

What's the difference? Models for assessing quality and value added in higher education

Hamish Coates

*Principal Research Fellow, Australian Council for Educational Research (ACER),
Private Bag 55, Camberwell, Victoria, 3124, Australia*

Abstract

This paper outlines two approaches being piloted by Australian universities in 2008 for assessing the quality and outcomes of higher education. The approaches offer alternative and complementary means of estimating the value that has been added by university education. They also provide a means of assuring the quality of the routine student assessment processes and results which may be used to underpin quality considerations. The paper concludes that the application of these approaches in Australian universities is important, for it flags innovative ways of thinking about how educational institutions measure and verify the quality of student learning.

Introduction

Measures of student learning are playing an increasingly significant role in determining the quality of higher education. In Australia, self-accrediting higher education institutions are responsible for assuring the quality of learning processes and outcomes (DETYA, 2000). It is a relatively recent development to consider using these learning outcomes to then draw inferences about the quality of the education provided by the institution (OECD, 2008; Spellings, 2006). Clearly, if data on learning outcomes is to be used in this way, institutions need a means of assuring the quality of the assessment processes and data on which their decisions are based.

This paper focuses on ways in which objective measures of student learning might be used by institutions to assess the quality of their educational provision. It presents two approaches for using data on student learning outcomes to estimate the value added by university education. It considers how each of these might be used to assure the quality of the routine assessments that are undertaken in regular subjects and hence manage risk associated with provision.

The first approach computes value added estimates by comparing predicted against actual performance using data from entrance tests and routine assessments. In the second, comparisons are made between outcomes from objective assessments that are administered to cohorts in the first and later years of study. Consideration is given to the kinds of instruments and analyses that underpin each approach, and pilot tests that are currently underway at Australian institutions. It is noted, by way of conclusion, how this work offers Australian higher education institutions new evidence-based approaches to quality assurance.

The emphasis on value-added measures in this paper is deliberate. 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. Value added analyses offer a powerful means of identifying the efficacy of an educational transformation. Education may be considered 'value adding' (Saunders, 1999; Meyer, 1997) 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, as opposed to those which are above or below expectation.

Comparing Predicted Against Actual Performance

As noted, the first approach compares 'predicted' against 'actual' estimates of performance. This approach depends on having data obtained at the point of entry that gives a basis for comparison. Such

data is likely to vary depending on the pathway that a student has taken to university. It might be results from the final year of school or from a prior tertiary qualification. Ideally, this baseline data would be derived from a common objective test that has robust measurement properties and offers firm foundations for statistical inference. This can be provided by an assessment of individual capability.

Several entrance tests are used in Australian higher education. In 2008, uniTEST (ACER & Cambridge Assessment, 2008) and STAT (ACER, 2008) are being tested in national pilot of a Student Aptitude Test for Tertiary Admission (SATTA) (DEEWR, 2008). Both uniTEST and STAT were developed to assess the kinds of generic reasoning and thinking skills that underpin successful higher education study. The assessment tests reasoning in familiar and less familiar contexts and does not require subject-specific knowledge. It provides measurement of quantitative and formal reasoning, critical reasoning, and verbal and plausible reasoning. uniTEST is designed to produce estimates of individual capability with known and appropriate levels of precision. These estimates can be used to predict subsequent university performance.

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 collected through routine assessment activities. Such assessments vary greatly both within and across institutions, even within similar fields of education. Confidence in the reliability of such assessments might be enhanced through the inclusion of common items across examinations that measure specific knowledge and skill or even generic capabilities.

With these data, the value added by a course of study might 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. A comparison of the simple difference between entrance scores and routine assessment results would also illuminate patterns of learning across an institution.

Methodology for such analysis might be explored in the SATTA pilot as part of the evaluation of the uniTEST and STAT instruments that is underway at four universities (Coates, Friedman, Marks & Ball, forthcoming a, forthcoming b). The evaluation will examine the predictive validity of the instruments and hence their capacity to provide a basis for estimating students' aptitude for university study. As such, it will investigate whether the uniTEST and STAT instruments can furnish common baseline data against which future comparisons of university performance might be assessed.

While beyond the scope of the SATTA evaluation, baseline data on individual ability might also be used to monitor and 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). Individual performance that is above expectation may indicate that larger gains have been 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.

Comparing First- and Later-Year Assessments

In the second approach, comparisons are made between the results of first- and later-year student assessments. In the simplest scenario, this analysis might be conducted using 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 that 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 objective psychometric 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 specific knowledge or skills or on competencies and capabilities which are more generic in nature.

The latter approach has been more common, perhaps surprisingly despite the large amount of assessment data available to institutions. This methodology 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, 2007) has been used in this context in the USA, again to measure generic capabilities which are core components of a university education.

Institutions are developing assessments that align with their educational strategies. Such developments are significant, for they offer an objective frame of reference while reinforcing the diversity and value of an institution's specific mission. In 2008, Victoria University is piloting assessments to measure students' work-, career- and future-readiness (Coates, 2007; Coates, Edwards & Nesteroff, forthcoming). This involves the assessment of a spectrum of constructs, from basic competencies and employability skills through to how students' position themselves professionally in the changing world of work. Together, these measures combine to measure the 'VU Student Dividend' (VU, 2007). The VU Student Dividend provides a measure of the value that VU education adds to students.

Rather than assess all students, a sampling approach has been developed to select students in a way that will enable the calculation of estimates for each of the 12 institutionally defined Industry and Community Clusters (ICC). Results of the VU Student Dividend assessment can be reported on the transcripts that are provided to students on graduation. The VU Student Dividend along with the absolute results of the assessment provides a foundation for drawing inferences about the quality of students' achievement. The objective results also furnish independent evidence that can be used to assure the quality of routine student assessments. The kind of moderation sketched for the first approach might also be undertaken.

Summary and Review

The approaches advanced in this paper emphasise new thinking about quality assurance in Australian higher education, if only through their explicit focus on validated student assessment. 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. Both approaches provide 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 student satisfaction with provision.

The first approach compares baseline data from entrance tests against routine student assessment results. The use of common data facilitates comparisons between individuals and courses that would not otherwise be possible. The approach works directly with routine student assessment data, and has the advantage of working this directly into quality considerations and also of providing a direct means of validating such results.

The second approach involves two assessments. This is exemplified in the direction being set by VU and involves assessing a range of basic competencies and high-level capabilities in first year, and repeating the assessment with a later-year cohort of students. It provides insights that align both with institutional strategy and with core aspects of student learning and development.

Both of these approaches provide a foundation that can be used to underpin statements about the quality of provision. Variations or combinations of these alternatives might be considered. For instance, an institution may choose to compare results on a selection test with those from an objective

assessment of later-year students. Items may be validated to measure more discipline-specific aspects of performance, and potentially incorporated into routine assessment tasks.

This paper has argued that if student learning outcomes are to be used to assure the quality of university education, then it is necessary to assure the quality of student assessment. Two approaches have been sketched that offer ways for estimating student performance and growth. Methods for using these approaches to monitor student assessment have also been proposed. These approaches are both currently being piloted in Australian universities, offering new approaches to assuring the quality of university education.

References

- ACER (2008). *Special Tertiary Admissions Test*. Camberwell: ACER.
- Coates, H. (2007). Australian Technology Network of Universities ATN Academic Standards Model. Adelaide: ATN.
- ACT Board of Senior Secondary Studies (ABSSS) (2008). *ACT scaling test. Board of Senior Secondary Studies*. Canberra: ACT Board of Senior Secondary Studies.
- Australian Council for Educational Research (ACER) & Cambridge Assessment (2008). *uniTEST*. Camberwell: ACER.
- Australian Council for Educational Research (ACER) (2001). *Graduate Skills Assessment*. Canberra: Department of Education, Training and Youth Affairs.
- Coates, H. (2007). *The Victoria University Student Dividend Options Paper: Assessing the value added by Victoria University's education and training*. Camberwell: ACER.
- Coates, H. Friedman, T. Marks, G. & Ball, S. (forthcoming a). *Evaluation of the criterion validity of uniTEST*. Canberra: DEEWR.
- Coates, H. Friedman, T. Marks, G. & Ball, S. (forthcoming a). *Evaluation of the criterion validity of STAT*. Canberra: DEEWR.
- Coates, H., Edwards, D. & Nesteroff, S. (forthcoming). *Assessing the VU Student Dividend*. Camberwell: ACER.
- Council for Aid to Education (CAE) (2007). *Collegiate Learning Assessment (CLA)*. New York: CAE.
- Department of Education, Training and Youth Affairs (DETYA) (2000). *The Australian Higher Education Quality Assurance Framework*. Canberra: DETYA.
- Meyer, R.H. (1997). Value added indicators of school performance: A primer. *Economics of Education Review*, 16(3), 283-301.
- OECD (2008). *The Assessment of Higher Education Outcomes*. Accessed 5 February from: http://www.oecd.org/document/51/0,3343,en_2649_35961291_40119475_1_1_1_1,00.html.
- Queensland Studies Authority (QSA) (2008). *Queensland Core Skills Test*. Brisbane: Queensland Studies Authority.
- Saunders, L. (1999). A brief history of educational 'value added': How did we get to where we are? *School Effectiveness and School Improvement*, 10(2), 233-56.
- Spellings, M. (2006). *A Test of Leadership*. Washington: U.S. Department of Education.
- Victoria University (VU) (2007). *Making VU*. Footscray: VU.
- Victorian Curriculum and Assessment Authority (VCAA) (2008). *What is the General Achievement Test?* Accessed 1 February 2008 from: <http://www.vcaa.vic.edu.au/vce/exams/gat/index.html>



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

COATES, HAMISH

Title:

What's the difference?: models for assessing quality and value added in higher education

Date:

2008

Citation:

Coates, H. (2008). What's the difference?: models for assessing quality and value added in higher education. In Proceedings of AUQF2008: Quality and standards in higher education: making a difference, Canberra, Australia.

Publication Status:

Published

Persistent Link:

<http://hdl.handle.net/11343/28861>

File Description:

What's the difference?: models for assessing quality and value added in higher education