

Defining and monitoring academic standards in Australian higher education

by

Hamish Coates

Australian Council for Educational Research (ACER), Australia

This paper outlines the need for adopting a more scientific approach to specifying and assessing academic standards in higher education. Drawing together insights from large-scale studies in Australia, it advances a definition of academic standards, explores potential indicators of academic quality and looks at approaches for setting standards. As learner outcomes need to be placed at the forefront of work on academic standards, this paper concludes by exploring the implications of this position for student assessment and institutional change.

Définition et suivi des standards académiques de l'enseignement supérieur en Australie

par

Hamish Coates

Australian Council for Educational Research, Australie

Le présent essai souligne la nécessité de l'adoption d'une approche plus scientifique de la spécification et de l'évaluation des standards académiques de l'enseignement supérieur. Partant des aperçus tirés d'études à grande échelle réalisées en Australie, il articule une définition de standards académiques, explore des indicateurs potentiels de qualité académique et examine des approches en vue d'établir des standards. À l'heure où les résultats des étudiants doivent constituer la priorité du travail sur la question des standards académiques, le présent essai aboutit à l'examen des implications de ce positionnement pour l'évaluation des étudiants et pour l'évolution des institutions.

Counting standards in higher education

Universities are responsible for maintaining academic standards. Indeed, having autonomy over standards is an intrinsic factor which contributes to their status. This privilege partly derives from the capacity of universities to conduct scholarly research and generate knowledge through teaching – to explore the foundations and contours of knowledge itself. Hence it could come as a surprise that while universities have responsibility for maintaining standards, relatively little scholarly research has been done on what this actually means, particularly in relation to student learning and outcomes. To impel progress in this area, this paper outlines a case for giving greater attention to the specification and assessment of academic standards in higher education. It argues, in particular, that significantly more emphasis needs to be placed on assessing learning outcomes.

Today, it is possible for students at different or even the same universities who have the same underlying competence to perform at comparable levels on the same assessment task, yet receive very different grades. This unfortunate state of affairs may arise from bias or lack of precision in marking processes. Alternatively, it may reflect variability in the tasks or in grading or reporting systems. Even when common tasks, marking processes and reporting metrics are used, these can still mask variability that goes unchecked by the absence of cross-validation. Such practice, where it occurs, is unsatisfactory. Assessment results are high stakes for students, play a significant and formative role in shaping individuals' outcomes, and impact graduates' broader social and professional experiences.

Given the various forms of validation that are used to ensure the quality of academic research, it is surprising that universities can lack procedures for verifying the standards of education. Of course, many excellent accountability mechanisms have been implemented widely by departments, institutions and systems. Continuous improvement, benchmarking, monitoring, risk assessment and auditing all refer to mechanisms for defining, measuring, assuring and enhancing academic standards. Yet while regulatory activity has grown along with the importance of higher education to the knowledge economy, there remains little formal conceptualisation of what is meant by “academic standards” and how these can be assessed. This is disquieting, for there is an increasing realisation (see, for instance, Salmi, 2009) that collegial approaches are no longer singularly sufficient to support the accountability

requirements of contemporary higher education. The drive to define and enhance standards is not new, but expansion of higher education institutions and systems creates new and increased pressure for change.

Close to the heart of much contemporary discourse in this area is a growing interest in developing new ways of documenting what students learn, know and can do. Of course, universities have always had students' learning and development as a core part of their mission. Discussions about higher education quality, however, have tended to focus on institution-level inputs or, in the last few decades, on teaching processes. Recent interest, however, flags a shift towards focusing more explicitly on students – on what they are doing and achieving. Developing, monitoring and reporting achievement and broader graduate outcomes is increasingly seen as an institutional and even system-wide responsibility.

This interest is international in scope. In Brazil, for instance, the pioneering national courses examination, the *Provaão*, has been administered since 1996, and now provides external data across more than a dozen disciplines. In the United States, the Collegiate Learning Assessment (CAE, 2009) has been used by over 400 institutions for around a decade to collect objective data on learning outcomes. *Measuring Up* (NCPPE, 2008) and related reports (see, for instance: Ewell, 2009; Ikenberry and Kuh, 2009; Spellings, 2006) have for some years flagged the absence of information on student learning metrics, and the Voluntary System of Accountability (VSA, 2009) has been developed to drive improvement in this area. In the United Kingdom, the Quality Assurance Agency (QAA) announced early in 2009 that it would look at how to build on the work of external examiners to make more explicit and comparable statements about achievement standards. Cross-nationally, the OECD's International Assessment of Higher Education Learning Outcomes (AHELO) feasibility study (OECD, 2009a) is examining the feasibility of assessing and comparing graduate learning across nations, systems, languages and cultures.

In line with these ideas, this paper brings together insights from large-scale research activities in Australia (AUQA, 2009; Coates, 2007a, 2007b, 2007c, 2007d, 2008, 2009a) to explore the foundations of a generalisable approach to specifying and assessing academic standards in higher education. The following section offers a working definition of academic standards, and is followed by a discussion of the methods used for selecting indicators of quality and establishing performance thresholds. The final sections affirm the importance of focusing on learner outcomes, and look at the implications of this for student assessment and institutional change.

Many of the ideas in this paper are provocative and would not be quick or easy to implement. It is contended, however, that they would result in a

significant improvement to academic work. With a more contextualised sense of how they have performed, for instance, students will have a better idea of how their work will be measured and reported, and a clearer picture of what their final reported performance signifies. Employers will have greater confidence in the meaning of the results that graduates present to them, regardless of the institution from which they come. Academic staff will be able to face and rebut with greater confidence the frequent accusations of falling standards and “dumbing down”.

What “academic standards” means

The concept of academic standards lies at the heart of higher education quality. Perhaps because of this it has proved difficult and even elusive to define. James (2003, p. 189) has suggested that this may be due to the “highly abstract nature of standards... [or because the] higher education sector has had little sustained discussion and analysis of standards in recent years”. Whatever the reason, it is useful to offer a brief deconstruction and working definition of the concept as a foundation for this paper and subsequent practice.

In education, as in many other industries, the term “standards” is used both in substantive and descriptive ways. Standards can refer to shared measures against which comparisons can be made. This is the “what” of standards. Examples include teaching quality, learning outcomes or information resources. In line with the OECD (2009b) and other common use, such phenomena are referred to as “indicators” in this paper.

“Standards” is also used to refer to varying levels of quality or performance, as in phrases such as low performance, high quality or teaching excellence. This is the “how much” of standards. This use is reinforced by technical methods for setting standards (Angoff, 1971). Similarly, the Australian Department of Education, Science and Training (DEST, 2002) states that “Academic standards usually refer to student performance and levels of achievement on a particular piece of assessment, in a subject, or at the end of a degree”. The Australian Universities Quality Agency (AUQA, 2007) defines standards as “A specification or other precise criteria designed to be used consistently as a rule, guideline or definition of a level of performance or achievement”. Anderson, Johnson and Milligan (2000) write that “It is assumed that standards can be represented on a scale, at least to the extent of distinguishing pass and fail and various classes of honours attainment”. Following such use, and consonant with the definition proposed by AUQA, the term “standards” is used in this paper to denote agreed levels of quality or performance on an indicator of student achievement.

As with standards, the term “academic” is used in many different ways. In broad terms, academic might be used to qualify a wide range of activities

that pertain to the production, accumulation and enhancement of knowledge. This work is conventionally distinguished into the three overlapping dimensions of education, research and service. This paper focuses on education and, as the subsequent discussion brings out, places particular emphasis on student learning and outcomes.

It is proposed, therefore, that a useful working definition of academic standards could be “agreed levels of academic quality” or, more fully, “agreed levels of performance on indicators of academic quality”. This implies the need to define indicators of academic quality.

Academic quality indicators

Developing indicators to measure complex phenomena is a complex task. Valued phenomena are not necessarily easy to define, measure and report. Even simple indicators must be accompanied by considerable qualifications and caveats. Indicators can carry different meanings in different contexts, be underpinned by different or differently collected data, or carry different implications for policy and practice. Along with the many educational and practical considerations, the definition of outcomes is a necessarily value-laden process and, as such, it can be difficult to reach consensus about which domains are relevant for a given situation.

Luckily, a considerable amount of work has been undertaken over the last 50 years to develop processes for the specification of higher education quality indicators (Bauer, 1966; Cuenin, 1988; Kells, 1993; Davis, 1996; Cave *et al.*, 1997; Linke, 1991; National Commission on Excellence in Education, 1983; Coates, 2009b). Any selection of indicators needs to balance technical, practical and substantive considerations. In summary, it is vital that indicators are valid, relevant to key phenomena, stable across contexts, transparent, non-trivial, responsive to change, auditable, efficient to collect, preferably readily available, as simple as possible, quantifiable and generalisable. Of course, higher education is a deeply contextualised activity, and indicators for universities should link with those used by schools, industry, the professions, vocational education and training and the broader community.

While data collections proliferated in the 1990s in many countries, in step with the global expansion of higher education and the quality movement, much of this proceeded without reference to an underlying set of indicators. In Australia, for instance, the national indicator system currently in use was developed in the late 1980s (Linke, 1991) for use in a very different kind of higher education system. Of course, institutions and institutional networks developed indicators to drive their own continuous improvement, but on the whole they are not based on the kind of research and consultation that facilitates cross-institutional generalisation.

Indicators are most powerful when they are located within a multidimensional framework. The educational framework that consists of input-process-output stages is one of the most general. This framework usually takes a hierarchical form, and further distinguishes each of the stages into individual, instructional and institutional levels. This basic but robust framework has been used across a large number of diverse education systems and contexts (Bottani and Tuijnman, 1994; Astin, 1985; Ewell and Jones, 1996; Jaeger, 1978). Its most general expression is found in the OECD's Indicators of Education Systems (INES) programme (OECD, 2009b).

In 2007, research was conducted with five Australian universities to develop an academic standards model (Coates, 2007c). An iterative process of review and consultation was undertaken and led to indicators given in Table 1.

Table 1. Indicators of education quality

	Outcomes	Processes	Inputs
Higher education learners	<ul style="list-style-type: none"> ● Graduation rates ● Graduate destinations ● Learning outcomes ● Graduate capabilities ● Work readiness ● Satisfaction 	<ul style="list-style-type: none"> ● Student engagement ● Retention and progress 	<ul style="list-style-type: none"> ● Entry levels ● Entry pathways ● Student diversity ● Student characteristics ● Student aspirations
Higher education teachers	<ul style="list-style-type: none"> ● Teaching experience ● Teaching resources 	<ul style="list-style-type: none"> ● Teaching processes ● Course management ● Support systems 	<ul style="list-style-type: none"> ● Staff characteristics ● University enculturation ● Educational resources ● Curriculum
Higher education institutions	<ul style="list-style-type: none"> ● Institutional growth ● Institutional reputation ● Community engagement 	<ul style="list-style-type: none"> ● Academic governance ● Academic management ● Academic culture ● Staff development ● Quality systems 	<ul style="list-style-type: none"> ● Institutional characteristics ● Institutional resources ● Industry engagement

Source: Coates (2007c).

This framework is not advanced as a definitive articulation of academic quality indicators, but as one means of illustrating the broad nature of phenomena that need to be considered to build a rich picture of the complex business of university education.

Indicators of academic standards should reflect high-level concepts. However, it is necessary to underpin indicators with tangible data elements, hereafter referred to as “measures”. These measures provide evidence of what is being achieved, and thereby play a very important role in the assessment of academic standards. It is important that the measures are carefully

considered, for without careful consideration of data requirements, there is a risk that data availability will dictate overarching quality objectives.

Each indicator might be underpinned by one or more measures. Data elements can be differentiated as follows: those based on fact, on subjective feedback, or on objective assessment. The distinguishing characteristic of the measures is that they are, or could be, operationalised in terms of specific data. Certain measures may be international or national in scope, while others may be relative to specific disciplines or institutions. It is essential that measures are formed in technically and educationally appropriate ways.

Inevitably, given the scale and complexity of higher education, an intricate information architecture is required to map measures onto indicators. It is useful to consider a few examples to illustrate the relationship between indicators and potential measures:

- Graduation rates may be assessed by calculating the percentage of students beginning a programme who graduate within a specified period of time.
- Student engagement is most commonly measured using surveys, or sometimes via time diaries or data collected through administrative systems.
- Teaching experience could be measured through reference to factors such as qualifications, contact hours, professional recognition, or excellence awards.
- University enculturation may be operationalised by measuring the effort institutions make to include teachers in the institution, particularly teachers who are new to the profession or have a contingent employment status.
- Community engagement may be defined variously in terms of alumni programme size, employer satisfaction feedback, participation in service learning, or sponsorship of public events.

Setting performance standards

The enumeration of operational measures facilitates the setting of academic standards. A range of methods are used in education to define quality standards. Standards can be set through reference to the performance of other individuals or groups, an approach which is often used in assessing and reporting student learning. Standards can also be set with reference to specific distributions, often specified in terms of percentiles, as is common in institutional benchmarking processes (McKinnon, Walker and Davis, 2000) or the calibration of student mark distributions. The Angoff standards-setting procedure (Angoff, 1971) has been developed as a robust method for defining the criteria which distinguish varying levels of performance. Standards can be set through empirical or normative analysis, through accreditation or moderation, or through trend analysis. A recent review of extant processes (AUQA, 2009) suggests that application of the word “appropriate” may currently

be among the most common means of setting standards, however this approach fails to provide an objective performance expectation.

A hybrid approach is commonly used that suits the indicators and measures under study and the contexts of development. Standards for high-stakes student assessments, for instance, might well be best set using psychometric modelling along with review by discipline experts. Standards for course curriculum and relevance might be set using rating criteria and expert moderation, either during course development, or on an ongoing and iterative basis. Professional accreditation processes play an important standards-setting function in many disciplines.

From a technical perspective, the standards-setting process should result in the specification of a series of thresholds that map out a continuum of increasing performance. Such thresholds operationalise the indicator in terms of its underpinning measures. The measures therefore play a role in defining the indicator in terms of current and potential planning and practice.

By far the most important standards from a public policy perspective are those that provide assurance that universities are exceeding minimum levels of performance. From a continuous improvement perspective, however, there would also be value in moving beyond minimal standards to set thresholds that measure gradations of increasing performance and potential. Specifying different levels of performance may assist institutions measure improvement against targets, and recognise particularly accomplished forms of performance excellence.

As with the indicators of academic quality, the standards would often be national or international in scope, although a few may be specific to individual institutions, courses or even teachers. Such standards provide a basis for review, audit, monitoring and enhancement processes, and they would derive much of their value from having external points of reference. It is important to stress that setting *a priori* standards does not imply or reinforce any move towards standardisation of institutional practice. Rather, as in the quality control of scholarly research, it provides scientific foundations that provide a springboard for diversity and excellence.

Developing approaches focused on learner outcomes

It is argued in this paper that learner outcomes are the most important of all the quality indicators categorised in Table 1. These, in many respects, can be considered an “educational bottom line”. Of course, many variables influence student performance, including student background, initial knowledge and skill, programme design, student effort, teaching resources and teaching quality. Gathering data about these types of input and process variables and evaluating them are very important exercises, particularly for

each institution's own continuous improvement. Limiting the scope of quality assessment and review strictly to these variables, however, cannot substitute for a direct focus on achievement itself. This is primarily because the various inputs and processes interact in complex ways, and are not deterministic. In large-scale analyses, for instance, it is common for context and demographic variables to explain a relatively small portion of student outcomes (Coates, 2008; Coates and Edwards, 2009).

An explicit focus on learner outcomes examines the net learning effect of all the variables operating together. It allows the attained level of achievement to be assessed and recorded (as grades on student transcripts, for instance), and it allows evaluation of how well the teaching and learning system is working. As Salmi (2009, p. 7) writes, "accountability should not focus on the way institutions operate, but on the results that they actually achieve". Learner outcomes are a basic focus of learners, teachers, governments and graduate employers. Despite considerable investment and development, if students fail to learn and achieve, it is unlikely that teaching and support resources could be considered to be of high quality. Institutional policies may appear coherent, but would be hollow unless related to student and graduate outcomes. Focusing attention on all facets of the educational process is important for quality management and improvement, but this is not by itself sufficient.

Strangely, given the basic importance of learner outcomes, they are demonstrably absent from many quality assurance systems around the world. This is no doubt partly because the assessment of learner outcomes has traditionally been seen as the private preserve of teaching staff. Academic or institutional autonomy are frequently cited reasons for upholding the privacy of student assessment, even though these principles are sustained in research contexts that involve extensive forms of peer validation and review. Another reason may be that unlike research, teaching staff have not typically been rewarded for developing high-quality assessments, which would have an understandable impact on the motivation of staff to invest effort in this area. Of course, very few academics are specialists in psychometrics or student assessment and without training could hardly be expected to have a grasp of key principles and practices. Such training could inform teachers of options for developing good practice. The balance of this paper explores what this may entail.

It is important to set the preconditions for such work by creating a culture that values learner outcomes. Engaging staff in conversations about learner outcomes has been developed as a mechanism for doing this. In the United Kingdom, this culture has come through the development of benchmark statements (QAA, 2009). It is reflected in the learning outcomes frameworks developed as part of the European Tuning Process (EC, 2009) and the "metarubrics" developed in the United States by the Association of American

Colleges and Universities (AACU, 2009). Developing a focus on learner outcomes can occur through professional accreditation, or via the use of objective assessments of learner engagement and capability. In Australia, for instance, many institutions have developed lists of “graduate attributes” (see, for instance, Barrie, 2009). While such work does not go so far as to yield evidence of performance, it can provide a foundation for institution-wide thinking about the broad outcomes expected of learners and graduates.

There have been enormous advances in educational assessment over the last hundred years, much of which is captured by Linn (1989), Keeves (1988), AERA (1999), NRC (2001) and OECD (2003). Important aspects of this assessment methodology, however, have yet to be applied to higher education. Universities and academics are responsible for monitoring and assuring academic standards, and it is critical that continuous efforts are made to enhance the standard of assessment itself. Documentation of a validated suite of measurement approaches lies at the heart of any work on academic standards, for it provides the means that can be used to assure academic quality and enhance teaching and learning practice.

Currently, as contended at the beginning of this paper, many tasks are developed by individual teaching staff for use in specific subjects whose content may change in various ways from year to year. Teaching staff often develop such resources over relatively short periods of time, for localised purposes and with limited resources or knowledge of assessment methodology. As a result, student knowledge and skill is often measured using uncalibrated tasks with unknown reliabilities and validities, scored normatively by different raters using unstandardised rubrics and then, often with little moderation, adjusted to fit percentile distributions which are often specified *a priori* by departments, faculties or institutions.

It is possible to develop validated assessment tasks for a large number of higher education subjects. Achieving consistency across tasks can be vital, because variations in task severity will register as variations in student achievement, regardless of actual competence. Broad subjects based on a single textbook, which take a “shrink-wrapped” approach, can be accompanied by assessment materials. These materials can incorporate formative assignments for continuous assessment as well as validated examinations or items. The tasks themselves could be supported by notes for managing the assessment, analysing data, interpreting results and reporting achievement. A degree of flexibility would presumably need to be designed into the tasks to both encourage and support local adaptations. These assessments could be designed to fit different levels and fields of study, and may include performance tasks, portfolios, open ended questions, constructed response items and multiple choice questions. The validated tasks for these mass subjects could take many different forms, their defining characteristic being

that they are designed to optimise the measurement, diagnosis and enhancement of learning.

Many higher education subjects are specialised in nature or small in scale, however, and it may not be feasible to develop fully validated assessments. It is important, nonetheless, that the resource-consuming nature of assessment task design does not inhibit high-quality practice. In such instances, the most appropriate approach may be to train academic staff. An awareness of basic principles of assessment design and advanced practice would develop the capacity of teaching staff to enhance their own assessment tasks and activities. It would also have more general pedagogical benefits, by requiring academics to think not just about what and how they teach, but about what students are expected to learn and how they should be assessed.

Training teaching staff in assessment could be coupled with a process of assessment task review, in which technical experts or academic colleagues offer feedback on assessment tasks and approaches, and ensure that tasks are of appropriate quality. This feedback may reference quality criteria for student assessment. Of course, this currently happens for many courses and assessments (see, for instance: QAA, 2008), but the process is by no means universal. The largely individualised development of assessment tasks can make it difficult to develop informed and generalisable criteria which map out thresholds of increasing performance. It can be difficult, as a result, for institutions to assure the quality of the tasks which are themselves used to set academic standards.

Moderation processes might be used to ensure the generalisability of assessment standards and outcomes. In general, moderation requires teaching staff to review samples of student work to assure the comparability of standards across contexts. Such moderation may be conducted on an *ad hoc* basis, as often already occurs. It is preferable to design robust and scalable management systems, however, to ensure that outcomes can be quality assured. Moderation could be managed by a cross-institutional agency, as in many senior secondary contexts, or perhaps by a cluster of discipline-specific agencies. The UK External Examiner system illustrates one implementation of moderation in higher education (QAA, 2008). It might involve statistical calibration processes to help equate standards, highlight unusual scores and to manage moderation processes.

Along with the development of formative assessment practice, objective tests can be used to measure critical thinking, problem solving and numeracy skills. Such tests have become popular over the last decade for monitoring the standards of institutional provision. The most widespread, the Collegiate Learning Assessment (CAE, 2009), “presents realistic problems that require students to analyse complex materials and determine the relevance to the

task and credibility". Responses are analysed to assess critical thinking, analytical reasoning, problem solving and communication. The Graduate Skills Assessment test (ACER, 2009) uses multiple choice questions and writing tasks to assess critical thinking, problem solving, interpersonal understandings and written communication. The ETS Proficiency Profile (ETS, 2009) measures critical thinking, reading, writing and mathematics. These tests have the advantage of providing objective estimates of each participant's performance. Data provide external points of reference which can help validate assessment processes and inform moderation and final grading. Similar triangulation may be obtained by drawing, where appropriate, on licensing examinations, consistent feedback from graduate employers or professional bodies, or other information about the performance of graduates.

Translating advances into practice

The notion of "academic standards" touches on most areas of a university's operations, yet is complex and difficult to define. There is, accordingly, no single or simple means of defining, assessing, monitoring and enhancing academic standards. Any single approach is likely to be overly simplistic and prescriptive and, as a result, is likely to underestimate the complexity and significance of the matters at hand, running the risk of promoting standardisation rather than diversification of practice.

A series of new policies and practices would likely be required to support any extensive movement towards implementing the ideas in this paper. In specific instances, new practices might be developed from existing pockets of excellence, or transferred from sectors and systems in which they have been developed and tested over many decades. However, certain forms of more general cultural change would likely be required, along with substantial investments in training and systems. These changes and investments may be costly, but their value lies in the significant information that they would provide on individual learning and growth. Of course, the deeply embedded nature of work surrounding academic standards means that universities have an intrinsic warrant for continuous improvement in this area.

It is almost essential that academics and institutions themselves take the lead in developing this growing aspect of higher education. As Salmi (2009, p. 7) writes, "accountability works better when it is experienced in a constructive way"; "... the most effective accountability mechanisms are those that are mutually agreed or are voluntarily embraced by tertiary education institutions". This is not just because institutions have the authority to accredit their own programmes, ensure academic standards and underpin quality assurance processes. To the extent that student outcomes are stressed, it is vital that progress in this area builds on, rather than breaks, the

authority of teachers and institutions over the development, dissemination and assessment of knowledge. It is important that any measurement of student learning and development is itself collaborative in nature, given the broader individual, social and economic roles such measures will play. Further, it is vital that performance information is reported in multilevel ways that inform and support practice.

This multilevel emphasis is important because, while there are pockets of excellence, developing assessments of performance that simultaneously provide sound information to students, institutions and systems remains a major challenge for higher education. It is common for classroom practice to play out in completely separate ways to the organisation-level assessments managed by institutional researchers, and for both of these to operate in relative isolation from external quality assurance activities. There are of course distinct contexts and foci that shape practice at each of these levels of analysis, yet they share the common basic goal of assessing what learning has been achieved. There would appear to be value, therefore, in identifying approaches that work towards an integration across these levels of analysis.

In this paper it has only been possible to offer a preliminary analysis of the specification and measurement of academic standards. No attempt has been made, for instance, to consider how performance data might be best analysed or how it might be reported to individuals, institutions or regulatory or funding agencies. Further, this paper has not attempted to examine many of the important pedagogical and course management considerations implied by the propositions made about outcomes measurement. Many important questions also surround the value of absolute performance measures compared with measures of individual growth and value added. While graduates and employers often review trends in performance over the duration of a degree, this perspective is rarely captured explicitly in academic transcripts or reports.

In order to encompass the many possibilities for further development in this area, this paper has focused squarely on developing mechanisms for defining and monitoring academic standards in higher education. Facets of the analysis may be considered controversial, but these have been explored with the assumption that much would flow for institutions, systems and – crucially – individuals from making substantive progress in this area. Importantly, it has been contended that standards of education are, or should be, linked in important ways to a capacity to demonstrate that individual learning and development have taken place.

The author:

Dr. Hamish Coates

Australian Council for Educational Research (ACER)

19 Prospect Hill Road (Private Bag 55)

Camberwell, Victoria, 3124

Australia

E-mail: coatesh@acer.edu.au; hamishc@unimelb.edu.au

Ideas expressed in this paper were formed through several projects sponsored by Australasian universities, peak bodies and government agencies.

References

- American Educational Research Association (AERA), American Psychological Association (APA) and National Council for Measurement in Education (NCME) (1999), *Standards for Educational and Psychological Testing*, AERA, Washington DC.
- Anderson, D., R. Johnson and B. Milligan (2000), *Quality Assurance and Accreditation in Australian Higher Education: An assessment of Australian and International Practice*, Australian Government Department of Education, Training and Youth Affairs, Canberra.
- Angoff, W.H. (1971), "Scales, Norms and Equivalent Scores", in R.L. Thorndike (ed.), *Educational Measurement* (2nd ed.), American Council on Education, Washington DC.
- Association of American Colleges and Universities (AACU) (2009), *VALUE: Valid Assessment of Learning in Undergraduate Education*, AACU, Washington DC.
- Astin, A.W. (1985), *Achieving Educational Excellence: A Critical Analysis of Priorities and Practices in Higher Education*, Jossey Bass, San Francisco.
- Australian Council for Educational Research (ACER) (2009), *Graduate Skills Assessment*, ACER, Camberwell.
- Australian Universities Quality Agency (AUQA) (2007), *AUQA Audits – 2007 Onwards: Overview of Audit Cycle 2*, AQU, Melbourne.
- Australian Universities Quality Agency (AUQA) (2009), *Setting and Monitoring Academic Standards for Australian Higher Education: A discussion paper*, AUQA, Melbourne.
- Barrie, S. (2009), *The National Graduate Attributes Project*, Australian Learning and Teaching Council, Sydney.
- Bauer, R. (1966), *Social Indicators*, MIT Press, Cambridge.
- Bottani, N. and A. Tuijnman (1994), "International Education Indicators: Framework, Development and Interpretation", in Centre for Educational Research and Innovation (CERI), *Making Education Count*, OECD Publishing, Paris.
- Cave, M. et al. (1997), *The Use of Performance Indicators in Higher Education: The Challenge of the Quality Movement* (3rd ed.), Jessica Kingsley, London.
- Coates, H. (2007a), "Developing Generalisable Measures of Knowledge and Skill Outcomes in Higher Education", in *Proceedings of AUQF 2007 – Evolution and Renewal in Quality Assurance*, Australian Universities Quality Agency, Melbourne.

- Coates, H. (2007b), "Excellent Measures Precede Measures of Excellence", in *Higher Education Policy and Management*, Vol. 29, No. 1, pp. 87-94.
- Coates, H. (2007c), *ATN Academic Standards Model*, Australian Technology Network of Universities, Adelaide.
- Coates, H. (2007d), "Universities on the Catwalk: Models for Performance Ranking in Australia", in *Higher Education Management and Policy*, Vol. 19, No. 2, pp. 1-17, OECD Publishing, Paris.
- Coates, H. (2008), *Attracting, Engaging and Retaining: New Conversations about Learning. Australasian Student Engagement Report*, Australian Council for Educational Research, Camberwell.
- Coates, H. (2009a), "Building Quality Foundations: Indicators and Instruments to Measure the Quality of Vocational Education and Training", *Journal of Vocational Education and Training*, Vol. 61, No. 4, pp. 517-34.
- Coates, H. (2009b), *What's the difference? A Model for Measuring the Value Added by Higher Education in Australia*, *Higher Education Management and Policy*, Vol. 21, No. 1, pp. 69-88, OECD Publishing, Paris.
- Coates, H. and D. Edwards (2009), *The 2008 Graduate Pathways Survey: Graduates' Education and Employment Outcomes Five Years after Completion of a Bachelor Degree at an Australian University*, Department of Education, Employment and Workplace Relations, Canberra.
- Council for Aid to Education (CAE) (2009), *Collegiate Learning Assessment (CLA)*, CAE, New York.
- Cuenin, S. (1988), *Performance Indicators in Higher Education: A Study of their Development and Use in 15 OECD Countries*, OECD Publishing, Paris.
- Davis, D. (1996), *The Real World of Performance Indicators: A Review of their Use in Selected Countries*, Commonwealth Higher Education Management Service, London.
- Department of Education, Science and Training (DEST) (2002), *Higher Education at the Crossroads*, DEST, Canberra, p. 15.
- Educational Testing Service (ETS), (2009), *ETS® Proficiency Profile*, ETS, Princeton.
- European Commission (EC) (2009), *European Tuning Process*, <http://unideusto.org/tuning>, accessed 1 February 2009.
- Ewell, P. (2009), "Stuck on Student Learning", in *Measuring Up 2008: The National Report Card on Higher Education*, The National Centre for Public Policy and Higher Education, San Jose.
- Ewell, P.T. and D.P. Jones (1996), *Indicators of "Good Practice"*, in *Undergraduate Education: A Handbook for Development and Implementation*, National Centre for Higher Education Management Systems, Colorado.
- Ikenberry, S. and G. Kuh (2009), *More Than You Think, Less Than We Need: Learning Outcomes Assessment in American Higher Education*, Indiana University, Bloomington.
- Jaeger, R.M. (1978), "About Educational Indicators: Statistics on the Conditions and Trends in Education", in *Review of Research in Education*, Vol. 6, pp. 276-315.
- James, R. (2003), "Academic Standards and the Assessment of Student Learning: Some Current Issues in Australian Higher Education", in *Tertiary Education and Management*, Vol. 9, No. 3, pp. 187-198.

- Keeves, J.P. (ed.) (1988), *Educational Research, Methodology, and Measurement: An International Handbook*, Pergamon Press, Oxford.
- Kells, H. (1993). *The Development of Performance Indicators in Higher Education: A Compendium of Twelve Countries*, OECD Publishing, Paris.
- Linke, R. (1991), *Performance Indicators in Higher Education: Report of a Trial Evaluation Study Commissioned by the Commonwealth Department of Employment, Education and Training*, Australian Government Publishing Service, Canberra.
- Linn, R.L. (ed.) (1989), *Educational Measurement*, Macmillan Publishing Company, New York.
- McKinnon, K., S. Walker and D. Davis (2000), *Benchmarking: A Manual for Australian Universities*, Department of Education, Training and Youth Affairs, Canberra.
- National Centre for Public Policy and Higher Education (NCPPE) (2008), *Measuring Up 2008: The National Report Card on Higher Education*, The National Centre for Public Policy and Higher Education, San Jose.
- National Commission on Excellence in Education (1983), *A Nation at Risk: The imperative for Educational Reform*, US Government Printing Office, Washington DC.
- National Research Council (NRC) (2001), *Knowing What Students Know: The Science and Design of Educational Assessment*, NRC, Washington DC.
- Organisation for Economic Co-operation and Development (OECD) (2003), *PISA 2003 Technical Report*, OECD Publishing, Paris.
- OECD (2009a), *International Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study*, www.oecd.org/edu/ahelo, accessed 1 February 2009.
- OECD (2009b), *Education at a Glance 2009*, OECD Publishing, Paris.
- Quality Assurance Agency (QAA) (2008), *Academic Infrastructure*, www.qaa.ac.uk, accessed 1 February 2009.
- QAA (2009), *Subject Benchmark Statements*, www.qaa.ac.uk, accessed 1 February 2009.
- Salmi, J. (2009), "The Growing Accountability Agenda in Tertiary Education: Progress or Mixed Blessing?", in *Education Working Paper Series*, No. 16, the World Bank, Washington DC, p. vii.
- Spellings, M. (2006), *A Test of Leadership*, US Department of Education, Washington DC.
- Voluntary System of Accountability (VSA) (2009), *Voluntary System of Accountability*, www.voluntarysystem.org, accessed 1 February 2009.

