IMHE
Programme on Institutional Management in Higher Education
Programme sur la gestion des établissements d'enseignement supérieur

Outcomes of higher education: Quality relevance and impact

8-10 September 2008
Paris, France

Enseignement supérieur : qualité, pertinence et impact

8-10 septembre 2008
Paris, France

Assessment Reform as a Stimulus for Quality Improvement in University Learning and Teaching: an Australian Case Study

Stuart Campbell, University of Western Sydney, Australia

ORGANISATION FOR ECONOMIC
CO-OPERATION AND DEVELOPMENT

ORGANISATION DE COOPÉRATION ET
DE DÉVELOPPEMENT ÉCONOMIQUES

ASSESSMENT REFORM AS A STIMULUS FOR QUALITY IMPROVEMENT IN UNIVERSITY LEARNING AND TEACHING: AN AUSTRALIAN CASE STUDY

Stuart Campbell, University of Western Sydney, Australia

Improving student assessment quality and practice is a particular challenge for higher education leaders. Internationally, assessment practice is seen as in need of improvement; at the same time there is a worldwide preoccupation with league tables and standards. A student assessment reform process at the University of Western Sydney is described. This year long process entailed the simultaneous development of a new assessment policy, an assessment guide, and communities of practice around assessment. The reform process had impacts beyond pedagogy: It impinged on management processes, curriculum renewal, attitudes to student centredness, higher education scholarship, governance arrangements, professional development, and industrial relations. In summary, it is argued that student assessment reform is a strong lever for quality improvement in learning and teaching and beyond, and that it poses challenges for higher education leaders in a broad range of management domains.

Introduction

It is argued in this paper that student assessment reform is a powerful stimulus for quality improvement in higher education, and that the effect of such reform goes well beyond the domain of assignments and examinations. Using as the point of reference an assessment reform process at an Australian university, we claim that such reform has the potential to impact on governance, curriculum, scholarship, professional development, industrial relations and other domains. The paper falls into three main sections: Firstly we discuss the critical status of student assessment for higher education leaders working in an environment of continual change. Secondly, a brief case study of assessment reform at an Australian university is presented in general terms. The third section enumerates the domains of university life and organisation that were impacted by this reform process. The paper concludes by advocating that higher education leaders use student assessment reform as a broad improvement strategy.

Student assessment in an environment of continual change

The essence of higher education leadership is the management of change in an increasingly complex environment. Contemporary challenges include rising student participation and graduation, increasing numbers of migrants participating in education, older people taking advantage of educational opportunities, a shrinking education workforce, shifts in interest in particular disciplines, the changing balance of private and government funding of tertiary study, and increasing movement of students to study abroad (OECD 2007a).

In this dynamic environment, student assessment occupies a critical position for higher education leaders given its capacity to impinge on numerous policy, management and cultural issues: We set the scene by reviewing some of these issues, including student assessment quality, the commodification of grades, the economic cost of student assessment, the relationship between assessment and academic values, and the current standards debate.

The quality of assessment practice internationally

There is ample evidence internationally that student assessment in higher education is an area of practice that needs improvement in a host of domains as evidenced by the findings of quality agencies around the world. In Britain, the Quality Assurance Review identified student assessment as a key area of concern (QAA 2003), while the recently formed UNDP Regional Bureau for Arab States reports in a quality audit of 23 universities that:

... assessment continues to be a weakness, particularly in three respects. First, too much emphasis is placed on the memory recall of descriptive knowledge. Second, not enough is done to test higher-level cognitive skills. Third, there is virtually no moderation either internal or external to ensure the fairness and transparency of marking. (UNDP/RBAS 2006:5)

Some of the most detailed evidence on the quality of assessment practice comes from Scott's CEQuery research, which analyses open ended comments written by 94,835 students on the annual Course Experience Questionnaire administered by fourteen universities in Australia (Scott 2006: iv). The CEQuery tool analyses student comments, assigns them to subdomains, and classifies them as best aspects (BA) or needs improvement (NI). For the five subdomains of Assessment (Relevance, Marking, Expectations, Feedback/return, Standards) the odds of a BA comment rather than a NI comment was mostly low, i.e. 1.8, 0.2, 0.2, 1 in 10, 0.3 respectively (Scott 2006: 27). Put simply, while students are – in their own words – equivocal about the relevance of assessment, they are more likely to be unhappy about the way their work is marked, unclear about what is expected of them, unhappy about the feedback they receive and unclear about the standards of assessment. There is no reason to believe that Scott's findings are unique to Australia, and that poor assessment practice is not a serious issue for university leaders around the world.

The commodification of marks and grades

Marks and grades are increasingly seen by students as commodities with a purchase price and spending power; in line with the OECD's findings, university students know that an investment in fees will generate a life time return:

Across 25 OECD countries and the partner economy Israel, individuals with university degrees and advanced research education had earnings that were at least 50% higher than individuals whose highest level of educational attainment was below the upper secondary level. (OECD 2007a:6)

The corollary of this is that marks and grades are worth money, and that as fees increase, students "may well shift further to see a degree as a commodity, bought to specification and with an expectation that the supplier will deliver" (Connolly *et al* 2006: 136). Commodification is impelled and facilitated by a number of factors such as the streamlining and automation of management information systems in mass higher education, modularisation of courses, the increasing sophistication of measuring learning outcomes, more accurate activity costing and fee setting, and of course competition among providers. A fully modularised undergraduate course with, say, thirty separate units of study that will each be allocated a grade entails thirty separate investments in the graduate's future and thirty different potential impacts on the return on the overall investment. Furthermore, grades have the awkward characteristic that they cannot normally be improved on: Once you have a pass grade in Biology 101 rather than the credit grade you hoped for, the economic effect stays

with you for ever, lowering your grade point average and diminishing your chances of competing for jobs and scholarships. It is no wonder that grades are so hotly contested and that universities set up elaborate systems for appeal and review. To try to imagine the scale of stress potentially generated by a university assessment regime, at the author's university, about 200,000 grades are generated annually, each with a potential impact on the economic futures and emotional wellbeing of students.

The economic value of assessment activity

The amount of time spent by students and staff on assessment activity must be considerable. At the author's university, the annual value of the work done by students could be in the region of twenty million euros, assuming that some 25,000 equivalent full-time students spend perhaps ten hours in each of their eight annual units of study directly engaged in assessment tasks, rather than spending the time in casual work at the rate of ten euros per hour. In other words, student assessment work (let alone staff assessment work) at this university represents the lost productivity opportunity to staff a fast food chain or a factory with 1000 employees. These alarming figures underline the moral imperative for assessment to be purposeful for the student, and for it to be productive – at least in the sense that goods or services will be generated by the graduate in due course. If an assessment task is irrelevant, the student would gain more by spending the time working at a supermarket checkout; and the capacity of the graduate to be successful in their future career will have been diminished because they have learned less than they could have.

Assessment as a proxy for values in higher education

As the crucible for the generation and contestation of ideas, higher education institutions problematise assessment, essentially through the enduring binary opposition of norm-referenced and criterion-referenced approaches. In crude terms, the norm-referenced approach privileges competition over opportunity, while the criterion-referenced approach aligns ranking with objective standards rather than individual superiority. In reality, the ideological lines are much more blurred (Orr 2007) and it is this blurring that provides an arena for nuanced debate about the purpose and nature of assessment. The debate is accentuated by the statistical foundation of much practice in assessment, for example the theoretically stark but practically blurred difference between statistics in norm-referenced marking systems that mirror natural distributions, and in criterion-referenced systems that allocate marks to benchmarked achievements.

Perhaps the current preoccupation with standards and the rising prominence of university league tables may steal some of the intellectual oxygen from this very old assessment debate. Australia's university quality agency AUQA is basing its second audit cycle on standards, and the likelihood is that the resulting increase in cross-university assessment benchmarking will raise new values debates focussed on the relative merits of educating existing elites on the one hand and adding value to the less advantaged on the other hand. In summary, with its potential to elicit debate about societal and educational values, assessment reform is good for the intellectual health of universities.

Assessment and standards

The increasing worldwide obsession with the relative position of universities in league tables and the "near-obsession with the status and trajectory of the top 100" (Hazelkorn 2007:82) highlights the importance of common standards: What standards can we use to know

whether university X is as good as university Y? What degree of agreement is there on such standards, and how valid are they? Notwithstanding the wide publicity about league tables, they are not unproblematic: University rankings on the whole "focus on inputs, activities and research outputs, such as resources used, classes taught, and articles published" (Nusche 2008: 3) and have little to say about what students actually learn.

Learning outcomes are more effectively assessed at the national or international level through such instruments as the Australian Course Experience Questionnaire, or the National Survey of Student Engagement used in the USA and Canada (see Nusche (2008) for a comprehensive account of such instruments). Such instruments cannot, however, assist in benchmarking standards at the fine grained level of subjects, units or modules; they cannot help university X know whether its history or physics or Chinese program is at a similar standard to those of university Y. And it appears that in Australia at least, not much has been done to use the assessment of student learning outcomes as a yardstick: The Australian Quality Agency in its first audit cycle, made numerous recommendations to institutions to, for example, "… review … processes for assessment, including moderation and benchmarking" (Stella and Woodhouse, 2007:28). The same report laments that in their audits "Institutions do make frequent reference to benchmarking but in many cases little has actually been done and what there is appears ad hoc" (23)

Benchmarking and standards occupy centre stage in the second audit cycle, for which

AUQA has been asked to adopt an "academic risk assessment" approach and pay attention to "evidence of setting, maintaining and reviewing institutional academic standards and outcomes, together with evidence of the institution's comparative national and international performance, …" so that it is better able to make unequivocal statements about graduate standards (Stella and Woodhouse, 2007:38).

Benchmarking is especially difficult with student assessment in university courses that are not accredited by national or international professional bodies. In Australia, for example, student assessment in courses such as engineering, medicine and translation is benchmarked against common national standards (although inputs are probably more commonly benchmarked than learning outcomes); disciplines like history, literature and languages tend not to have documented standards.

Industry's increasing interest in student assessment includes and transcends the issue of standards. Employers are more and more concerned with ensuring that universities embed assessable generic and specific skills in their courses. A report prepared for the Business, Industry and Higher Education Collaboration Council was concerned that universities in Australia have no uniform system for assessing employability skills, and recommended the development and delivery of "a professional development program, with a suite of supporting resources which specifically address best practice in teaching and assessment of the eight employability skills in the higher education context" (Precision Consultancy 2007: 52).

Given such interest from industry and the increasing prominence of league tables, it is very much in the interests of higher education leaders to devise and implement student assessment systems that demonstrate the quality of their graduates in relation to the rest of the higher education sector nationally and internationally, and to insist that the emphasis is on benchmarking outcomes rather than inputs.

An Australian case study of assessment reform

Background

Like higher education around the world, Australian higher education system is in its own state of flux. The system comprises

- 39 universities of which 37 are public institutions and 2 are private;
- 1 Australian branch of an overseas university;
- 4 other self-accrediting higher education institutions; and
- non-self-accrediting higher education providers accredited by State and Territory authorities, numbering more than 150 … (OECD 2007b:x)

In this increasingly competitive system, characterised by non-university institutions offering degrees - a function formerly seen as university territory - the government university sector in the difficult position of relying on international students to subsidise the shortfall between the cost of teaching local students and the income received on their behalf. Over the whole higher education sector, 26% of student load in 2005 was from overseas (OECD 2007b: xiii). With current weak local demand for university places, softening entry scores, and continual financial belt-tightening, the Australian university sector finds itself in the dilemma of having to do more with less. Universities await with interest the details of the "Education Revolution" promised with the change of Federal Government in 2007 (Marginson 2008).

Inevitably, student assessment is squeezed between the tectonic plates of pedagogy and financial sustainability: Assessment of student learning outcomes is an indispensible facet of university work; on the other hand it is an expensive activity that is easily identified as one where economies can be made under the management exhortation of doing less but doing it more effectively. This squeeze on student assessment is no doubt one of the factors that have led a number of Australian universities to undertake renewal of student assessment practise and policies in recent years.

In 2007, the University of Western Sydney (UWS) embarked on a major review of its student assessment regime. The review was conceived as the centrepiece among a number of other projects in the university's Learning & Teaching Action Plan for 2007. The annual plan itself was developed in top-down and bottom-up fashion; broad directions were set by the senior management group and the implementation measures were developed in consultation with academic staff at all levels. As the Action Plan rolled out through the year, it had a high degree of recognition and acceptance among staff, and the assessment review came as no surprise. The positive spin off effects of the review were probably not fully understood when the project was conceived; as we will explain in the latter part of this paper, the assessment review had implications for many facets of the university's business.

While we describe the strategies employed in the review here, the outcomes of staff discussions and data gathering processes are not reported. Rather, we focus on the impact of the reform process beyond the immediate domain of student assessment.

Rationale for reform

The decision to undertake a major review of student assessment was based on a number of factors, including a high volume of student complaints about assessment relative to other

matters and negative student feedback about assessment on UWS student data collection instruments (AUQA 2007:20). Some if not all of the negativity about assessment derived from the university's policy on assessment, which included some inconsistencies, particularly about its philosophical basis; it was equivocal and difficult to interpret on whether it required a norm-referenced or criterion-referenced approach, or a hybrid approach. Another motivation for the review was the strong belief of the leadership group involved that assessment at our university should underpin learning (Boud and Falchikov 2006, James *et al.* 2002, Scott 2005). Running parallel with the assessment reform project was another project aimed at revising the kind of documentation supplied to students relating to individual units of study; this project aimed to reengineer the traditional unit outlines to make them assessment-focussed. It was thought that the two projects could (as they did) evolve symbiotically.

A three-pronged approach: Policy, communities of practice, assessment guide

The review was conceived as a three-pronged approach lasting for the whole of 2007. The central component was the review of the policy itself. This was undertaken by a small senior committee chaired by the Pro Vice Chancellor (Learning & Teaching) with accountability to the Senate and to the University Executive. This double accountability was to ensure that both the academic and executive management domains of the university approved the new policy. Accountability to the managerial domain was essential because of the resource demands of implementing the new policy.

Parallel to the work of the policy review committee, a process of awareness raising and staff development took place. This work was undertaken in three domains. Firstly, a program of internal grants had been established to permit teams of academics to undertake learning and teaching improvement projects. "Assessment for learning" was one of the themes under which grant applications could be made, and a number of grants were approved to improve assessment. Another category "Flexible and blended learning" attracted applications to work on assessment focussed learning guides. As a result, during 2007 numerous academics were working on funded projects dealing with student assessment in some way (see http://www.uws.edu.au/__data/assets/excel_doc/0008/17459/LTAP_Action_Projects_080508. xls). The second domain was activism in the Schools and Colleges (UWS has three large Colleges, each comprising between five and six Schools). This work was undertaken by Teaching Fellows, leading teachers seconded to the role of learning and teaching improvement, and consisted largely of organising local forums dealing with assessment. The third domain was centrally organised activism comprising an on-line staff questionnaire on assessment, a full day university-wide forum on criterion and standards based assessment, and a full-day forum to elicit the views of the Heads of Program; this group has been identified as a crucial stratum of education managers without whom change cannot be implemented. Typically senior lectures administering whole degree programs or parts of very large degree programs, these staff are routinely consulted on major strategic moves because of their knowledge of operational contexts. The UWS Head of Program Network features in the AUQA Good Practice Database (http://www.auqa.edu.au/gp/search/index.php). In summary, all of this activity helped to develop communities of practice around student assessment.

As the year proceeded, lines of communication were kept open between the policy committee and the communities of practice by way of cross membership of key groups. In addition, the policy committee had the benefit of high quality management information on assessment from the UWS Tracking and Improvement System for Learning and Teaching (TILT)(http://www.uws.edu.au/opq/planning_and_quality/tracking_and_improving_performa nce#1).

The third of the three prongs was the compilation of an Assessment Guide to accompany the new policy. This guide was produced by the university's Teaching Development Unit, which had cross membership with the policy committee and the communities of practice. With these three activities running in parallel, a considerable number of staff came into contact with the assessment review in some way or in multiple ways during 2007.

Early in 2008 the policy was approved by the university's Senate after proceeding successfully through the governance and executive management hierarchy. Apart from a large amount of administrative streamlining, the key differences between the old and new policies can be summarised as:

- Elimination of ambiguity about the philosophical basis for assessment, and clarification that the new policy is underpinned by a criteria and standards approach.
- A requirement in the new policy that the standards for different levels of pass (pass, credit, etc.) are explicitly described in the documentation given to students for each unit of study (e.g. each of the four units taken each semester).
- Removal of any requirement that grades conform to a normal distribution; the previous policy had a weak requirement of "curviness" in grade distributions, but there was some folk belief that the university required grades to strictly "follow the curve".
- Clarification of the term "moderation" and a requirement for moderation to occur.
- Proscription of statistical scaling except in exceptional circumstances.

The significance of all of these changes is that they were not merely technical or bureaucratic; they challenged the established beliefs of some academic staff, and required individuals and systems to face quite significant change. For this reason, the university recognised that the implementation of the policy would be a significant leadership task. Consequently a major project was established to rework all the university's assessment schedules over a two year period.

In the following section we discuss the many ways in which this case of assessment reform, and potentially any such example, impinges on and disrupts university systems and processes in multiple ways. We will argue, however, that this disruption should not be regarded as a negative; rather, assessment reform can send an energising jolt through organisational culture, stimulating reflection and encouraging positive change and renewal.

Assessment reform as a centrepiece of quality improvement

An outcome of the assessment reform process as UWS has been that it can serve as the centrepiece of learning and teaching improvement, warranting a considerable allocation of management attention and leadership capacity. In this section we enumerate various facets of university work where we have found assessment reform to be a powerful force for positive change in learning and teaching, and for the engagement of leaders at various levels of seniority.

Engaging and linking multiple constituencies in learning and teaching improvement

Our experience at UWS was that various constituencies have their own special interests in assessment, and that shepherding policy change through these constituencies involved a progression towards consensus as each group accommodated the special interests of the other. As Pro Vice Chancellor (Learning &Teaching) the author played a networking role in negotiating the policy through these constituencies (Smith *et al.* 2007:37). For example, the

program heads had a sharp perspective on operationalising the policy directly with students, and were able to anticipate pitfalls and to spot naiveties. The university's executive on the other hand had a concern with the cost of implementation, one that was far from the concerns of program heads with no budget responsibilities. At each exposure to the particular constituencies, consensus grew. The reform process had the effect of consolidating the practice of top-down bottom-up change management.

Encouraging curriculum renewal

The development of communities of practice led seamlessly into curriculum renewal. Our expert teachers and educational developers took the line that assessment reform could not take place in isolation. In operational terms, much of the curriculum renewal work occurred through small grant projects to develop assessment-focussed learning guides for each unit of study. Over a dozen project teams developed such guides under the guidance of a Teaching Fellow. The work was very much influenced by Biggs' Constructive Alignment model (Biggs 2003). These learning guides were eventually made mandatory and they are now required to conform to the new assessment policy.

Focussing on student centredness

UWS is not alone in placing the student experience at the heart of its mission. The assessment renewal process provided additional opportunities for university management to stress the importance of student centredness. A key strategy in this regard was to expose academic staff to student survey data that evaluated assessment practice poorly. The very effective message to staff was that (a) students are unhappy about assessment, and (b) we collectively know how to fix the problem. This assessment focus on student centredness assisted efforts by senior university leaders to develop a statement about the distinctive student experience as a keystone of the next iteration of its strategic plan

Stimulating scholarship in learning and teaching

An important element of the UWS improvement strategy for learning and teaching is to encourage scholarship in the area. The assessment reform process stimulated scholarship among the communities of practice, and helped to shape the research agendas of some of our expert teachers. It was also one of the factors in encouraging senior staff to begin to develop a more coherent approach to higher education research in the university, for example by fine tuning research ethics approval processes to better accommodate projects using student related data.

Testing management and governance structures

The centrality of assessment to the complex life of the university meant that the reform process tested the effectiveness of the management and governance structures, a particular issue in a university like UWS that clearly demarcates between academic governance and executive management. Essentially the reform process required negotiation between two clusters of interests: The academic governance domain had responsibility for the academic integrity of the new assessment regime, while the executive management stream had the responsibility for resourcing and implementing the regime. One of the tasks of the PVC (Learning and Teaching) was to manage this negotiation.

Improving professional development

Recognising that the new policy would not be simply absorbed via email, the university enhanced its existing professional development program to incorporate the new regime. One effect was that the training program for new academic staff was updated to incorporate the new assessment regime. More profoundly, the scale of the task of assessment renewal impelled the university to refashion its centralised development model to become a hybrid of centralised and distributed provision. The process was not without tensions, and in fact provoked discussions at senior levels about the relative merits of centralised and distributed provision of academic support services.

Focussing on industrial relations and the work of university teachers

The assessment reform process impinged on the university's industrial relations regime by potentially unsettling the existing formulation for calculating the workloads of teachers. Formerly it was common practice to separate out assessment as a separate workload activity from teaching; the new policy encouraged more integration of the two, thus posing the challenge to university leaders of coming up with a new consensus on how to measure the value of academic work.

Concluding remarks

The experience of the University of Western Sydney has been that student assessment reform has proven – somewhat unexpectedly – to have had quite far reaching impacts on the work of the university, and has exposed senior staff to new management challenges. Our advice is that when a student assessment system needs renewal, the issue should not be compartmentalised as a narrow matter of pedagogy; the preferable approach is to use it to test the resilience of existing systems and beliefs, to stretch and refine the leadership capacity of the institution, and to leverage quality improvement on a broad front.

REFERENCES

Australian Universities Quality Agency (2007), *Report of an Audit of University of Western Sydney* Australian Universities Quality Agency, Melbourne.

Biggs, J.B. (1999), *Teaching for Quality Learning at University* Open University Press, Buckingham, UK.

Boud, D. and N. Falchikov (2006), "Aligning assessment with long-term learning", *Assessment & Evaluation in Higher Education*, Vol. 31, No. 4, pp. 399-413

Connolly, M., N.Jones and D.Turner (2006), "E-learning: A Fresh Look", *Higher Education Management and Policy*, Vol. 18, No. 3, pp. 128-139.

Hazelkorn, E. (2007), "The Impact of League Tables and Ranking Systems on Higher Education Decision Making", *Higher Education Management and Policy*, Vol. 19, No. 2, pp. 81-105.

James, R., C. McInnis and M. Devlin (2002), *Assessing Learning in Australian Universities* Centre for the Study of Higher Education, The University of Melbourne, Melbourne.

Marginson, S. (2008), "Conditions for an Education Revolution", *Dialogue,* Vol. 27, No. 1/2008, pp. 3-16

Nusche , D. (2008), *Assessment of Learning Outcomes in Higher Education: A Comparative Review of Selected Practices. OECD Education Working Paper No. 15* OECD, Paris.

OECD (2007a), *Education at a Glance 2007: OECD Indicators. Executive summary* OECD, Paris

OECD (2007b) *Thematic Review of Tertiary Education - Country Background Report - Australia* OECD, Paris

Precision Consultancy (2007) *Graduate Employability Skills* Business, Industry and Higher Education Collaboration Council, Canberra.

Orr, S. (2007), "Assessment moderation: constructing the marks and constructing the students", *Assessment & Evaluation in Higher Education,* Vol. 32, No. 6, pp. 645-656.

Quality Assurance Agency (QAA) (2003), *Learning from subject review 1993–2001: sharing good practice* Quality Assurance Agency for Higher Education, Gloucester.

Scott, G. (2006), *Accessing the Student Voice: Using CEQuery to identify what retains students and promotes engagement in productive learning in Australian higher education* Department of Education, Employment and Industrial Relations, Canberra.

Smith, D., J. Adams and D.Mount (2007), *UK Universities and Executive Officers: the Changing Role of Pro Vice Chancellors. Final Report* Leadership Foundation for Higher Education, London.

Stella, A. and D. Woodhouse (2007), *Benchmarking in Australian Higher Education: A Thematic Analysis of AUQA Audit Reports, Occasional Publications Series No: 13* Australian Universities Quality Agency, Melbourne

UNDP/RBAS (2006), *Quality Assessment of Education Programmes in Arab Universities Regional Overview Report* United Nations Development Programme Regional Bureau for Arab States.