their experience and outcome. Measurements of these areas are seen by Brooks to be comprised of four main categories, namely, programme characteristics, programme effectiveness, student satisfaction and student outcomes.

1. Programme characteristics include counts of the degrees awarded, amount of student financial support available and quality of students entering the university.
2. Programme effectiveness is a measure of the timelines of completion of students.
3. Student satisfaction is often measured by national survey or by examining students' engagement and career preparation.
4. Student outcomes are often operationalised by either determining students' career path and job satisfaction or their generic skills, such as critical thing, analytical reasoning and written communication (Brooks 2005).

The combination of the teaching indicators used in national rankings with these seven areas constitutes a comprehensive array of what might be the basis of a teaching index. But how close do these elements or even the full array, match what is now known about high quality university teaching?

9.5 What Is High Quality Teaching?

A common explanation as to why there are so few indicators of good teaching in global ranking systems is that high quality teaching (and learning) is too difficult to define and to measure. A part of this explanation is no longer valid. It is now quite clear what good teaching is, and what constitutes poor teaching. International research studies consistently point to a set of factors associated with good teaching (and learning) and these are addressed in more detail below. However, there are still scholars who hold the belief that *quality* teaching can never be fully understood since teaching itself is a never-ending process (Argyris and Schön 1974; Hau 1996). For instance, Hau (1996) contends that in order to ensure quality teaching in higher education, a never-ending process should be undergone to eliminate defects in quality.

As Biggs (2001) suggested, 'quality' is unarguably a multi-layered concept, which can be understood as an outcome, a property, or a process. Quality teaching also encompasses multiple meaning. The contentiousness of 'quality' also stems from its 'stakeholder-relative' characteristics, as suggested by Tam (2001), who proposes that all stakeholders hold their own perspective on what is meant by quality in education. This idea is also supported by Harvey and his associates, who agree that different ways of defining quality in higher education results from different stakeholders in higher education (Harvey et al. 1992). According to Harvey et al. (1992), 'stakeholders' include students, teaching and non-teaching staff, government and funding agencies, assessors and even the community.

Harvey and Green (1993) describe four kinds of definitions of quality teaching according to its history of development. Traditionally, quality teaching is a conception closely related to 'excellence', which is still a dominant concept among many tertiary educational institutions. More recently, quality teaching is associated with 'value for money'. Thirdly, quality teaching can also be defined as the efficiency and effectiveness of fitness for purpose, which suggests that a quality educational

provider needs to engage students to acquire knowledge effectively. A final conceptualisation of quality has to do with 'transforming' students' perceptions of applying their knowledge to solve real world problems.

Quality teaching may be hard to define, and excellent teaching may be difficult to achieve, but researchers do generally agree that the following list includes those qualities that are essential to good teaching (Ramsden et al. 1995):

- Good teachers are also good learners; for example, they learn through their own reading, by participating in a variety of professional development activities, by listening to their students, by sharing ideas with their colleagues and by reflecting on classroom interactions and students' achievements. Good teaching is therefore dynamic, reflective and constantly evolving.
- Good teachers display enthusiasm for their subject, and a desire to share it with their students.
- Good teachers recognise the importance of context, and adapt their teaching accordingly; they know how to modify their teaching strategies according to the particular students, subject matter and learning environment.
- Good teachers encourage learning for understanding and are concerned with developing their students' critical thinking skills, problem-solving skills and problem-approach behaviours.
- Good teachers demonstrate an ability to transform and extend knowledge, rather than merely transmitting it; they draw on their knowledge of their subject, their knowledge of their learners and their general pedagogical knowledge to transform the concepts of the discipline into terms that are understandable to their students. In other words, they display what Shulman has termed 'pedagogical content knowledge' (Shulman 1987, 1988).
- Good teachers set clear goals, use valid and appropriate assessment methods and provide high quality feedback to their students.
- Good teachers show respect for their students; they are interested in both their professional and their personal growth, encourage their independence and sustain high expectations of them. (p. 24)

9.5.1 Scholarship of Teaching

It has been argued by Boyer (1990) that the elements of university teaching should be rendered a higher status and that one way to do that is to encompass the 'scholarship of teaching' as one of the four types of scholarships in the academy (teaching, discovery, integration and application). Trigwell et al. (2000) have reviewed the meaning of scholarly teaching and presented it as an achievement in four broad dimensions (such as communication and ways of thinking about teaching). A generic approach to the assessment of these scholarships, including teaching, is described by Glassick et al. (1997) in work, building on that of Boyer. Using the concept of scholarship, Glassick and colleagues proposed that academics practising

- Clear Goals - Does the scholar state the basic purposes of his or her work clearly? Does the scholar define objectives that are realistic and achievable? Does the scholar identify important questions in the field?
- Adequate Preparation - Does the scholar show an understanding of existing scholarship in the field? Does the scholar bring the necessary skills to her or his work? Does the scholar bring together the resources necessary to move the project forward?
- Appropriate Methods - Does the scholar use methods appropriate to the goals? Does the scholar apply effectively the methods selected? Does the scholar modify procedures in response to changing circumstances?
- Significant Results - Does the scholar achieve the goals? Does the scholar's work add consequentially to the field? Does the scholar's work open additional areas for further exploration?
- Effective Presentation - Does the scholar use a suitable style and effective organisation to present his or her work? Does the scholar use appropriate forums for communicating work to its intended audiences? Does the scholar present her or his message with clarity and integrity?
- Reflective Critique - Does the scholar critically evaluate his or her own work? Does the scholar bring an appropriate breadth of evidence to her or his critique? Does the scholar use evaluation to improve the quality of future work?

(Glassick *et al*., 1997: 36).

Fig. 9.1 The Carnegie Foundation's six standards of scholarship

the 'scholarship of teaching' experience a symmetry between teaching and research through the common elements of clear goals, adequate preparation, appropriate methods, significant results, effective publication and reflective critique (Glassick et al. 1997). If these criteria are used to assess the quality of the scholarship of discovery (research), can they also be used to assess the quality of the scholarship of teaching? And if they can, is there a relation between scholarship of teaching and good teaching or high quality teaching? The symmetry between the two forms of scholarship is illustrated in Fig. 9.1.

9.5.2 Qualitative Variation in University Teaching

It is apparent from the above analysis that teaching involves much more than what happens in a classroom or on-line; it is oriented towards, and is related to, high quality student learning, and includes planning, compatibility with the context, content knowledge, being a learner and above all, a certain way of thinking about teaching and learning. It is also no longer a solo pursuit. Not only do the planning elements apply to individual subjects, they must also be seen to be part of a whole programme. And there is an alignment between teachers' teaching and student learning. Good teaching includes all these elements.

Higher education research studies suggest that there is a way of conceiving university teaching which is more strongly associated with higher quality student learning than other ways of thinking (Prosser and Trigwell 1999). Some teachers keep more of a focus on their students in their planning and their activities. These teachers tend to be teaching students who describe using a higher quality approach to their learning. Teachers adopting this approach see their role as helping their students develop and change their conceptions or world views. As a result of this thinking, their focus is on the bigger picture – an overview of the topic or how the components of the information are related to each other, and on students' prior knowledge – what students bring to the situation. Their planning and teaching methods are in alignment with this conception.

This thinking is in contrast to that of teachers who work with a view where the focus is on what they do as teachers, or on the detail – individual concepts in the syllabus or textbook, or the teachers' own knowledge structure – without acknowledgment of what students may bring to the situation or experience in the situation. They see their role as transmitting information based upon that knowledge to their students.

With respect to the concept of alignment, a teacher who holds the former conception is more likely to adopt an approach which has the student as the focus of activities. It matters more to this teacher what the student is doing and learning and experiencing than what the teacher is doing or covering. This teacher is one who encourages self-directed learning, who makes time (in 'formal teaching' time) for students to interact and to discuss the problems they encounter, who assesses to reveal transformed knowledge (not only to judge and rank students), who provokes debate (and raises and addresses the taken-for-granted issues) and who uses a lot of 'lecture' time to question students' ideas, and to develop a 'conversation' with the students.

These strategies may differ from those used by a teacher with a teacher-focused approach, but this is not always so. For example, two teachers can use the same strategy (say, buzz groups during a lecture – a buzz group is a short discussion between a small group of students, initiated by the teacher). It is the teachers' intention (aligned with their conception) that constitutes the main difference in this case. Using a student-focused approach, a teacher may see the buzz groups as a means by which students can compare their understandings of the lecture topic, and give feedback to the teacher on that understanding. In a teacher-focused approach, the teacher may see buzz groups as a way of giving her or himself a break from talking, and students a break from note-taking in a one-hour lecture. The differences in student learning, from the use of the same strategy, may be substantial.

Much of the good teaching literature, and some of the more common forms of evaluating teaching, are focused on teaching strategies, such as the clarity of explanations, or the availability of teachers for consultations. The conception-based research described above suggests that unless the teacher is using a student-focused, conceptual development conception, the emphasis on strategies may be misplaced. For example, on advice about using an online teaching strategy for a component of a course, from a student-focused, conceptual development conception

a teacher might ask the following two questions: (a) Is this strategy likely to achieve the student learning aims? and (b) What type of learning is likely to be encouraged using this strategy? From a teacher-focused perspective, the questions asked are more likely to include: (a) Is this strategy likely to be the most efficient method of dissemination? and (b) What amount of coverage is likely to be achieved using this approach?

This variation in thinking or conceptualising is a key element in variation in teaching quality (and student learning). Teaching indicators need to focus beyond how well the teacher is conducting teaching activities, and how well those activities are received by students. There is a need to also consider the nature of those activities and how they align with variation in student learning. Activities that are student-focused are more likely to align with higher quality outcomes of learning.

This analysis of teaching and learning gives some insight into the gap that exists between the ways teaching/learning is assessed, the ways it is proposed to be assessed and what researchers agree as the characteristics of good teaching/learning.

9.6 Surveys of Student Experience

Some of these elements of good teaching have been incorporated into surveys of the student learning experience. Students' descriptions of their learning experience have also been used as a proxy for learning quality, for example, in the National Student Survey (NSS) used in the UK and the Course Experience Questionnaire (CEQ) in Australia. The research underpinning the Australian Course Experience Questionnaire (Ramsden 1991) shows that there is a significant positive correlation between the quality of the outcomes of student learning and their experience of the teaching they receive, of the appropriateness of their workload, of the nature of their assessment and of the clarity of what is expected of them (Lizzio et al. 2002).

The CEQ, as currently used as a national indicator of learning experience, contains 13 items – six that make up a scale related to students' experience of the quality of teaching in their course, six on their acquisition of a range of generic graduate skills and one item on their overall satisfaction with the quality of the course they have just completed. In a symposium designed to explore the characteristics of courses that had high scores on the Australian CEQ Good Teaching Scale, most of the representatives of the Business/Commerce/Law courses, selected to present their information, were not sure why they received the positive response from students that they did. There was enormous variation in the contexts they described. The contexts included distance learning, large undergraduate courses, small postgraduate courses, and in the disciplines of accounting, marketing, business administration, management, economics and law. At first sight, there appeared to be little that the programmes as described had in common, and that could account for the common high quality experience of their students. However, what was

apparent was that the presenters themselves, who in most cases were the course coordinators, had a lot in common. They were teachers who showed that they were concerned about students, no matter what their context.

A common theme in all courses described was closely related to the idea of using the objectives of the students as the design/presentation feature. While the focus of this theme is still on the students, it is on those students who were enrolled and doing the course rather than on those who might enrol in the course. This meant that there was almost always close contact of some sort between student and teacher as perceived by the student, and it is likely that this is perceived as personal attention, even in the larger classes.

Given the large variation in these courses in so many other aspects, this perception of personal attention may be an important factor in explaining the high CEQ rating in the courses featured at the symposium, and in providing support for the use of CEQ scores in rankings. The way the comments were made in the presentations included:

- An investigation of the nature of potential students and their objectives and the use of this information in developing the curriculum.
- Dedicated staff who make time to interact with the students.
- Personal attention given to students during industrial placement.
- An intimate rural context with significant contact between staff and students.
- Small (postgraduate) classes where students felt they had personal contact.
- Use of full-time teaching staff rather than sessional, casual or part-time staff which meant staff supporting the programme were usually more contactable.
- Programmed individual consultation time with students which was based on a formula that increased available time with increasing student numbers.
- Focus on the students, rather than on the teacher and what the teacher does.

In all the cases described above, the students received more attention than they would in a 'normal' programme. While it was not all the same sort of attention (not all classes were small or had dedicated staff), it may have been what students wanted, and it may have matched the needs of the students in the course. It is likely that what came across to students was care and support in their context. The variations in context are illustrated by extremes of a student-based approach to care and support in developing student learning independence in one university versus offering care and support in negotiating student demands for a teacher-based approach to teaching in another university.

Conclusions which can be drawn from these observations are:

1. Good teaching in any one university may be achieved when students feel they have received personal attention from the staff managing their learning programme. This personal attention may come in a variety of forms depending on the context, and it may be more important to students than issues such as class size, coherent curricula and so on.
2. High CEQ scores result from a match between course and context. Using CEQ results (or any ranking system) in advertising to attract students to the course may lead to the recruitment of students who are not those who would normally

enrol and may therefore not fit well in the existing course. Given the diversity of university missions, this one observation may be sufficient grounds to argue that global rankings based on teaching are inappropriate.

In 2005, the NSS was introduced to English universities on a voluntary basis, to be completed by undergraduate students in their final year. Since then most universities have administered the survey annually. The stated aim of the survey 'is to gather feedback on the quality of students' courses, to help inform the choices of future applicants to higher education, and to contribute to public accountability' (Higher Education Academy 2010a). The survey consists of 22 items in six broad educational areas, including teaching, assessment and feedback, academic support, organisation and management, learning resources and personal development. There is also an 'overall satisfaction' question about how satisfied the respondent is with the quality of their course (Surridge 2008). In the USA, The National Survey of Student Engagement (NSSE), first launched in 1991 and distributed annually to four-year college and university undergraduate students, asks them about 'participation in programs and activities that institutions provide for their learning and personal development' (Kuh 2001; National Survey of Student Engagement 2010a). The NSSE is comprised of 42 questions measuring students' behaviours and institutional features, which can reflect students learning and personal development. The benchmarks used in the NSSE include level of academic challenge, active and collaborative learning, student–faculty interaction, enriching educational experiences and supportive campus environment (National Survey of Student Engagement 2010b). It is currently being used in voluntary (and private) ranking exercises in the USA and Canada, and interest in an Australian version is growing (Australian Council for Education Research 2010).

Brooks (2005) proposes student surveys as an important element in choice of teaching quality indicators. Of course, there are as yet no surveys of this type that are global, and even if there were, they may prove to have the same two limitations that Marsh and Cheng (2008) describe in their report on the English National Student Survey. First, that there is not substantial variation between different universities in terms of overall satisfaction; and differences between universities explain only about 2.5% of the variance (variance component based on multilevel analyses, controlling for student characteristics and discipline). However, because the number of students from each institution is so large, the differences between institutions are quite reliable (for all but a handful of universities with small numbers of respondents). Second, that there is substantially more variance explained in overall satisfaction by differences between discipline groupings within universities than by differences between universities.

Results from surveys such as the Australian Course Experience Questionnaire, the English National Student Survey and the National Survey of Student Engagement in the USA all provide opportunities for the collection of nationwide teaching/learning related data that could potentially be used in national ranking systems, but the difficulties in developing even national teaching indicators are illustrated in the following Australian case study (Prosser M., 2009, Personal communication).

For over 5 years in Australia, the Course Experience Questionnaire has been used as the principal component of a teaching/learning performance-based funding system. In April 2004, the Australian Government released its discussion paper on its Learning and Teaching Performance Fund (Department of Education, Science and Training, Australian Government 2004). It canvassed various ways of assessing excellence based upon performance indicators, peer review and value added approaches. In the end, it adopted a two-stage process.

In stage 1, institutions had to meet a number of more qualitative criteria to be eligible for the second stage. These criteria included such things as evidence of systematic support for improving teaching and learning. Having met these conditions, institutions were deemed eligible for inclusion in stage 2, which was based upon a rank ordering of institutions on a set of quantitative learning and teaching performance indicators. In 2008, the performance indicators for the scheme were: Student Progress Rate, Student Retention Rate, Graduate Full-Time Employment, Graduate Further Full-Time and Part-Time Study, Graduate Satisfaction with Generic Skills, Graduate Satisfaction with Good Teaching and Graduate Overall Satisfaction. A sophisticated statistical process for adjusting these indicators to take account of institutional variation is outlined in the paper. The adjustment factors included: Mix of disciplines, Admission basis criteria, Type of attendance, Mode of attendance, Indigenous status, Gender, Disability, Non-English Speaking Background, Age, Location, Socio-Economic Status, Unemployment rate, Number of students, Level of study and New to higher education. Within each broad field of study, institutions are placed in bands based upon the combination of indicators, and funding is allocated to institutions. Even with the availability of a standardised, evidence-based, effectively compulsory, student experience survey (Course Experience Questionnaire) the establishment and maintenance of a national teaching comparator has proved difficult and controversial. The scheme is not being continued from 2010.

9.7 High-Quality Learning Outcomes

Given the problems with surveys of student experience, it is of little surprise that the focus is shifting to establishing common national and international ground in the outcomes of student learning. Since learning is the main reason why teaching is carried out, an assessment of the quality of learning is also a valid component of ranking systems. But similar quality questions remain – what is high quality learning? In a major UK study on student engagement and high quality learning, the Enhancing Teaching-Learning (ETL) Environments research team describes desired student outcomes more in terms of ways of thinking and practising in the discipline or professional area (Hounsell et al. 2005). Ways of thinking and practising ‘capture the richness, depth and breadth of what the students could learn through engagement with a given discipline or subject area. It rests on a secure foundation of subject knowledge and understanding and could also encompass

subject-related skills, conventions and practices for communicating within the subject, and salient values and attitudes' (p. 5).

To achieve what Hounsell and colleagues suggest is likely to be extremely difficult, but the struggle has begun. Marginson and van der Wende (2007) describe the developments as follows: 'Based on these trends in relation to the job market, there are moves to apply the terminology for normal education to higher education. One example is the standardisation of higher education in curriculum, testing, credit hours and degree awarding requirements. Through the Bologna Process, European countries are moving towards a common structure for bachelors, master and doctoral degrees, with similar achievement criteria and credit hours systems. In some countries (e.g. the UK), benchmarks for each discipline area have been set within quality assurance systems. Based on the notions of standardisation, the OECD is developing standardised tests for college graduates similar to those for secondary school students (the so-called 'PISA'). If standardisation is widely applied, the shape of higher education will be totally different from what it is now. Curricula will be standardised in each discipline area, standardised tests will be given to most college students, each course will carry the same credit hours, and degree requirements will be similar across national borders' (p. 319).

Similar developments are happening in the USA. For example, the Middle States Commission on Higher Education has a long-standing commitment to outcomes assessment and to student learning outcomes in particular. In their publication *Characteristics of Excellence in Higher Education* (Middle States Commission on Higher Education 2006), excellence is defined as meeting 14 accreditation standards. Seven of the standards are related to educational effectiveness (Admissions and Retention; Student Support Services; Faculty; Educational Offerings; General Education; Related Educational Activities (Basic Skills; Certificate Programmes; Experiential Learning; Non-Credit Offerings; Branch Campuses, Additional Locations, and Other Instructional Sites; Distance or Distributed Learning; Contractual Relationships and Affiliated Providers) and Assessment of Student Learning). Each of the standards is defined in terms of outcomes. For example, for Student Admissions and Retention outcomes are: Statements of expected student learning outcomes and information on institution-wide assessment results, as appropriate to the programme offered, available to prospective students (Prosser M., 2009, Personal communication).

But here too, there will be difficulty as the focus is likely to be on achievement of a minimum or threshold level, and such an approach cannot therefore be used to assess the degree of 'value added'.

9.8 Conclusions

The analysis described in this chapter does not provide any clear guidance for national ranking systems based on university teaching that could be of use to inform potential students, or the development of global indicators of teaching

quality. Part of the problem could be that the search for the answers is focused mainly on the easier options rather than looking at what might be needed. Even a cursory glance at the way the research indices are measured (and accepted) reveals the enormous amount of international peer review resource that is expended in the process. It is highly likely that the same international peer review framework is needed for teaching. At a national level, this has been growing in some countries for several decades, for example in the allocation of national teaching grants in the 1990s (Australian Committee for the Advancement of University Teaching and the Swedish Council for the Renewal of Undergraduate Education) and in teaching awards (e.g. National Teaching Fellowship Scheme (Higher Education Academy 2010b)). The scholarship of teaching movement (ISSOTL 2010) is also encouraging international peer review of scholarly teaching artefacts. Despite this progress, descriptions of international standards and universal acceptance of them are still many years away.

If and when international teaching/learning standards are developed, their use in rankings may still be seen to be undesirable. As Teichler (2008) has commented, ranking reports guide vertical classification of HEIs that is likely to have adverse effects on institutional diversity. In addition, ranking indicators are overly simplified and do not necessarily reflect the quality of education. Instead, international ranking surveys (e.g. Shanghai Jiao Tong's Ranking) put major emphasis on research performance instead of teaching.

From the perspective of preparing acceptable global ranking systems that include a prominent place for teaching, there are clearly many obstacles to overcome, but from the point of view of enhancing the quality of student learning, attempts at ranking systems may not be such a bad idea. 'The savviest universities are using the rankings phenomenon to catalyse and establish keener performance evaluation internally at a faculty, department and individual staff member level. Driving it down to this level can help build actionable metrics as opposed to abstract statistics and this can lead to a university being able to revolutionise their performance in education and research, and in time, as a side-effect rather than an objective, improve their performance in rankings' (Sowter 2009).

An example of what Sowter has described exists at the University of Sydney. Nationally, the Course Experience Questionnaire is used in Australia to capture the comparative experience of university graduates. The release of data, at times as much as 6 years after students have begun their first year of study, is not sufficiently current to support change (some courses will have gone through one and sometimes two revisions in that period). To facilitate the improvement process, the University of Sydney introduced a Student Course Experience Questionnaire (SCEQ) that is completed by students in all years of the programme on their experience of their course to that point. This provides the 'bigger picture' data that is up to date, and is a measure of the quality of the course at that time. The SCEQ includes the same scales as the CEQ (Good Teaching and Generic Skills) but also scales on Clear Goals and Standards, Appropriate Assessment and Appropriate Workload. To enable change to be addressed at the level of the individual teacher, a third level of questionnaire (Unit of Study Evaluation, USE) was designed with 12 questions

including one on each of the five scales of the SCEQ. As such, an individual teacher has a 'direct line of sight' from the USE to the CEQ. Relations of this sort are an essential criterion of any national indicator of university teaching.

There is additional evidence that teaching indicators can generate positive change. The Institute for Higher Education Policy notes that the institutions included in their case studies '… continue to point to their changing practices that alter input indicators – increasing selectivity, favouring research over teaching, and strengthening the faculty profile – [while] a number of institutions are also reporting changes to practices directly related to student learning and success. Institutions that use their rankings to prompt change in areas that directly improve student learning experiences demonstrate that rankings can lead to positive change in teaching and learning practices' (Institute for Higher Education Policy 2009: 3).

References

Altbach, P. (2006). The dilemmas of ranking. *International Higher Education, 42*, 2–3.

Argyris, C., & Schön, D. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey Bass.

Astin, A., & Chang, M. J. (1995). Colleges that emphasize research and teaching. *Change, 27*(5), 44–49.

Australian Council for Education Research. (2010). Australasian Survey of Student Engagement. Retrieved June 23, 2010, from http://www.acer.edu.au/documents/AUSSE_Brochure.pdf

Benowski, K. (1991). Restoring the pillars of higher education. *Quality Progress, 24*(10), 37–42.

Biggs, J. (2001). The reflective institution: Assuring and enhancing the quality of teaching and learning. *Higher Education, 41*(3), 221–238.

Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton: Carnegie Foundation for the Scholarship of Teaching.

Brooks, R. L. (2005). Measuring university quality. *The Review of Higher Education, 29*(1), 1–21.

Clarke, M. (2005). Quality assessment lessons from Australia and New Zealand. *Higher Education in Europe, 30*, 183–197.

Department of Education, Science and Training, Australian Government. (2004). *Learning and teaching performance fund: Issues paper*. Retrieved June 23, 2010, from http://www.dest.gov.au/NR/rdonlyres/8F857836-F4AE-4D74-952E-3906FF5807B4/10254/ltpf_issues_2004.pdf

De Miguel, J. M., Vaquera, E., & Sanchez, J. D. (2005). Spanish universities and the ranking 2005 initiative. *Higher Education in Europe, 30*, 199–215.

Dill, D. D., & Soo, M. (2005). Academic quality, league tables, and public policy: A cross-national analysis of university ranking systems. *Higher Education, 49*, 495–533.

Eccles, C. (2002). The use of university rankings in the United Kingdom. *Higher Education in Europe, 27*, 423–432.

Federkeil, G. (2002). Some aspects of ranking methodology – The CHE-ranking of German universities. *Higher Education in Europe, 27*, 389–397.

Filinov, N. B., & Ruchkina, S. (2002). The ranking of higher education institutions in Russia: Some methodological problems. *Higher Education in Europe, 27*, 407–421.

Gibbs, G. (1995). The relationship between quality in research and quality in teaching. *Quality in Higher Education, 1*(2), 147–157.

Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the professoriate*. San Francisco: Jossey-Bass.

Guarino, C., Ridgeway, G., Chun, M., & Buddin, R. (2005). Latent variable analysis: A new approach to university ranking. *Higher Education in Europe, 30*(2), 147–165.

Hattie, J., & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. *Review of Educational Research, 66*, 507–542.
Hau, H. (1996). Teaching quality improvement by quality improvement in teaching. *Quality Engineering, 9*(1), 77–94.
Harvey, L., Burrows, A., & Green, D. (1992). *Criteria of quality in higher education report of the QHE project*. Birmingham: The University of Central England.
Harvey, L., & Green, D. (1993). Defining quality. *Assessment and Evaluation in Higher Education, 18*, 8–35.
Higher Education Academy. (2010a). *National student survey*. Retrieved June 23, 2010, from http://www.hefce.ac.uk/learning/nss/
Higher Education Academy. (2010b). *National teaching fellowship scheme projects*. Retrieved June 23, 2010, from http://www.heacademy.ac.uk/ourwork/professional/ntfs/projects
Hounsell, D., Entwistle, N., Anderson, C., Bromage, A., Day, K., Hounsell, J., Land, R., Litjens, J., McCune, V., Meyer, E., Reimann, N., & Xu, R. (2005). Enhancing teaching-learning environments. *Final report to the Economic and Social Research Council on TLRP project L13925109*. Retrieved June 23, 2010, from http://www.etl.tla.ed.ac.uk/docs/ETLfinalreport.pdf
Institute for Higher Education Policy. (2009). *Impact of college rankings on institutional decision making: Four country case studies*. Retrieved June 23, 2010, from http://www.ihep.org/assets/files/publications/g-l/ImpactofCollegeRankings.pdf
ISSOTL. (2010). *International Society for the Scholarship of Teaching and Learning*. Retrieved 23 June, 2010, from http://www.issotl.org/
Kuh, G. D. (2001). *The national survey of student engagement: Conceptual framework and overview of psychometric properties*. Bloomington: Indiana Center for Postsecondary Research and Planning.
Ledesma, J. R. (2002). Latin America: A different reality. *Higher Education in Europe, 27*, 467–474.
Liu, N., & Cheng, Y. (2005). The academic ranking of world universities. *Higher Education in Europe, 30*(2), 127–136.
Liu, N., & Liu, L. (2005). University rankings in China. *Higher Education in Europe, 30*, 217–227.
Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27*, 27–52.
Marsh, H. W., & Cheng, J. H. S. (2008). *National student survey of teaching in UK universities: Dimensionality, multilevel structure, and differentiation at the level of university and discipline: Preliminary results*. Retrieved June 23, 2010, from http://hlst.ltsn.ac.uk/assets/York/documents/ourwork/research/surveys/nss/NSS_herb_marsh-28.08.08.pdf
Marsh, H. W., & Hattie, J. (2002). The relation between research productivity and teaching effectiveness: Complementary, antagonistic, or independent constructs? *The Journal of Higher Education, 73*(3), 603–641.
Marginson, S., & van der Wende, M. (2007). To rank or to be ranked? The impact of global rankings in higher education. *Journal of Studies in International Education, 11*(3), 306–339.
Middle States Commission on Higher Education. (2006). *Characteristics of excellence in higher education: Eligibility requirements and standards for accreditation*. Retrieved June 23, 2010, from http://www.msche.org/publications/CHX06060320124919.pdf
Morse, R., & Flemingan, S. (2009). *How we calculate the college rankings*. Retrieved June 23, from http://www.usnews.com/articles/education/best-colleges/2009/08/19/how-we-calculate-the-college-rankings.html?PageNr=1
National Survey of Student Engagement. (2010a). *About the national survey of student engagement*. Retrieved June 23, 2010, from http://nsse.iub.edu/html/about.cfm
National Survey of Student Engagement. (2010b). *Benchmarks of effective educational practice*. Retrieved June 23, 2010, from http://nsse.iub.edu/pdf/nsse_benchmarks.pdf
Patrick, W., & Stanley, E. (1998). Teaching and research quality indicators and the shaping of higher education. *Research in Higher Education, 39*(1), 19–41.

Perspektywy, W. S. (2002). Ten years of rankings. *Higher Education in Europe, 27*, 399–406.
Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching: The experience in higher education*. Buckingham: Open University Press.
Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*, 129–150.
Ramsden, P., Margetson, D., Martin, E., & Clarke, S. (1995). *Recognising and rewarding good teaching in Australian higher education*. Canberra: Committee for the Advancement of University Teaching, Australian Government Publishing Service.
Shulman, L. S. (1987). Assessment for teaching: An initiative for the profession. *Phi Delta Kappan, 69*(1), 38–44.
Shulman, L. S. (1988). A union of insufficiencies: Strategies for teacher assessment in a period of educational reform. *Educational Leadership, 46*(3), 36–41.
Sowter, B. (2009). *University rankings: There can be no "right answer"*. Retrieved June 23, 2010, from http://qsiu.wordpress.com/2009/04/24/university-rankings-there-can-be-no-right-answer/
Stromquist, N. P. (2007). Internationalization as a response to globalization: Radical shifts in university environments. *Higher Education, 53*, 81–105.
Surridge, P. (2008). The National Student Survey 2005–2007: findings and trends. *A report to the Higher Education Funding Council for England*. Retrieved June 23, 2010, from http://www.hefce.ac.uk/pubs/rdreports/2008/rd12_08/
Tam, M. (2001). Measuring quality and performance in higher education. *Quality in Education, 7*(1), 4–54.
Taylor, M. (2003). Teaching capabilities and professional development and qualifications framework project: stage one. Unpublished Report, Melbourne: RMIT University
Teichler, U. (2008). Diversification? Trends and explanations of the shape and size of higher education. *Higher Education, 56*(3), 349–379.
The Princeton Review. (2010). *Best value colleges methodology*. Retrieved June 23, 2010, from http://www.princetonreview.com/bvc-methodology.aspx
The Times Higher Education (2010). Methodology: A simple overview. Retrieved June 23, 2010, from http://www.topuniversities.com/university-rankings/world-university-rankings/methodology/simple-overview
Trigwell, K., Martin, E., Benjamin, J., & Prosser, M. (2000). Scholarship of teaching: A model. *Higher Education Research and Development, 19*, 155–168.
Yair, G. (2008). Can we administer the scholarship of teaching? Lessons from outstanding professors in higher education. *Higher Education, 55*(4), 447–459.
Yonezawa, A., Nakatsui, I., & Kobayashi, T. (2002). University rankings in Japan. *Higher Education in Europe, 27*, 373–382.