Outcomes of higher education: Quality relevance and impact

What’s the Difference?
A Model for Measuring the Value Added by Higher Education

Hamish Coates, Australian Council for Educational Research (ACER), Australia
WHAT’S THE DIFFERENCE?
A MODEL FOR MEASURING THE VALUE ADDED BY HIGHER EDUCATION

Hamish Coates, Australian Council for Educational Research (ACER), Australia

Measures of student learning are playing an increasingly significant role in determining the quality and productivity of higher education. This paper evaluates approaches for estimating the value added by university education, and proposes a methodology for use by institutions and systems.

The paper argues that value-added measures of learning are important for quality assurance in contemporary higher education. It reviews recent large-scale developments in Australia, methodological considerations pertaining to the measurement and evaluation of student learning, and instruments validated to measure students’ capability, generic skills, specific competencies, work readiness and student engagement.

Four approaches to calculating value-added measures are reviewed. The first approach computes value-added estimates by comparing predicted against actual performance using data from entrance tests and routine course assessments. In the second approach, comparisons are made between outcomes from objective assessments administered to cohorts in the first and later years of study. Comparisons of first-year and later-year students’ engagement in key learning activities provide a third and complementary means of assessing the value added by university study. Feedback on graduate skills provided by employers is a fourth approach which provides an independent perspective on the quality of education.

Reviewing these four approaches provides a basis for their synthesis into a robust and potentially scalable methodology for measuring the value added by higher education. This methodology is advanced, along with its implications for instrumentation, sampling, analysis and reporting. Case studies are presented to illustrate the potential of the methodology for informing comparative analyses of the performance of higher education systems.

An evidence-based perspective on higher education quality

The economic and social role of higher education has expanded and diversified rapidly over the last twenty years. Flowing from this is an increased need from government, university leaders, business, students and the public for more focused information on whether graduates have the capabilities required to engage productively in the global knowledge economy. The perspective driving this paper is that such practice will doubtless grow in an expanding and increasingly competitive higher education environment.

Yet our understanding of the learning outcomes of higher education, and of the difference that a university education makes, is limited. In Australia as in many countries, conversations about higher education quality have expanded over the last 20 years beyond institutional resources and reputations to include teaching processes and institutional supports. But until recently there has been a surprising lack of information and emphasis on learning processes and outcomes. Capturing and using generalisable measures of learning in quality improvement and accountability calculations requires sophisticated methodology. But as our understanding of measurement and evaluation methodology unfolds, it is critical to develop more sophisticated approaches for managing and improving the quality of higher education.

To this end, the paper discusses four approaches for measuring the value that higher education adds to individual students and their communities. The first approach calculates change by comparing students’ expected against actual university performance. The second
uses assessments of first- and later-year student performance to estimate learning gains. The third approach assesses the extent to which students are engaging in productive learning activities. It is argued that the fourth approach, employer feedback, provides an independent and highly valuable perspective on the outcomes of student learning.

It is proposed that these approaches, which are discussed in terms of current developments in Australian tertiary education, provide complementary means for estimating the value that has been added by higher education to students’ learning and development. Having considered each, the paper investigates their contribution to a robust and potentially scalable methodology for monitoring educational quality. This methodology, which is ambitious in nature, is detailed in terms of its implications for instrumentation, sampling, analysis and reporting. This model does carry controversial implications, and concluding remarks focus on the likelihood of change in quality assurance activities and hence in educational policy and practice.

As noted, a distinguishing feature of these approaches is that they provide direct or very sound proxy measures of student learning outcomes. This is important, as it moves our focus beyond resources and teaching – the focus of most quality indicators in Australian higher education (Coates, 2007b) – and emphasises what students themselves are doing and achieving. So far, in Australia as in most other countries, quality assurance has involved a considerable amount of data collection from students, but little collection of data about students and their learning. A shift in focus towards student-level outcomes is important as quality processes may increasingly be assumed, as it becomes more strategically important for countries to assure the competence and capability of highly trained individuals (James, 2003; The Economist, 2006), and as outcomes data are seen to offer more pointed insight for monitoring the quality and risks of educational provision.

These data collections, and the integrated methodology, offer ‘evidence-based’ approaches to quality assurance, a position that needs brief qualification. In general, ‘evidence-based’ implies forms of professional practice based on data collected using scientific methods. Such data provides a robust foundation for professional diagnosis, decision-making and action, when carefully designed and collected. This implies a certain way of thinking that can play out in various ways in practice. Evidence-based management can denote senior executives making decisions based on data about the quality of provision. It may also involve academic staff using locally collected data to analyse student performance and to help target teaching and support. In education it should, as advanced in this paper, involve analysis of students’ learning processes and outcomes.

Of course the merit of evidence-based practice hinges on the relevance and validity of the data on which decisions are made. This can be more problematic in education than in other professions, as teaching and learning can be difficult to define, measure, analyse and report, particularly in ways that are generalisable across fields of study, let alone across institutions, states or nations. Not all that counts can be easily measured, and not all that is measured necessarily counts. This area is complex and has been explored in a range of recent analyses (see, for instance: Coates, 2007a; Nusche, 2008; Dwyer et al., 2006; Millett et al., 2006; Millett et al., 2008). With such complexities in mind, a considerable amount of work has been done to develop the collections described in this paper as valid, authoritative, relevant, efficient and timely measures that hold weight across diverse contexts.

As the title suggests, this paper emphasises the importance of assessing the ‘value added’ by university education (Saunders, 1999; Meyer, 1997). Measures of absolute performance
are important as they provide information on graduate capability. They do not, however, index the growth in student learning that may be attributed to an educational process. Analyses of the value added by education offer a powerful means of identifying the efficacy of an educational transformation. Education may be considered value adding if, while controlling inputs, it produces a gain in student learning that is above expectation. Of course, it may also be of interest to consider the absolute gains in student learning, the simple difference between outcomes and inputs, as well as those which are above or below expectation.

**Approaches for measuring outcomes and change**

*Comparing expected against actual performance*

In 2008, the Australian Government Department of Education, Employment and Workplace Relations (DEEWR) introduced a pilot program of the Student Aptitude Test for Tertiary Admission (SATTA) (DEEWR, 2008). This program includes funding for an evaluation of the criterion validity of uniTEST, an assessment managed by the Australian Council for Educational Research (ACER). The evaluation (Coates & Friedman, forthcoming) will examine the extent to which uniTEST correlates both with alternative concurrent measures used for university entrance and with performance in the first year of study. This latter evaluation involves analysis of the predictive validity of the instrument.

uniTEST (ACER & Cambridge Assessment, 2008) has been developed to assist universities with the often difficult and time consuming processes of student selection. The test is designed for use with school leavers to complement the existing achievement-oriented measures that form the basis of many selection decisions. uniTEST assesses the kinds of generic reasoning and thinking skills that underpin successful higher education study. It provides measurement of quantitative and formal reasoning, critical reasoning, and verbal and plausible reasoning. Reasoning is assessed in familiar and less familiar contexts and does not require subject-specific knowledge. The instrument is designed to estimate individual capability with known and appropriate levels of precision.

While not the primary purpose of the instrument, or of the 2008 validation, objective measures of individual aptitude provide a basis for estimating subsequent performance. Hence, they provide an inferential foundation for estimating the value added by university study.

In addition to a robust baseline measure, it is necessary to have measures of actual student achievement that are gathered after a period of university study. These are routinely collected through ongoing assessment activities. Such assessments vary greatly both within and across institutions, however, even within similar fields of education. While there are pockets of excellence in assessment, knowledge and skills are often measured using uncalibrated tasks with unknown reliability and validity, scored normatively by different raters using unstandardised rubrics. In addition, the rubrics themselves are often applied with little moderation, adjusted to fit percentile distributions which are often specified *a priori* by departments, faculties or institutions. Confidence in the reliability of such assessments might be enhanced through the inclusion of common and psychometrically validated items across examinations that measure specific knowledge and skill or even generic capabilities.

Such limitations aside, these data allow for the value added by a course of study to be assessed statistically by comparing predicted with actual measures of individual performance.
Performance above expectation suggests value-added growth. Performance below expectation indicates that less value has been added than expected. As noted earlier, a comparison of the simple difference between entrance scores and routine assessment results would also illuminate patterns of learning across an institution.

In addition to any assessment of value added by university study, baseline data on individual ability might also be used by an institution to monitor, scale or even moderate grade distributions across an institution. Such work is undertaken routinely in three senior secondary systems in Australia (VCAA, 2008; QSA, 2008; ABSSS, 2008). In such an analysis, individual performance that is above expectation could be taken to indicate larger gains made through university education. When performance is above expectation for a whole group, however, this may indicate grade inflation or that assessment tasks are too easy. If so, adjustments may then be made for risk management purposes so as to assure the quality of data used for quality assurance decisions.

Assessing change in performance across years

The quality of university education may be examined by comparing objective assessments of first-year and later-year students’ performance and potential. In the simplest scenario, this might involve analysis of routine student assessment data. A first year grade point average, for instance, might be compared against a third year grade point average. This approach is attractive as it involves the use of extant student assessment data. The limitations of the approach, however, stem from uncertainties associated with the psychometric properties of routine assessment data, and whether the assessments have been assured by the educational processes that they are being called upon to evaluate. The process is not grounded by an objective assessment of student competence or capability.

A preferable approach, therefore, involves making comparisons between two psychometrically validated and linked assessments. Data from such assessments provides points of reference from which value-added estimates can be derived. This requires assessment of first- and later-year students, either of the same students as they progress through a course of study, or of a later-year matched cohort of students. The assessments might focus on broad discipline-specific knowledge or skills, such as on reasoning in the biological, mathematical or social sciences, in business, in general studies, or in the humanities. Alternatively, or in addition to this, the assessments might focus on competencies and capabilities that are considered to be more ‘generic’ in nature.

The latter approach, involving the use of standardised assessments, has been more commonly used in Australia. This is perhaps surprising, given the large amount of routine assessment data available to institutions. In Australia, this approach was seeded during development of the Graduate Skills Assessment (ACER, 2001) which measures written communication, critical thinking, problem solving and interpersonal understandings. The Collegiate Learning Assessment (CAE, 2008) has been used in this context in the USA, again to measure generic capabilities which are core components of a university education.

The measurement of generic competencies is important, but there is value too in focusing on phenomena that align with an institution’s specific mission. In 2008, for instance, one Australian university is piloting a Work Readiness Assessment Package (WRAP) which has been designed to measure students’ work-, career- and future-readiness (Coates, Edwards & Nesteroff, forthcoming). This involves the assessment of a spectrum of constructs, from basic competencies such as numeracy and literacy, to job searching, workplace reasoning and
career management, through to how students’ position themselves professionally in the changing world of work.

As detailed above, statistical comparison against learning data collected at two points in time can be used to derive estimates of individual growth against expectation, or the value added by university study. In addition to use in quality assurance determinations, results from such assessment can be reported on the transcripts that are provided to students on graduation, benchmarked as necessary by level of qualification and field of education. They also provide a foundation for drawing inferences about the quality of students’ achievement. As with the previous approach, they furnish independent evidence that can be used to assure the quality of routine student assessment.

**Measuring student engagement**

‘Student engagement’, defined as students’ involvement with activities and conditions likely to generate high-quality learning (NSSE, 2008; Coates, 2006), is increasingly understood to be important for higher education quality. The concept provides a practical lens for assessing and responding to the significant dynamics, constraints and opportunities facing higher education institutions. It provides key insights into what students are actually doing, a structure for framing conversations about quality, and a stimulus for guiding new thinking about best practice.

Student engagement is an idea specifically focused on learners and their interactions with university. The idea touches on aspects of teaching, the broader student experience, learners’ lives beyond university, and institutional support. The concept of student engagement is based on the premise that learning is influenced by how an individual participates in educationally purposeful activities. It operationalises research that has identified the educational practices linked empirically with high quality learning and development (see, for instance: Astin, 1979, 1985, 1993; Pace, 1979, 1995; Chickering & Gamson, 1987; Pascarella & Terenzini, 1991). While students are seen to be responsible for constructing their knowledge, learning is also seen to depend on institutions and staff generating conditions that stimulate and encourage involvement. Learners are central to the idea of student engagement, which focuses squarely on enhancing individual learning and development.

Surprisingly, given its centrality to education, information on student engagement has not been readily available to Australasian higher education institutions. The Australasian Survey of Student Engagement (AUSSE) (ACER, 2008b), conducted with 25 institutions for the first time in 2007, provides data that Australian and New Zealand higher education institutions can use to engage students’ in effective educational practices. The AUSSE builds on foundations laid by the North American National Survey of Student Engagement (NSSE, 2008). By providing information that is generalisable and sensitive to institutional diversity, and with multiple points of reference, the AUSSE plays an important role in helping institutions monitor and enhance the quality of education.

The AUSSE collects data from institutionally representative samples of first- and later-year students, and provides a foundation for analysing change over time. Although these are not assessments of value added in the statistical sense, examining change across year levels provides insight into the extent to which people are being challenged and pushing themselves to learn. An increase in engagement in active learning practices, for instance, indicates that learners are investing more time constructing new knowledge and understanding. It also
indicates that learners are intrinsically more engaged in their work, and hence more likely to be developing their knowledge and skill.

In 2008, ACER is piloting the Staff Student Engagement Survey (SSES) as a complement to the student data collection. The SSES is a survey of academic staff about students, which builds on the foundations set by the Faculty Survey of Student Engagement (FSSE) (FSSE, 2008). The Staff Student Engagement Questionnaire (SSEQ) measures academics’ expectations for student engagement in educational practices that have been linked empirically with high quality learning and development. Data is collected from staff, but students remain the unit of analysis.

Compared with student feedback, relatively little information from academic staff is collected in Australasian higher education. The SSES builds on processes developed in recent surveys of staff and leaders (Coates et al., 2008; Scott, Coates & Anderson, 2008). Information from staff is important, as it can help identify relationships and gaps between student engagement and staff expectations, engage staff in discussions about student engagement and in student feedback processes. It can also provide information on staff awareness and perceptions of student learning and enable benchmarking of staff responses across institutions.

In summary, the AUSSE provides information about students’ intrinsic involvement with their learning, and the extent to which they are making use of available educational opportunities. As such, it offers information on learning processes, is a reliable proxy for learning outcomes, and provides diagnostic measures for learning enhancement activities. This data can be a powerful means for driving educational change, particularly when linked with feedback from staff.

**Recording employer satisfaction**

Graduate employers are important stakeholders in tertiary education who have the capacity to offer independent information on the quality of student outcomes. Surprisingly, this relevance is not generally reflected in formal quality monitoring activities. In Australia, a large amount of data is collected on learners and educational providers. Considering employers’ relevance to education and the lack of available data, there would appear to be much value in collecting data from these stakeholders for use in quality assurance activities.

A model for an employer data collection was developed in Australia in 2006 and 2007, and is scheduled for national deployment in 2008. Three ‘Quality Indicators’ were developed to underpin a new outcomes-focused and evidence-based approach to monitoring quality in Australia’s vocational education and training system (PhillipsKPA, 2006; Coates & Hillman, forthcoming; NQC, 2007). ‘Employer satisfaction’ was defined as one of the three indicators, along with ‘learner engagement’ and ‘competency completion’. Instruments and collection systems were developed to assist organisations collect data in each of these areas.

After design and national validation (Coates & Hillman, forthcoming), the Employer Questionnaire (EQ) was developed to measure three domains (training quality, work readiness and training conditions) and the following sub-scales: trainer quality, overall satisfaction and the effectiveness of assessment, training relevance and competency development, training resources and the effectiveness of support. The EQ is designed to support training organisations to collect data from employers on the quality of education and, more generally, to enhance relationships between education providers and this key stakeholder group.
While the EQ instrument and associated collection systems were developed for use by vocational providers there is, as noted, an important need for such feedback in higher education. Given the increasing economic importance of higher education to the global knowledge economy, it is difficult to see why and how employer perspectives could not become more valued. Employers see graduates in context and are in a unique position to assess their capability and performance. Further, it is likely that many of the same phenomena might be measured in higher education as in more vocational types of training, including employers’ perceptions of teaching quality, graduates’ work readiness and educational conditions.

It is unlikely that feedback from employers could be used in isolation to assess change in learner competence that results from university study. Employers could assess whether a graduate has reached a required level of proficiency, and hence whether sufficient learning growth has occurred. Without information on learner capability prior to study, however, it would not be possible to isolate change due to an educational process. To identify change, employers would need to assess learners at the start and end of their higher education, an uncommon arrangement in many qualifications and fields. It is possible, however, that change over time could be registered at the institution level. Employer feedback on graduates could be compared across areas within an institution, for the institution over time, or against cross-institutional points of reference (Coates, 2007c). Each of these approaches would enable statistical calculation of performance against expectation, and hence the value added by university study.

A model for assessing change

The approaches considered above exhibit key qualities of a robust and potentially scalable model for measuring the value-added contribution of higher education. Key instrumentation, sampling, analysis and reporting characteristics of this model are detailed below. Much of the methodology that underpins this model is appropriated from the Programme for International Student Assessment (PISA) (OECD, 2005b). The PISA methodology has been rigorously developed and tested at an international level, and key aspects of its structure and detail could be sustained and adapted for higher education.

As with PISA’s model, data should be collected from a range of sources and be centred around the assessment of student learning. Working from the OECD INES framework (OECD, 2005a), an assessment framework could be developed that specifies the learning outcomes and contextual themes to be assessed, and provides a basis for specifying constructs and then generating items and instruments (see, for instance: OECD, 2005b). The development of such a framework would, in itself, likely advance conversations about quality assurance. In Australia, for instance, the most recent systematic definition of performance indicators for higher education is nearly two decades old (Linke, 1991).

Development of this framework is beyond the scope of this paper, which instead proposes that data from employers, staff and learners can underpin indicators of learner engagement and outcomes which, in turn, can be used to calculate estimates of outcomes and value added. This approach has desirable methodological properties, including that it allows for triangulation among a manageable number of indicators, focuses on important aspects of education, and contains indicators that are both robust and responsive to change. While it does carry the implication of requiring new data, this may be an unavoidable consequence of any shift towards outcomes-focused quality assurance, and such data would offer direct ancillary benefits to providers.
Student assessments, as detailed above, could focus on generic capabilities relevant to individuals work- and career-readiness, as well as on discipline-specific forms of reasoning. The disciplinary focus is important for establishing the validity and relevance of the assessment instruments and results. These instruments need to be designed and developed in rigorous and consultative ways and involve item preparation, panelling, cognitive interviewing, consultation, pilot testing, expert review, translation, and composition into rotated or modular forms (OECD, 2005b). The instrumentation may be varied depending on whether the assessment was undertaken prior to or at the start of university study. As discussed, a role might be played by routine student assessments, either in terms of providing data for the calculation of value added, or in terms of using objective assessment data to monitor assessment results.

A series of psychometrically linked questionnaire instruments would be required in addition to the objective tests to collect context data from students, staff and employers. These might follow, or be based on, the SEQ, SSEQ and EQ outlined above. They could capture data on student engagement, staff perceptions of students’ engagement and outcomes, and employers’ perspectives on student outcomes. Collection of data from students, staff and employers would enable triangulation of results, identification of process factors that managed to drive educational change, and the multilevel analysis of pedagogical, institutional and environmental effects. It would also play the important role of engaging these stakeholders directly in institutional assessment processes and hence in quality monitoring and improvement.

A sample design would be needed that is both generalisable across institutions and also sensitive to organisational and systemic diversity. Sampling methodology has developed considerably in the last few decades, with stratified and multistage cluster designs used routinely in studies of school achievement (Kish, 1965; Ross, 1992, 1999; OECD, 2005b). Despite such progress in sampling, the census is by far the most common means of collecting data from university students in Australia. In certain contexts, a census may offer a scalable approach for collecting data in different educational and organisational contexts. Yet a census can lead to inefficient use of resources, waste respondents’ time, and produce results with unknown levels of reliability and precision. As data-driven quality assurance becomes more important in higher education, it is necessary to leverage sampling methodology into this operating context. Complex sample designs are required that are feasible and more efficient to implement, sensitive to the phenomena being measured, and methodologically robust.

The assessments outlined in this paper help identify key attributes of a robust and generalisable approach. To begin, target populations of institutions, staff, employers and students would need to be carefully defined. This involves specification of the desired population and of excluded elements. A generic sampling frame would need to be specified that contains all elements required for consistent specification of target population, stratification and weighting. In many if not most cases, lists of students, staff and employers will be held by, and perhaps only by, individual institutions. The frame specification needs to provide a basis for aggregating this information in comparative and contextually meaningful ways. It needs to be possible, for instance, to identify students who are full-time and campus-based students in their first-year of medical study.

The sampling procedure would need to be probabilistic in nature and designed to select a sufficient number of students into the study to generate powerful and representative statistical estimates at the institutional level. While institutions or fields of education would be the ‘level of analysis’, students rather than teachers, staff or institutions would be defined as the primary
‘unit of analysis’. Certain analyses may focus on instructional or organisational factors, however the design would need to be specified to provide information on individual students’ learning interactions and outcomes. It would likely need to be multistage and involve the sampling of institutions, employers, staff and students. Implicit and explicit stratification should be used to enable different sampling designs to be applied for different parts of the population and improve the representativeness of the sample and hence accuracy of results. As with PISA (OECD, 2005b), it is likely that while a series of generic strata could be defined, these would need to be varied to suit the contours of different institutions and systems. Sampling students in clusters may be efficient in various contexts, but in many institutions it may be equally or more feasible to sample students, staff and employers randomly from across the institution. The size of the samples would hinge on technical considerations, the implications of any clustering and oversampling, institutional characteristics and reporting considerations.

Analysis, as with instrumentation and sampling, need to be carefully designed and aligned with the aims of the assessment. Advanced psychometric procedures need to be used to validate and produce scores for composite variables, account for any item sampling that underpins linked rotated forms of the assessment instruments, produce item statistics for equation of scores across time and contexts, facilitate analyses of bias, and calibrate benchmarks or thresholds of increasing performance on the variables being measured. The mixed co-efficients multinomial logit model (Adams et al., 1997; Wu et al., 1997) has been developed for such analysis.

It is vital that the statistical methods used account for the psychometric, distributional and structural properties of the data. Assessing the value added by educational processes is methodologically and educationally complex. Regression models that enable calculation of growth estimates, make appropriate adjustments for individual demographics and educational contexts, and account for any cohort effects and implications from sampling need to be specified. Multilevel modelling may be required, depending on the extent of clustering in the observations.

Well designed reporting plays a critical role in ensuring that assessment results drive effective educational change. A suite of reports are required, including for institutions, groups of institutions, fields of education and individual students. Each of these reports must be technically robust and at the same time presented in informative formats. Scaled change and outcomes scores should be reported, as appropriate, to provide normative baseline data for tracking improvement and drawing cross-group comparisons. To generate change in routine educational practice, however, it would be critical to report results in terms of the criterion referenced benchmarks produced during psychometric validation of the items and scales. This would make clear to students and teachers what levels of competence are tied to different grades.

Other aspects of the overall approach would, of course, require careful consideration in addition to those canvassed here. Key constraints and scope limitations would be essential to analyse. Cultural and linguistic translation, field operations, coding and data management, and quality monitoring, for instance, all play important roles in ensuring assessment processes and outcomes are robust (OECD, 2005b; Coates, Tilbrook, Guthrie & Bryant, 2006).
Forecasting change in quality assurance practice

Together, the four specific approaches advanced in this paper and the overall assessment model emphasise new thinking about quality assurance in Australian higher education, if only through their explicit focus on student learning and development. The application of these approaches in Australian universities is important, for it flags innovative ways for institutions to measure and verify what their students have learned. Each of the four approaches provides institutions with empirical foundations for drawing inferences about the quality of higher education. They provide concrete data that moves beyond prevailing metrics, which focus on graduation rates and subjective measures of student satisfaction with service provision.

While relatively early days, the relevance of these new developments will depend on the extent to which they shape institutional policy and, more importantly, educational practice. Universities and higher education systems evolve slowly and in complex ways but several trends would appear to be driving change more rapidly in this area.

The first of these trends is an increasing emphasis on evidence-based and outcomes-focused approaches in formal quality assurance activities. Spellings (2006), the OECD (2008a) and Callan, Ewell, Finney and Jones (2007) highlight such trends internationally. The direction is emphasised in Australian tertiary education by policy papers released by the Australian Universities Quality Agency (AUQA) (2007) and, in terms of vocational education, by the National Quality Council (NQC) (2007). The capacity to measure the value that has been added by university study is embedded within this discourse. Of course, this trend follows developments in school education over the last few decades, which have culminated in collections such as PISA (OECD, 2008b) and the Trends in International Mathematics and Science Study (TIMSS) (IEA, 2008).

In many respects, this first trend reflects more a general overarching need for objective evidence on the quality of institutional provision and on student outcomes. In Australia, aside from administrative data on student enrolment and completions, the quantitative data used in quality assurance determinations is overwhelmingly derived from students’ perceptions of the quality of teaching and institutional services (Coates, 2007b). While important, such information provides only an indirect and subjective proxy measure of the quality of student learning. Objective assessments, even if of more ‘generic’ rather than discipline-specific phenomena, provide much more direct and robust information and, further, can be used to moderate or monitor routine assessments.

A further driver of change is the need for greater diversification in the data that is collected by institutions for quality assurance purposes. Australian institutions have developed very sophisticated means of capturing feedback on student satisfaction over the last few decades, which has driven important changes in practice. But subjective information on student satisfaction provides just one perspective on education (ACER, 2008a). With a more complex and integrated role in contemporary society, and more differentiation between individual institutions, comes a need for more diversified, robust and educationally significant information. As the widespread adoption of the Australasian Survey of Student Engagement (ACER, 2008b) suggests, institutions need data that helps shape understanding of the student and industry markets in which they operate.

The ideas sketched in this paper are important and ambitious. As noted at the start of this paper, as our understanding of university education and evaluation methodology unfolds, it is critical to consider more effective approaches for managing and improving the quality of
higher education. The proliferation of university league tables in recent years and their attraction as a source of information on quality, despite often serious limitations in methodology (Hazelkorn, 2007; Locke et al., 2008; OECD, 2008c), underlines the need for more evidence-based approaches, and for more rigorous methodology. While the challenges are not small, the successful testing of pertinent methodologies at institutions offers promise, as does the tradition of large-scale studies of student achievement studies in school-level education (Husén, 1996; Schleicher, 1994).

It is essential that academics and institutions themselves take the lead in developing this area of higher education. This is not just because most institutions have the authority to accredit their own programs and ensure academic standards and underpinning quality assurance processes. Rather, it is vital that progress in the measurement of student outcomes and value added builds rather than breaks the link held by teachers and institutions between the development, dissemination and assessment of knowledge. It is important that any new measurement of student learning and development is itself collaborative in nature, given the broader individual, social and economic roles such measures will play.

The capacity to measure the value-added by university education – the difference that it makes – hinges on the provision of robust measures of learner capability and performance. Evidence-based quality assurance requires data that can be used to target enhancement and improvement activities. Such evidence-based approaches are required as institutions grow and diversify, and as it becomes less feasible and even effective to support all areas of provision. A data-driven approach helps identify areas of risk, target limited resources, focus improvement activities, and monitor change. It offers insight for identifying areas of good practice. The perspective driving this paper is that such practice will doubtless grow in an expanding and increasingly competitive higher education environment.
REFERENCES


Organisation for Economic Cooperation and Development (OECD) (2008a). *Programme for International Student Assessment (PISA)*. Accessed 16 February 2008 from: [http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1_1_00.html](http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1_1_00.html).


