The AAGLO Project

The ‘Assessing and Assuring Graduate Learning Outcomes’ (AAGLO) project has been funded through the Australian Learning and Teaching Council’s (ALTC) Strategic Priority Grant scheme to gather evidence about the type of assessment tasks and assurance processes that provide convincing evidence of student achievement of and progress towards graduate learning outcomes. AAGLO is one of a number of related ALTC projects and fellowships that reflect increasing international attention to the quality of student learning outcomes.

This shift of emphasis from teaching inputs to learning outcomes has been a characteristic of considerable international activity aimed at articulating discipline-specific statements of the learning outcomes university students should demonstrate by the time of graduation. Such statements vary according to the perspectives and purposes of their creators. Some reflect academic learning outcomes, some lean more towards employability, while others indicate transformational aspirations related to the development of democratic perspectives and global citizenship. Terminology around graduate learning outcomes is also diverse – ‘graduate attributes’, ‘graduate capabilities’, ‘competencies’ and ‘competences’ being the most common. Though the definitional distinctions among these terms have long been the focus of debate, for the purposes of this project, ‘graduate learning outcomes’ will address whatever a program, institution or discipline has articulated as the intended learning outcomes of a particular university experience.

Background

The articulation of graduate outcomes is not new and universities in Australia and elsewhere have for some time used these statements to construct, communicate and market institutional identity. Recent international movements however are driven to a certain extent by the desire to establish, enhance and assure academic standards, provide a basis for demonstrating their comparability within and between institutions and countries and to encourage and support increased student and staff mobility. Extensive, large-scale, consultative processes have been employed in the development and implementation of graduate learning outcomes that will serve these purposes in centres around the world (Harris 2009).

Over a decade ago the UK Quality Assurance Agency (QAA) initiated a process whereby disciplinary communities developed statements of expected learning outcomes called Subject Benchmark Statements which vary in their level of detail and in the levels of award targeted. Resources have also been developed to guide the use of subject benchmark statements through processes which include a well-established system of external examiners.

Similarly, a European collaboration The Tuning Project - now referred to as a Process – was established and articulated in the 1999 Bologna Declaration as a strategy for increasing collaboration in higher education (the expression tuning chosen to convey the objective of harmonising rather than standardising higher education systems among participating countries). A key component of Tuning has also been the generation of external reference points - threshold-level learning outcomes and competences - by communities across disciplines such as history, chemistry, nursing and business. A subsequent activity in 2004 was the development of the Dublin Descriptors, broad statements of learning outcomes that distinguished among Bachelor, Master and Doctoral level awards in five areas of learning.

As a result of interest in the European Tuning Process, a Tuning Latin America Project was initiated in 2004 to encourage dialogue around teaching, learning and assessment in order to facilitate the exploration of points of agreement on qualifications within the region and to establish links with European systems.

Tuning USA, an initiative supported by the Lumina Foundation to highlight the process for establishing and demonstrating standards of academic achievement, is also related to Tuning Europe. The initial focus is on graduate learning outcomes in six disciplines (biology, physics, chemistry, history, education and graphic design) across four states (Indiana, Minnesota, Utah and Texas). The USA has also expanded the use of standardised instruments such as the Collegiate Learning Assessment (CLA) to assess graduate achievement.

Canada and areas of South-east Asia - Hong Kong, Taiwan, Singapore and Japan - have shown interest in monitoring these developments but as yet no specific processes have been put in place to replicate Tuning-type activity in these regions.

Since 2010 Australia’s participation in the international standards movement has been supported by the Australian Learning and Teaching Council’s (ALTC)
Learning and Teaching Academic Standards (LTAS) project. To date this project has facilitated discipline communities in the articulation of threshold learning outcomes in Architecture and Building; Arts, Social Sciences and Humanities; Business, Management and Economics; Creative and Performing Arts; Engineering and ICT; Health, Medicine and Veterinary Science; Law; and Science. The LTAS project has generated considerable activity both during and after the endorsement of completed threshold learning outcomes. Disciplines scholars have initiated follow-on projects, frequently as collaborative efforts with discipline deans, and the ALTC has also generated further activity through its grants and fellowships schemes.

Developments in the assessment and assurance of graduate learning outcomes

This point-in-time overview summarises the current state of activity in an area with heavy exposure to the influences of constantly evolving movements in policies and procedures related to quality assurance. To date the main take-up of graduate learning outcomes has been for purposes of program development and approval. There is now evidence of increased attention to the collection of convincing assessment evidence that students have achieved them.

The Organisation for Economic Co-operation and Development (OECD) is conducting an AHELO (Assessment of Higher Education Learning Outcomes) project to investigate options for the direct assessment of student achievement that are applicable across different types of institutions, countries and languages. The initial focus of AHELO is a feasibility study of the suitability of assessment instruments developed for the disciplines of economics and engineering. In addition, the USA-developed CLA instrument is being adapted and trialled for application to the testing of generic skills.

In Australia a new regulatory body, the Tertiary Education Quality and Standards Agency (TEQSA) has now been established. TEQSA’s approach to the incorporation of graduate learning outcomes in undertaking its responsibilities will involve the work of the soon to be established Standards Panel. However, whatever approach is taken will undoubtedly have some level of impact on institutional practices in the sector. In anticipation of future directions some institutions have already taken the initiative in demonstrating capacity for taking responsibility for self-regulation in maintaining academic standards.

Issues and concerns

Unsurprisingly, as with any significant educational innovation on this scale, some concerns have been associated with the standards movement. There has also been criticism of the approach based on the experiences of other educational sectors where the introduction of ‘standards’ has generated considerable bureaucracy with little evidence of improvement in student learning experiences.

Other concerns include the possibility of standardisation of program curriculum and erosion of standards as a consequence of the implementation of threshold learning outcomes. Doubts have been expressed regarding the adequacy of traditional forms of assessment or standardised tests for the assessment of complex learning outcomes. Criticism of some forms of standardised testing has included the influence on results of a large number of uncontrolled factors such as institutional entry requirements, student motivation and a capacity to produce league tables with their potential for misrepresentation and misuse. The assumption that decontextualised standardised testing provides valid evidence of achievement of generic graduate learning outcomes is also open to challenge.

The AAGLO Project

The AAGLO project will investigate issues in the assessment and assurance of graduate learning outcomes and provide evidence-based options for practice. The analysis of assessment tasks and processes collected across a range of disciplines in national and international contexts will identify effective practice and make it accessible across the sector in the form of principles and examples.

AAGLO outcomes will be developed through an extensive consultation process incorporating the views and experience of a wide range of key stakeholders. The project team will also invite collaboration with teams from completed or current projects in related areas.

You are invited to contact the project manager if you would like to contribute information or opinion to the AAGLO project or to register your interest in being included in planned project consultation activities.

Key references


AHELO (OECD) http://www.oecd.org/document/22/0,3746,en_2649_35961291_40624662_1_1_1_1,00.html#What_is_ahelo

Tuning Europe http://www.unideusto.org/tuningeu

Tuning Latin America http://tuning.unideusto.org/tuninalg/index.php?id=0&option=content&task=view&id=168&Itemid=196&lang=en

Tuning USA. http://www.luminafoundation.org/?s=tuning

AAGLO Project Team

Simon Barrie (The University of Sydney): simon.barrie@sydney.edu.au

Clair Hughes (The University of Queensland): clair.hughes@uq.edu.au

Geoffrey Crisp (RMIT University): geoffrey.crisp@rmit.edu.au

Anne Bennison (The University of Queensland): a.bennison@uq.edu.au
AAGLO Summary 2: Assurance of graduate learning outcomes through external review

The AAGLO Project

Australia’s new regulatory body, the Tertiary Education Quality and Standards Agency (TEQSA), has responsibilities for articulating, reviewing and reporting on teaching and learning standards in Australian higher education. In anticipation of the impact of this new regulatory environment on the sector, the Australian Learning and Teaching Council (ALTC) commissioned a series of projects related to the establishment and assurance of academic standards. One of these, AAGLO - Assessing and Assuring Graduate Learning Outcomes has been funded to investigate:

- What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes? and,
- What processes best assure the quality of assessment of graduate learning outcomes?

The project employs a broad range of consultative processes and the collection of examples of practice in national and international educational contexts.

External review of graduate learning outcomes

One of the areas of interest to the AAGLO project is external review as a strategy for assuring academic standards; the term ‘external review’ here interpreted broadly to encompass practices concerned with the review, verification or moderation of the standard of work produced in taught undergraduate and postgraduate coursework. External review is receiving current attention through inclusion in a TEQSA (2011) discussion paper; several ALTC-funded investigations of forms of peer and expert review; and the Group of Eight (G8) Quality Verification System (QVS) pilot.

The UK QAA external examination system

The UK QAA External Examining system provides a well-established and well-documented example of external review with origins dating to 1832. The UK system has been closely scrutinised through much of its recent history in response to periodic questioning of the capacity of what was once a small ‘elitist’ system to maintain broad comparability of standards across a higher education sector at a time of significant expansion, structural change and greater assessment diversity. A series of reviews of the external examining system has been conducted to ensure its ongoing relevance to quality assurance in the sector, the most recent of which was reported earlier in 2011 (UUK).

In addition to consulting review and research reports, and related practices in the school sector, the AAGLO project team investigated the external examining system in a number of Scottish universities because of the strong emphasis on enhancement which permeates the Scottish QAA approach to quality assurance.

By and large, the external examining system is viewed favourably by those in the sector who associate it with a number of benefits including the evidence it provides of the sector’s willingness and ability to maintain appropriate standards without government regulation. Key strengths of the external examination system can be summarised as:

- enhancement of teaching, learning and assessment standards through a requirement for appropriate and timely senior level responses to external reports
- sharing of effective discipline-specific assessment practice among institutions
- broader perspective on teaching, learning and assessment developed by all participants
- initiation of dialogue around teaching, learning and assessment
- provision of evidence to contribute to regular internal subject and institutional reviews
- affirmation of teaching, learning and assessment strengths
- flexible arrangements to accommodate diverse institutional requirements and practices
- the leverage reports can provide in supporting curriculum leaders and teachers in arguing for assessment reform or appropriate resourcing.

Issues and options

Researchers and reviewers of the UK external examining system have identified important issues to be considered when introducing similar systems of external review elsewhere. The 2011 UUK review highlighted three broad areas of concern: the selection and appointment of external reviewers, consistency of standards and the ways in which review reports are used to enhance practice and communication with all stakeholders groups but particularly with students.

There are options for addressing these issues or concerns and those selected will depend on the outcomes intended for any particular external review scheme and the level of resourcing available.
The selection and appointment of external reviewers

External reviewers can be either appointed directly by program or course coordinators or other appropriate staff in accordance with institutional guidelines, or selected or assigned from a pool of authorised reviewers. The choice of method is largely dependent on the relative importance of factors such as the necessity to avoid the ‘cosy’ relationships that can ‘flavour’ reviewer reports and the administrative implications of establishing and maintaining an up-to-date pool of reviewers with appropriate expertise. The UUK (2011) review rejected proposals for a national register but provided in its report a set of ‘National criteria for the appointment of external examiners’. As the participation of experienced, discipline experts is essential to the validity of an external review process, administrative and remuneration arrangements and institutional recognition should provide an incentive for the regular involvement of senior academics. Broadening participation to include less experienced academics as reviewers is also considered beneficial in providing a professional development experience and in constituting a form of induction to the reviewer role and related processes (e.g. mentoring and team membership) to ensure system sustainability. In addition, broad participation has the potential to forge or strengthen discipline networks and collaboration within the sector, an advantage that to some extent balances concerns that current or future research collaborations may inhibit review reporting.

The conduct of external reviews

External review processes can vary in scope from the extremely comprehensive - program approval, verification of judgement and assessment processes, random sampling of student work, arbitration (QSA) – to ‘light touch’ processes concerned chiefly with the broad comparability of assessment practices and student outcomes.

Review reports will reflect the scope and purpose of any individual external review system. Consistency and comparability of reporting can be boosted through the provision of minimum system requirements (e.g. UUK 2011 “External examiners” report checklist”) while individual institutions are able to retain flexibility through specifying additional reporting requirements that address current priorities and ensuring that these are communicated through institutional induction.

External review arrangements also need to be responsive to the growing repertoire of assessment methods adopted by many disciplines and supported by a range of technologies. The ‘signature’ assessments of some disciplines may therefore require field visits, attendance and interactive oral presentations or access to audio and video files of student performance in clinical or work-integrated learning contexts.

Concerns have been raised about inconsistency of standards applied by reviewers in commenting on the comparability of standards among institutions (Bloxham 2009, UUK 2011). While it is unrealistic to expect guarantees of comparability from external reviewers, the use of common external reference points (e.g. the discipline-specific benchmark statements developed in the UK or the more recently endorsed ALTC threshold learning outcomes) has been proposed as a way of boosting the objectivity or trustworthiness of external reviewer judgements.

Putting external review reports to good use

Institutions also need to consider the matter of ‘closing—the-loop’ on external review reporting through developing processes for responding to both individual external review reports and any themes emerging at a program, school or institutional level. The public release of at least part of external reports has been suggested as a communication mechanism for demonstrating institutional accountability, and for dispelling the many myths about the process that can exist among student populations. Reviewers need to know how their reports are to be disseminated and advised on the inclusion/omission of staff or student names and other confidential details. Processes for raising serious concerns also need to be considered.

Finally, external review is only one of a number of possibilities for the assurance of academic standards. External review policies and practices therefore need to recognise their contributions to other institutional strategies; for example, the inclusion of review reports as formal components of subject and institutional review processes. Exemptions from external review processes might also be appropriate where they replicate information provided through alternative processes such as rigorous peer review undertaken in preparation for promotion.

Some forms of external review can have a positive impact not only on the assurance of broad comparability of standards but more comprehensively in strengthening discipline communities and in providing a worthwhile professional development experience. A TEQSA decision to take the Australian higher education sector in this direction presents a rare opportunity to open teaching, learning and assessment issues to conversations conducted on a national level.

Contribute to the AAGLO Project

This paper has mapped the territory of external review of academic standards as part of the AAGLO project. You are invited to contact the project manager if you would like to contribute further information or opinion or to register your interest in being included in planned project consultation activities.

AAGLO Project Team

| Simon Barrie (The University of Sydney): | | | | Clair Hughes (The University of Queensland): | Anne Bennison (The University of Queensland):  |
| Geoffrey Crisp (RMIT University): | | | | a.bennison@uq.edu.au | |
**Key readings**


AAGLO Summary 3: Challenges of assessing Graduate Learning Outcomes (GLOs) in work-based contexts

This is an abridged version of an invited presentation to the Informa WIL Forum in May 2011.

The AAGLO Project

The AAGLO - Assessing and Assuring Graduate Learning Outcomes project has been funded by the Australian Learning and Teaching Council (ALTC) to investigate the types of assessment practices most likely to provide convincing evidence of student achievement of graduate learning outcomes (GLOs). The project defines ‘graduate learning outcomes’ as the intended learning outcomes of a particular university program as specified by an institution or discipline.

Work as the context for assessment

Convincing evidence of student achievement in many health and educational fields has traditionally been generated through assessment conducted during practical work placements. An increasing focus on employability as a key outcome of higher education has expanded opportunities for students in additional disciplines to develop and demonstrate their learning in the world of work. However, it has been observed that “the assessment of work based learning brings into particularly sharp focus some challenges that all assessments face”; adding graduate learning outcomes to this mix intensifies the level of challenge.

Assessment in the curriculum

Assessment is only one element of an interdependent set of design elements that comprise the curriculum ‘system’ illustrated in Figure 1. While many assessment challenges are inherent to assessment itself, addressing these in isolation overlooks other challenges to assessment that originate in the assumptions and decisions that shape all elements of the curriculum. The following summary addresses those challenges of most relevance to assessment in work-based contexts.

1. Personal beliefs about graduate learning outcomes

Research reports that agreement on the nature of graduate learning outcomes is far from universal with academics’ conceptualisations ranging from basic communication skills to the transformational outcomes that shape personal and professional identity. Vastly different perspectives such as these will have vastly different implications for assessment so a fundamental challenge is:

- to ensure shared stakeholder understanding of the nature of the outcomes to be developed and demonstrated in work-based contexts.

2. External and internal influences on GLO assessment

Work-based learning and assessment are subject to a number of internal and external influences including the constantly evolving priorities, policies and practices of institutions, government, regulatory authorities such as TEQSA and professional and employing bodies. Ongoing and rapidly evolving challenges are to:

- to respond to movements in the assessment-related requirements of institutional policy and external authorities and agencies
- to balance immediate employability requirements (e.g. professional accreditation) with broader and longer-term institutional mission and goals.

Educational research has advanced both the theory and practice of assessment, graduate learning outcomes and work-based learning and consequently posed a challenge to academics:

- to draw on authoritative research as the basis for decisions related to work based curriculum and assessment development.

3. Articulation of GLOs

A major challenge for curriculum planners is to clearly articulate program outcomes and course objectives and to employ them in establishing, maintaining or restoring key points of curriculum alignment -between individual course objectives and overall programme outcomes, and, between class-based and work-based learning objectives.
Related challenges are:

- to express GLOs clearly to ensure accessibility by a range of stakeholders including students and work-based supervisors, particularly those actively involved in assessment processes
- to specify an appropriate ‘grain size’ for the articulation of GLOs - neither so general or vague as to be of little practical value in guiding curriculum development or assessment, nor so precise that they fragment the curriculum
- to avoid lengthy, reductionist and disaggregated checklists of individual workplace competencies that fail to acknowledge the complexities of GLOs and which obscure the relationship between program GLOs and work-based learning objectives

4. The assessment of GLOs
Designing plans and tasks suited to generating evidence of GLO achievement in work-based contexts intensifies the challenges inherent in more traditional assessment environments. Those challenges of most significance are:

- to devise assessment tasks appropriate to the ‘wicked’ outcomes valued in the world of work – unpredictable and relatively unbounded problems, incomplete information and requiring engagement with others
- to achieve parity of esteem between work-based and other course assessments and to uphold the fidelity and credibility of work-based assessment standards by avoiding oversimplified approaches (e.g. Pass/Fail or minimum completion)
- to require demonstrations of learning beyond the context-specific, behaviours performed routinely in the workplace
- to provide multiple opportunities for students to demonstrate GLOs and combine evidence from work and tasks completed in other contexts in defensible ways when awarding grades
- to assure the quality of assessment tasks and judgements of the quality of work completed in work contexts through appropriate policy, approval and moderation processes
- to articulate assessment criteria and standards that support consistency of judgement and whose definitional value specifies the contribution of work based learning to GLO development
- to develop feasible options for resourcing assessment undertaken in diverse locations and under conditions over which the institution has varying levels of control.

6. Teaching activities
Enabling students to benefit from work-based learning experiences involves teaching challenges:

- to systematically embed GLOs throughout the years and levels of the program
- to develop the assessment capacity of workplace supervisors and assessors
- to provide effective and timely feedback to students in work-based locations.

7. Evaluation
Timely and responsive evaluation is a key mechanism for monitoring and enhancing the quality of assessment of GLOs in work-based contexts with challenges:

- to capture the complexity of assessment
- to involve workplace mentors, supervisors and assessors meaningfully in evaluation processes
- to complement routine evaluations with additional approaches and instruments customised to the work-based learning environment.

A final challenge
Meeting the challenge of assessing graduate learning outcomes in work contexts will provide an intellectual and political journey that takes protagonists over terrain that, at times, will prove inhospitable. There is, manifestly, plenty of work to be done.

Contribute to the AAGLO Project
You are invited to contact the project manager if you would like to contribute further information or opinion or to register your interest in being included in planned project consultation activities.

AAGLO Project Team
Simon Barrie (The University of Sydney)
Clair Hughes (The University of Queensland)
Geoffrey Crisp (RMIT University)
Anne Bennison (The University of Queensland): a.bennison@uq.edu.au
References


Other recommended reading


AAGLO Project interest in standardised testing

The possibility of a role for standardised testing of both generic and discipline specific graduate learning outcomes in the Australian context has been raised in a TEQSA (2011) discussion paper and through Australia's participation in the OECD Assessment of Higher Education Learning Outcomes (AHELO) project. Standardised testing as an approach to the assurance of student achievement of graduate learning outcomes (GLOs)1 have therefore been considered by the ALTC AAGLO - Assessing and Assuring Graduate Learning Outcomes project team as one aspect of their investigation of two key questions:

- What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes? and,
- What processes best assure the quality of assessment of graduate learning outcomes?

Our investigation involved a literature review, discussion with representatives of the OECD AHELO and Collegiate Learning Assessment (CLA) project teams and consultation with members of the AAGLO reference group whose responses are embedded in this summary.

A review of standardised testing

The practice of standardised testing on a national or international scale is more common to the school sector than to higher education. Nevertheless the experience of the school sector is useful in describing practice and identifying issues that have emerged.

Morris (2011, 5) in an extensive review of standardised testing in OECD countries defines them as “tests that are designed externally and aim to create conditions, questions, scoring procedures and interpretations that are consistent across schools”. She distinguishes between standardised tests with high stakes for students – where results determine access to or graduation from an institution or program of study - and those with no stakes for students – but with possible stakes for teachers, programs or institutions if linked to funding arrangements. Decisions to implement standardised testing are attributed to five primary drivers: 1) New public management; 2) Standards based assessment; 3) International competition; 4) Increasing demand for 21st Century Skills; 5) Test industry pressure (ibid, 7).

Morris (ibid, 10) argues that “the purpose behind a standardised test should guide the rationale for the assessment and feed into the design and implementation of the test as well as steer the use of the test results”; and that inferences drawn from specific test results will not be accurate or valid for purposes other than that for which the test was designed. In addition to tests with high stakes for students’ academic or professional careers, there are four distinct purposes identified for administering standardised tests, each associated with implications for use of test results.

<table>
<thead>
<tr>
<th>Standardised tests purpose</th>
<th>Use of test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>To monitor and evaluate the education system</td>
<td>To inform policy</td>
</tr>
<tr>
<td>To hold the education system (and/or its components) accountable</td>
<td>To reward or sanctions institutions or teachers</td>
</tr>
<tr>
<td>For public information (i.e. by those outside the education system)</td>
<td>To compare and rank schools</td>
</tr>
<tr>
<td>For formative purposes</td>
<td>To identify learning needs and instructions</td>
</tr>
</tbody>
</table>

The practice of using single, large-scale, standardised tests to serve multiple purposes – e.g. the determination of national standards and the determination of rewards or sanctions for teaching quality - is considered problematic in that any increase in the ambiguity of result validity has the potential to undermine the assessment system (ibid).

Test design and development involves numerous decisions too detailed for a elaboration in this summary. They relate to the scope of the test, development of test items, test frequency and timing, sample or census-based participation, the method of analysis of test results - norm referenced, criterion referenced or growth measures - the use of ICT, implementation and scoring. There are also decisions related to if and how results are reported -are they to be publicly available? aggregated or disaggregated? adjusted to account for factors outside institutional control?

Advantages of standardised testing

Many standard tests such as the US Collegiate Learning Assessment and the Australian Graduate Skills Assessment focus on generic skills such as critical thinking, written communication and problem-solving, though the AHELO project has discipline specific tests for engineering and economics in development in addition to its adaptation of the CLA (Tremblay 2011).

1 Graduate learning outcomes: the intended learning outcomes of a particular university program as specified by an institution or discipline.
Agencies engaged in the development and implementation of standardised tests claim a range of significant educational benefits. These include the provision of comparisons with similarly situated institutions to use as the basis for a formative approach to the enhancement of students’ educational experiences and learning outcomes. AHELO also aims to inform students, governments and employers about the effectiveness of resource deployment and graduate employability (AHELO 2011, 2) while it is suggested that information reported from the Graduate Skills Assessment is used for fields of study comparisons, measures of growth, early identification of students in need of support, entry into postgraduate courses or for seeking employment after graduation.

AAGLO reference group (ARG) members expressed a number of reservations as well as commenting that:

- **Standardised testing measuring differential between skills upon commencement in schools and graduation could form part of an overall analysis of the University’s skills improvement mechanisms, both as embedded in the curriculum and in the form of other language and academic skills programs (ARG1).**

- **Introducing a standard test across the whole of Australia will allow for ease of benchmarking (ARG2).**

- **I can see a possible use of common tests for assessing small aspects of knowledge-based discipline specific learning outcomes in early foundational aspects of a higher education degree (ARG3, emphasis in original).**

To date the use of major standardised tests is voluntary, a factor that providers consider crucial to effective use.

**Issues around the use of standardised testing**

Standardised testing is a complex undertaking that has limitations and possibly undesirable consequences (Banta 2006, 2007a, 2007b; TEQSA 2011; Sadler 2011) as well as potential benefit.

**Limitations of standardised testing**

It is questionable whether generic skills can validly be assessed “in isolation from the discipline specific contexts in which students have gained their education and socialisation” (ARG4), as “very little is being written about what tests of generic skills are actually measuring and with what accuracy (ARG5)”.

Doubts have been raised regarding the risk that “if absolute student scores are not appropriately controlled, they will fail to distinguish the contribution of the actual teaching from the entry standards of the institution” (TEQSA 2011, 12). In the United States it has been reported that “institution level correlation between student scores on the tests of generic skills and entering SAT/ ACT scores is so high that prior learning accounts for at least 2/3 of the variance in institutional scores” (ARG5), and that the effects of age, gender, socioeconomic status, race/ethnicity, college major, sampling error, measurement error, test anxiety and student motivation to perform account for a proportion of the remainder. “We must argue for multiple measures of institutional effectiveness ... and ... specific guidance for improving curriculum and instruction” (ARG5, emphasis in original).

**Undesirable consequences**

There is evidence that standardised tests have led to unintended and undesirable consequences in the school sector and there is potential for their replication in higher education, particularly when student test results are used in accountability systems or published in ways that encourage unjustifiable or demoralising comparisons.

Morris (2011) reports an increase in strategic behaviours such as “teaching to the test”. This can involve both item teaching - teaching test taking skills and using test or similar items as instructional material - or organising the curriculum around test items rather than a body of content. In either case, this can inflate scores without reflecting an actual increase in student understanding and therefore provide a misleading measure of student achievement. “If this is the case then it is not the instrument itself that needs examining but the way in which is used in the higher education sector” (ARG2).

A related consequence is the narrowing of the curriculum where efforts to improve test scores will narrow instruction to those aspects of the curriculum most likely to be tested. “A perverse outcome that could result is the incorporation into curricula of generic communication or problem-solving courses for example, at the expense of skill development in areas such as a clinical communication skills or skills in specific approaches to solving engineering problems” (ARG4).

When test results have implications for institutional or teacher sanctions and rewards there is evidence of manipulation of the student population through the exclusion of low performing students from taking the tests. International research and media coverage of the Australian school NAPLAN testing has reported instances of cheating in which teachers have changed student responses, filled in blank answers, allowed additional time, or provided students with answers.

In addition “Universities draw on people from vastly different demographics students, and having standardised testing that is only done upon exit or graduation can disincenivetise universities from enrolling students from disadvantaged backgrounds or students who have attained a lower ATAR score and are, bluntly assumed to have a lower grasp of generic skills” (ARG1).

**Contribute to the AAGLO Project**

You are invited to contact the project manager if you would like to contribute information or opinion on the issue of standardised tests or to register your interest in being included in planned project consultation activities.

**AAGLO Project Team**

Simon Barrie (The University of Sydney):
Clair Hughes (The University of Queensland):
Geoffrey Crisp (RMIT University):
Anne Bennison (The University of Queensland):

a.bennison@uq.edu.au
References

AHELO Project website www.oecd.org/edu/ahelo


Collegiate Learning Assessment website http://www.collegiatelearningassessment.org/


Assurance of assessment quality

The AAGLO - Assessing and Assuring Graduate Learning Outcomes project team has conducted interviews with academics from seven disciplines in a range of Australian universities as part of their investigation of the question:

What processes best assure the quality of assessment of graduate learning outcomes?

This summary reports preliminary analysis of responses concerning internal processes for assurance of the quality of assessment of graduate learning outcomes (GLOs) – other processes such as professional accreditation and external review are dealt with separately. The examples provided fall into two broad categories; approaches to assuring the quality of assessment tasks and approaches to assuring the quality of assessment judgements. Each of these two major categories is further subdivided according to whether assurance practices are undertaken before or after assessment plans are implemented and before or after assessment judgements (e.g. ‘marking’) occur. The summary draws on selections from AAGLO interviews to illustrate the practices identified.

Approaches to assuring the quality of assessment tasks

It is evident that there is significant disparity in the assurance of the quality of tasks used for the assessment of GLOs. In some instances, academics enjoy high levels of professional autonomy with few or no restrictions placed on either the type of assessment tasks or plans developed for individual courses or on the substitution of new ones. In other contexts, assessment tasks and plans are subject to approval processes before being assigned to students and before substitutions are allowed; systematic processes are in place to ensure post-implementation evaluation.

Pre-implementation task quality assurance practice

In the first instance, policy is the key mechanism for assuring task quality prior to implementation. Interviewees referred to elements of policy as having relevance to the assurance of assessment task quality but these are more frequently in relation to task parameters such as number of tasks, word length and weighting than to general indicators of task quality or the provision of credible evidence of achievement of GLOs. Though general policy analysis is outside the scope of the project, policy-related issues that emerged from the interviews are addressed in “AAGLO Summary 6”.

Mapping is another approach to quality assurance though a whole-of-program approach to mapping of assessment tasks against GLOs is uncommon. Examples include the embedding of GLO assessment tasks throughout the program (e.g. in capstones) and coordination of the timing and variety of all first-year assessment.

Most respondents reported a requirement for some form of quality assurance prior to task implementation and many also described processes for the approval of task changes. These are undertaken at different administrative levels and with different degrees of formality. Some tasks are subjected to relatively informal peer review only and some require no approval beyond that of the course coordinator. However many institutions have established variously titled teaching and learning committees whose formal approval is to be obtained before tasks are presented to students in course/subject outlines. Such committees operate at discipline, school, faculty or institutional level and their focus ranges from general policy compliance to the detail of examinations or the closer scrutiny described below:

The Faculty of Law Education Committee (EC) reviews unit outlines at the beginning of each semester to check general compliance with provided guidelines such as the alignment of learning objectives and assessment tasks (e.g. that oral communication is not assessed through written examination). Appropriate intervention is undertaken where ‘slippages’ from approved approaches are detected. The (EC) also checks across courses for coordination and timing of assessment to ensure an appropriate spread of tasks for students and comparable weighting of assignments of similar length. A review of all assessment across programs has identified assessment types and items that are overused and encourages consideration of desirable alternatives where appropriate.

Mark Israel and Natalie Skead:
The University of Western Australia

Formal task approval can require successful passage through two or more committees. Those committees close to the point of implementation – e.g. program or discipline committees – are most likely to attend to the significant detail of assessment quality and are best positioned to take a whole-of-program approach to the assessment of GLOs.

Post-implementation task quality assurance practice

Assurance processes are also undertaken after assessment has been completed. This can take the form of informal sharing among course teams or focus group consultation with students. More formal practices generally incorporate the student satisfaction data.
collected by most universities though not all universities use the data in the same way. Task quality may be reviewed at course/unit level only or subject to more comprehensive institutional evaluation processes.

The Associate Dean (Academic) and Director of First-Year Engineering (DFYE) draw extensively on institutional student satisfaction survey ratings and electronic course profiles to gain an overall picture of assessment across first-year courses and also for cross course comparisons. They have also helped initiate a formal reflection process for all staff to identify what has gone well and what assessment (and other teaching and curriculum) improvements could be made. Reflections are provided to the Chair of the appropriate School Teaching and Learning Committee for discussion and cross school dissemination of issues and effective practice is ensured through the membership of the DFYE in all school teaching and learning committees.

Lydia Kavanagh: The University of Queensland

Institutional quality assurance policy can also require additional processes such as publication of assessment quality data, development of school and faculty action plans and inclusion of reports in cyclical formal reviews.

**Approaches to assuring the quality of assessment judgements**

The terms ‘calibration’ and ‘consensus moderation’ (Sadler in review) distinguish between pre-and post processes for assuring quality in assessment judgements.

**Calibration – pre-judgement quality assurance**

‘Calibration’, the promotion of shared understanding of standards among assessors, can take forms such as training sessions or workshops. These activities commonly draw on exemplars of student work from previous years as a basis for judgements and subsequent dialogue in relation to standards. They also support the consistent application of resources such as assessment guides or criteria and standards rubrics developed to inform the assessment process.

**Consensus moderation – post-judgement quality assurance**

Considerable moderation activity was reported with the aim of achieving consensus in GLO assessment judgements. Effective practice however was not necessarily complex or expensive:

> Staff who are marking accounting tasks are offered the opportunity to do this together in one room. This encourages dialogue through providing easy access to colleagues in a nonthreatening environment.

Graeme Wines: Deakin University

Approaches to moderation at the course level included assigning assessors to specific papers or examination items; double-assessment - all papers are routinely assessed twice or course coordinators reassess randomly selected papers; and meetings of teaching teams to consider all judgements, representative samples of each standard or borderline and disputed cases only.

The inclusion of academics from drama and performance arts disciplines in the interview sample provided an important reminder that assessment of GLOs takes other than written forms and that this has implications for the design of moderation activities. Judgements of students’ demonstrations of learning through oral presentation or dramatic performance are moderated through a variety of means which include public or open performance where industry representatives contribute to assessment judgements and also provide feedback. The notion of double-assessment also applies in performance contexts.

The convenor of courses in the Applied Theatre degree with practical assessment, sits in on at least one session of all the tutorial groups (in addition to conducting her own) and compares judgements among tutors to ensure equivalent standards. In addition, all practical work such as presentations are filmed to allow the later review of grades, particularly where they concern strong or weak work, and to support moderation.

Madonna Stinson: Griffith University

Cross-course or program moderation is mostly undertaken through comparisons of grade distributions and a consequent requirement for non-conforming course distributions to be satisfactorily explained or adjusted. In some instances a ‘normal’ distribution is the mechanism used to achieve consensus of assessment judgements.

Where calibration activities have been effective they can significantly reduce the time and effort required in achieving consensus in assessment judgements. However, opportunities for calibration and moderation are limited by a number of factors including large class sizes, high levels of casual staffing and consequent staff turnover and the difficulty in finding convenient meeting times for busy academics. Some institutions are exploring technological solutions to these problems.

**Conclusions**

An increasing emphasis on teaching and learning standards in Australia’s developing regulatory arrangements has implications for the ways in which universities collect and report evidence of student achievement of graduate learning outcomes. Preliminary analysis of AAGLO interview data indicates considerable activity in the assurance of the quality of assessment of GLOs and also suggests areas where more attention may be beneficial. These include shifts of emphasis from (1) course to whole-of-program approaches to assessment of GLOs and (2) from consensus moderation to calibration as an approach to shared understanding of assessment standards.

**Contribute to the AAGLO Project**

You are invited to contact the project manager if you would like to contribute information or opinion on the issue of the assurance of quality assessment of graduate learning outcomes or to register your interest in being included in planned project consultation activities.

**AAGLO Project Team**

Simon Barrie (The University of Sydney):
Clair Hughes (The University of Queensland):
Geoffrey Crisp (RMIT University):
Anne Bennison (The University of Queensland):
a.bennison@uq.edu.au
Reference
AAGLO Summary 6:
Assessment policy issues in the effective assessment and assurance of GLOs

The AAGLO project and assessment policy

The focus of the AAGLO - Assessing and Assuring Graduate Learning Outcomes project is the investigation of two key questions.

• What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes (GLOs)? and,
• What processes best assure the quality of assessment of graduate learning outcomes?

The project drew on literature reviews, institutional visits, interaction with teams working on related projects, and interviews with 48 academics from seven disciplines across a range of Australian universities. A number of issues concerning universities’ assessment policy emerged from the data analysis. Broader analysis of Australian university assessment policy has been undertaken elsewhere (Duck and Hamilton, 2008; Orrell and Parry 2007) however the policy issues particularly relevant to GLOs are noted in this paper.

Influence of Policy

Policy had the potential to both facilitate and limit the effectiveness of GLO assessment. Facilitation of effective practice was through specification of task design features such as alignment of assessment with learning objectives; the requirement for assurance processes relating to approval of tasks prior to finalisation of course\(^1\) plans; requirements for moderation of judgements to ensure consistency of standards; mandatory conduct of formative assessment and provision of feedback; and guidelines for ensuring comparability of reward for comparable tasks (see also AAGLO Summary 5).

However there were examples where well-intentioned efforts to prevent poor practice had resulted in the development of assessment policy with negative, unintended, consequences. There were also instances where laudable assessment policy was not implemented either because of incompatibility with other institutional policies or the absence of relevant implementation requirements, monitoring or infrastructure.

The risks to successful GLO implementation associated with addressing the issue in isolation from other relevant factors have been reported in earlier projects (e.g. The National GAP: Barrie, Hughes and Smith 2009).

The following overview further illustrates the limiting impact of unforeseen policy shortcomings and the failure to establish productive relationships between assessment policy and complementary policies and procedures.

Key policy issues related to the assessment and assurance of GLOs

Eight key policy issues were identified:

1. Fragmented program assessment design
2. Policy gaps and inconsistencies
3. Specification of standard grade cut-offs
4. Norm-referenced moderation
5. Mandatory provision of detailed criteria and standards for assessment judgements
6. Mandatory variety in assessment tasks
7. Specification of number of assessment tasks
8. Tacit approval or requirement for inclusion of non-achievement factors in grade calculations.

Table 1(overleaf) illustrates each issue, identifies its significance and suggests ways for achieving a positive policy influence to better assure the assessment of GLOs.

Conclusion

As the Australian higher education sector enters a new regulatory era, institutions will need to review and revise the quality of their arrangements for collecting convincing evidence of student learning outcomes. A key aspect of this will be a consideration of the policies surrounding assessment. There will be a need to clearly articulate the principles shaping assessment as institutional policy and ensure that local guidelines or procedures are neither overly prescriptive nor inconsistent with these principles. Assessment policies that underpin institutional quality arrangements should not only promote effective practice but also minimise the risk of unanticipated negative outcomes and be supported through appropriate implementation structures and mechanisms. Fundamental to the achievement of this are systematic and cohesive arrangements for institutional quality assurance and whole-of-program assessment.

AAGLO Project Team

Simon Barrie (The University of Sydney):
Clair Hughes (The University of Queensland):
Geoffrey Crisp (RMIT University):
Anne Bennison (The University of Queensland):
\(a.bennison@uq.edu.au\)

\(^1\)Course: A subject or unit, generally semester-long
Programme: A set of courses taken over several years

Support for this project has been provided by Office for Learning and Teaching. The views expressed in document do not necessarily reflect the views of the Australian Government Office for Learning and Teaching.
Table 1: Key policy issues in the assessment and assurance of graduate learning outcomes

<table>
<thead>
<tr>
<th>Issue</th>
<th>Background</th>
<th>Why this is an issue</th>
<th>How the issue can be addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fragmented program assessment design</td>
<td>Autonomy in course design and assessment has been a traditional feature of academic life</td>
<td>Unlikely to provide a coherent program experience for students or support the collection of convincing evidence that students have achieved minimum learning outcomes by the time of graduation (ALTC 2009; Gibbs 2009; Graff 2009).</td>
<td>Whole-of-program planning for curriculum and assessment. Formal and typically multi-level processes for the approval of assessment plans and subsequent modifications.</td>
</tr>
<tr>
<td>2. Policy gaps and inconsistencies</td>
<td>Lack of coordination for policies and guidelines when developed in different institutional sections</td>
<td>Contradictory policies and local assessment practices which do not support GLO assessment widely reported. Limited evidence of effective monitoring of local implementation of policy or enforcement of assurance of assessment quality strategies. Implication that assessment is relatively unimportant if not explicit and enforced in quality assurance policies and guidelines.</td>
<td>Assessment monitored and regulated through supportive quality assurance arrangements, e.g. requirement for non-ritualistic (Brennan 2012) action at the local school or faculty level in response to analysis and interpretation of institutional evaluative data.</td>
</tr>
<tr>
<td>3. Specification of standard grade cut-offs</td>
<td>Prevention of unjustifiably low “Pass” marks and/or inexplicable variations in grade cut-offs in the same program or school</td>
<td>Inadequacy of overall 50% “Pass” grades to provide a credible basis for assurance of multiple graduate learning outcomes. “Pass” cut-offs of 50% allow students to avoid completion of some course assessment components if their 50% has already been achieved through earlier tasks – a further challenge to the confidence with which program assessment assures the achievement of GLOs.</td>
<td>Agreement on standards or levels of achievement relating to GLOs. Decisions reached through a consideration of student work in relation to these standards rather than marks or percentages only.</td>
</tr>
<tr>
<td>4. Norm-referenced moderation</td>
<td>Mechanism for achieving comparability of assessment judgements within or across courses based on achieving a ‘normal’ grade distributions</td>
<td>Achievement of a normal distribution of grades is inconsistent with standards based assessment. Standards are the underlying mechanism for effective assessment of GLOs. Inconsistent with the goal of aligning teaching, assessment and learning objectives: university teaching should be concerned with improving overall student performance rather the ranking of students (Biggs 1999). Incorporation of ‘efficiencies’ such as statistical comparisons and rescaling of scores that bypass the need for argument, justification and other interactive processes that interviewees cited as effective change initiatives.</td>
<td>Explicit, standards-based assessment policy. Professional development activities undertaken in preparation for assessment judgements such as workshops involving judgement and discussion of exemplars (calibration) or discussion-based moderation activity integrated into the judgement process (Sadler 2012).</td>
</tr>
<tr>
<td>5. Mandatory provision of detailed criteria and standards for assessment judgements</td>
<td>Assurance of transparent and consistent assessment judgements</td>
<td>The provision of criteria and standards have a role in developing student capacity for self assessment and for supporting the assessment transparency missing from opaque, connoisseurship approaches (Ecclestone 2001). Drawbacks have been associated with an over-reliance on detailed criteria and standards (Price and Rust 1999) and arguments made for the appropriateness of more holistic approaches (Sadler 2009).</td>
<td>Supplementing criteria and standards with additional activities such as peer-assessment and dialogue, sometimes based on exemplars of work at different standards. Use of holistic approaches with detailed feedback.</td>
</tr>
<tr>
<td>6. Mandatory variety in assessment tasks</td>
<td>Assured exposure to the range of task types characteristic of the discipline. Relief from monotony of repetition of a small number of task types</td>
<td>A variety of assessment tasks is needed to assure the achievement of the diversity of GLOs. However many disciplines employ only a relatively narrow range of tasks. Conversely, too great a variety of tasks reduces scope for students to make repeated attempts at a single type with opportunities for the application of feedback to inform improvements in subsequent performance (Gibbs and Simpson 2004; Price and O’Donovan 2006; Taras 2006).</td>
<td>A whole-of-program approach to assessment that provides systematic development of task types characteristic of the discipline, including opportunities for repeated attempts and feedback provided to achieve competent or expert performance.</td>
</tr>
<tr>
<td>7. Specification of number of assessment tasks</td>
<td>Prevention of over assessment. Provision of adequate opportunities for timely feedback</td>
<td>Gathering sufficient evidence on which to base a judgement of achievement of a GLO may require more than one task. Ambiguity in ‘task’ definition: Research proposal comprising a series of components – literature review, poster, final submission with commentary on poster feedback – interpreted as a (compliant) single task or as three distinct (non-compliant) tasks. Over-assessment is a meaningless term when formative and summative assessments are balanced (ASKe 2007).</td>
<td>Rather than specify number of tasks emphasise achieving ‘confidence’ of judgement. Approval processes that draw on appropriate expertise in evaluating the quality of course assessment intentions and the flexibility to make appropriate approval decisions.</td>
</tr>
<tr>
<td>8. Tacit approval or requirement for inclusion of non-achievement factors in grade calculations</td>
<td>Grade component used to encourage tutorial attendance and participation Limits on grades for resubmitted work as penalty for academic integrity breaches</td>
<td>Trustworthy grades represent a student’s level of academic achievement. However, ‘many elements that are technically non-achievements are routinely incorporated into grades and thereby act as contaminants’ (Sadler 2010, 727). Common ‘contaminants’ can be informal such as subconscious boosts or reductions of marks to reward or penalise tutorial attendance or effort; or formal inclusions in grade calculation through the allocation of specific marks for tutorial participation.</td>
<td>Identification of alternative ways to reward or penalise attendance or participation and of ways of dealing with breaches of academic integrity that still allow for the award of grades that provide an accurate reflection of achievement.</td>
</tr>
</tbody>
</table>

Contribute to the AAGLO Project

You are invited to contact the project manager if you would like to contribute information or opinion on policy or guidelines issues or to register your interest in being included in planned project consultation activities.
References


AAGLO Summary 7: Characteristics of tasks effective in the assessment of GLOs

The AAGLO project

A key focus of the AAGLO - Assessing and Assuring Graduate Learning Outcomes project is the investigation of the types of assessment tasks most likely to provide convincing evidence of graduate learning outcomes (GLOs). The investigation has incorporated a range of strategies including 48 telephone interviews with academics from Australian universities selected in consultation with leaders of ALTC discipline projects. This enabled the compilation of a collection of assessment tasks from seven disciplines - Business, Chemistry, Drama, Engineering, History, Law and Veterinary Science.

Disciplinary traditions were evident in the types of tasks nominated by interviewees as effective in the assessment of GLOs. Report writing of the type used in professional practice in conjunction with oral presentations was commonly required of business and engineering students. Chemistry students also produced reports of laboratory experiments. Drama students engaged in performance while history and law students produced critical reviews or essays. Veterinary students were assessed through written and/or clinical examination.

There were however also examples of more ‘generic’ tasks which were often associated with the assessment of relatively underemphasised GLOs such as those related to research skills, attitudinal development and self-regulation. Tasks such as reflective journals, diaries, role plays, workshops, posters or letters of advice could be incorporated across a range of disciplines.

This summary highlights significant characteristics of tasks identified by interviewees as effective in providing students with opportunities to demonstrate achievement of or progress towards significant learning outcomes at various stages of their programs. Representative examples of specific tasks provided by interviewees (including those from a pilot project in archaeology Hughes 2012) are included with permission (Appendix A) to illustrate the key characteristics addressed in this discussion. The tasks provided were diverse and of high quality and many more could also have been included in this selection had succinctness not been a priority for the AAGLO Summary series.

Characteristics of effective tasks

The tasks described by interviewees were consistent with recommended practice contained in the general literature on effective assessment. This summary is however confined to factors that characterised the distinctive features of this particular group of tasks. References to illustrative examples are indicated by bracketed numbers.

Assessment for learning

Assessment tasks were worthwhile activities in themselves as they provided opportunities for students to develop learning as well as to demonstrate it, hence the wide acceptance of the term ‘Assessment for learning’ (Samball et al 2012). Many of the examples involved students in substantial tasks which they were unable to complete without undertaking related learning activities. Some tasks were clearly research focused (5, 11, 14) while others such as the project-based tasks engaged students not only in research but also in collaborative planning and decision-making; and performance tasks (13) where quality depended on students’ creative input and participation in rehearsal.

Relevance to professional practice

Students were frequently assigned tasks similar to those undertaken in professional practice. For example, the most common assessment type nominated, a report, was used in all but one discipline. The actual format of reports varied according to discipline e.g. project reports in engineering (1, 2), laboratory reports in chemistry (3, 4) and consultancy reports in business (9, 10) and archaeology (15, 16).

Many tasks were multi-component in that written reports were often accompanied by oral presentations which again reflected ‘real world’ practice (1, 2, 16).

Tasks such as performance in a drama (13) and preparation of a museum exhibition (12) also demonstrated relevance to professional practice.

When examinations were discussed in relation to graduate learning outcomes, the emphasis was again on relevance to professional practice as well as problem-solving (as opposed to regurgitation of facts) and the importance of thorough student preparation through class exercises and formative feedback (6).
**Authenticity of role and audience**

Task relevance was generally achieved through authenticity of role and audience. Tasks cast students in a variety of roles such as consultant (1, 2, 9, 10, 15, 16), performer (13), curator (12) or researcher (5, 7, 11, 14). In some cases there were authentic audiences for task outcomes (10): where this was impracticable the task description itself suggested an audience such as consumer association or government client for consultancy reports, a jury for the forensic expert witness report (15), a theatre audience for dramatic performance (13) and judges for the engineering competition (1). In many instances peers served as the audience as in the case of oral presentations that accompanied reports, rehearsals (13), poster presentations (7), or workshops (14). Students’ sense of an audience was further developed through requiring them to develop several versions of a text such as a report for different audiences (9).

The provision of work-based contexts for learning can greatly facilitate authenticity of task, role and audience though assessment when undertaken in such contexts does come with its own particular set of challenges (see AAGLO Summary 3).

**Student engagement**

Authenticity of tasks, role and audience is particularly effective in engaging students in their learning. However innovative assessors have demonstrated that there is also a place for novel tasks in engaging students - the making of action figures and trading cards in an archaeology course being a particularly noteworthy example (Smith and Burke 2005). Example 8 demonstrates how a range of non-traditional activities such as the production of diary entries, ministerial briefings, scripts and letters of advice could be used to effect in developing and demonstrating learning outcomes (Easteal 2008). Students need support in mastering the text types through which they are to demonstrate learning whether the tasks assigned are traditional or novel. Also task variety needs to be carefully considered at the program level to ensure the development of competence in the forms of communication valued within a discipline and to avoid fragmentation and consequent student confusion (TESTA).

**Careful design and management of group tasks**

The articulation of disciplinary learning outcomes indicates a growing emphasis on the development of team skills such as collaboration and negotiation which is reflected in the fact that a third of the tasks nominated by interviewees were group tasks. Group work can be contentious, particularly when students find it difficult to work with others and when they perceive they are not fairly rewarded for effort. Effective group task management addressed the need for skill development and mentoring (2), structuring through the allocation of team roles and responsibilities (16) and provision of class time for some aspects of the task to get students off to a positive start and to enable the early identification of possible group dysfunction (15, 16). A fundamental condition of effective task design is a requirement for students to arrive at a specific perspective, action or recommendation rather than a general discussion or summary of possibilities or options (Michaelson et al 1997). This condition was evident in each of the examples: students produced specific solutions to engineering problems (1, 2), determined a ‘best value for money’ commercial product (3), provided expert forensic opinion to a jury (15); provided advice to a council (16); created a performance (13 and prepared a museum exhibition (12).

Fair reward for effort was addressed through incorporation of peer review processes (2, 15) to moderate product grades when determining appropriate grades for individual students.

**Explicit task relationships**

While assessment tasks could be ‘stand alone’ or independent of other components of the course assessment plan, academics often built in strong relationships among individual tasks. These could be cumulative where the completion of each and every one of a series of interdependent sub-tasks culminated in a coherent outcome. For example the completion of a project could be staged through progressive submission of separate components (2), through work placements requiring negotiation of plans and evidence of the effectiveness of planned outcomes through both product (e.g. poster) and reflective account (7).

Other tasks were strongly linked in that successful completion of a task increased the likelihood of success in following tasks. Success in an exam was more likely if students had participated effectively in class exercises (6), the quality of written ‘learning chronicles’ was largely determined by participation in class ‘buzz session’ (8); the appropriateness of material included in workshops prepared for peers was influenced by the quality of the research task that preceded it (14); peer review of draft literature reviews or essays provided opportunities for students to incorporate feedback into final submissions (5, 11).

A repetitive pattern occurs when students undertake tasks such as laboratory reports several times, usually with a heavier weighting for later iterations to indicate an expectation that feedback from earlier attempts will be applied. Though not repetition of strictly identical tasks, examples from drama (13) and business (9) illustrate how expectations can be progressively increased as students complete cycles of similar activity.

Task relationship patterns are relevant to planning for feedback. While not all feedback can or should be immediately applicable within the time span of the course in which it is provided, making task relationships explicit is one way of establishing with students the nature of feedback to be provided and its intended use.

**Focus on ‘Reflection: Turning experience into learning’**

Several examples incorporated explicit reflective or metacognitive components to enhance student awareness of their learning and learning behaviours...
(Boud et al. 1985). Project reports required accounts of both process and product (1) or group reflection on learning and areas in need of improvement along with an individual learning portfolio (2); a work experience culminated in a reflective journal (8), a business project involved students in planning and evidencing their learning (10).

Active student roles
Active assessment roles for students are widely advocated as contributing to the development of student capacity for self-assessment that underpins life-long learning (Boud et al. 2010, Price et al. 2013, Stefani 1998). Interviewees reported a range of student roles in assessment which can be located somewhere between the ‘Passive’ and ‘Active’ endpoints in Figure 1.

**Figure 1: Active and passive student assessment roles**

At the ‘active’ end of this range students were given opportunities for self-directed learning when participating in internships requiring negotiation of a placement plan (7), and a proposal incorporating selected learning outcomes and the method by which they would be achieved and evidenced (10). Various tasks incorporated a reflective component (1, 2, 8, 10) and opportunities for self or peer-assessment (5, 11). Tasks that were posed as problems allowed for an open range of responses and provided choice and flexibility (3, 4, 8, 12, 13, 14, 15, 16).

**Conclusions**
All examples provided to the AAGLO project clearly demonstrate that effective practice in the assessment of GLOs exists throughout the Australian higher education sector. This explication of effective assessment practice has been developed to support the ongoing efforts of those engaged in assessment enhancement projects within and across institutions.

**AAGLO Project Team**
Simon Barrie (The University of Sydney);
Clair Hughes (The University of Queensland);
Geoffrey Crisp (RMIT University);
Anne Bennison (The University of Queensland)

**References**
Transforming Student Experience Through Assessment (TESTA) Project. The University of Winchester. Retrieved from [http://www.testa.ac.uk/](http://www.testa.ac.uk/)
Appendix A: Selected Examples of Assessment Tasks

The following tasks have been selected to illustrate significant characteristics of tasks nominated as effective in providing credible evidence of GLOs.

ENGINEERING

1. Students participate in the ‘Engineers Without Borders’ challenge which involves first year students from about 20 universities. The challenge integrates skills such as creative thinking and research, communication and teamwork by presenting students with a problem located in an authentic context e.g. the rebuilding of houses in a flooded area of Thailand, developing a region’s tourist potential or establishing an artists’ retreat. Students produce a report using a specified template to document both project outcomes and management details (meeting records, actions and progress, individual log books), make an oral presentation and, where appropriate, produce artefacts such as models (e.g. a coal burner) or pieces of software. Teams compete in a state-wide heat with winners progressing to a national final.

John Roddick, Flinders University

2. Students complete an authentic engineering design task for a client in teams of six or seven. Teams produce three written reports each of which documents progress on the design task as well as meta-level reflection on what has been learned and what needs to be improved. The first report sets out the team understanding of the task as the basis for negotiation and clarification with the client, the second is a progress report and the third is an industry-standard final report. All reports are accompanied by an oral presentation. Weighting of reports is progressively increased to maximise the application of feedback provided. Students also document individual learning in a portfolio. Group support is provided through explicit skill development and mentoring which enables timely responses to emerging problems such as social loafing. A group mark is allocated for the report with individual student marks moderated using an anonymous, online peer-assessment process.

Lydia Kavanagh (& Crosthwaite, C. 2007), The University of Queensland

CHEMISTRY

3. Students carry out a simulated consumer market research activity by analysing three brands of a commercial product. Individuals analyse one sample and use pooled class results to calculate the concentration of active ingredient in each brand. They then decide whether the brand that offers the highest concentration or the one that contains the greatest amount offers ‘best value for money’ when cost is taken into account. In their report students must also comment on other consumer aspects such as shape and colour of the bottle and the way the product is marketed. The task is developmental as students are required to reach a certain level of communication and thoroughness which may involve resubmission to address feedback provided in which case the maximum grade is a credit.

Roy Tasker, The University of Western Sydney

4. Students analyse the composition of a chemical compound using funds provided through a virtual budget based on realistic costs. They are provided with some initial information about the compound and the costs of various analytical processes but need to use their notional funding as efficiently as possible to acquire all the data needed to reach a conclusion within budget. This task develops problem solving ability and the communication skills required for laboratory reports.

Brian Yates, Jason Smith, The University of Tasmania

VETERINARY SCIENCE

5. Students select and critically review a pair of recently published papers on the physiology of pain in livestock. They must provide a draft for peer review and have an opportunity to respond to the feedback provided prior to final submission.

Rosanne Taylor (2008), The University of Sydney

6. Students draw on content from earlier modules of the course to solve problems that they may encounter in clinical practice. They are presented with six scenarios and must respond to four of these under examination conditions. This task is supported by class exercises and formative feedback provided by teaching staff and access to a discussion board that facilitates discussion and through which assistance can be sought.

Glen Coleman, The University of Queensland

LAW

7. Students complete a three-component task during an elective internship.

Part A: Negotiation with workplace supervisor of a placement plan incorporating administrative arrangements and the identification of tasks to be undertaken to develop course Learning Goals.

Part B: Poster presentation of research into topic of contemporary significance and relevance to the placement and that goes beyond a description of the law to also consider policy implications.

Part C: Reflective journal demonstrating ability to recognise themes from practice and how they impact on student development as legal professionals of the future. The content is flexible and may include communication skills, networking, career plans, and appropriateness of application of the law in practice.

Judith McNamara, Queensland University of Technology

8. A series of small group discussions called ‘buzz groups’ creates opportunities for developing learning and providing formative feedback. Students are then assessed through written ‘learning chronicles’ developed through the buzz activities in which they examine topics or issues from multiple viewpoints. ‘Chronicles’ are thought-provoking in substance and format in that students are able to go beyond traditional options such as research essays, in adopting creative text types which can include diary entries (e.g.
11. Students complete a three-part report (3000 words) either individually or in a group (in which case the word count for Part A below is increased).

**Part A:** A typical academic task - an explanation of a component of Australia’s accounting system

**Part B:** Conversion of the explanation in Part A to a brief report suited to a Board of Directors explaining particular application to their company

**Part C:** Production of a version of the same report for non expert shareholders of the company.

*Note: Task has been influenced by the ALTC ‘Achievement matters’ project and relates to Accounting Threshold Learning Outcome (TLO) ‘Justify and communicate accounting advice and ideas in straightforward collaborative contexts involving both accountants and non-accountants’.

Graeme Wines, Deakin University

10. Tourism students undertake an internship during which they design a proposal incorporating selected learning outcomes and the activities they will use to develop, demonstrate and evaluate them. The proposal is documented using timelines, Gantt charts, project/internship management details. For example a student visiting a traditional village in Fiji chose ecotourism as the project focus. After researching a similar business in Australia (Tangaloooma Wild Dolphin Resort) and skill development in how to collect evidence of their learning outcomes, two reports were produced – one containing research and specific recommendations for the community visited and another describing the outcomes of the internship experience addressing nominated criteria.

Gayle Mayes, University of the Sunshine Coast

**BUSINESS**

9. Students work in groups of 3 to produce a Forensic Expert Witness Report based on their actual excavation and laboratory analysis of material from a mock crime scene. This task encourages group interdependence by being beyond the scope of what a single student could achieve by working in isolation. Written reports must conform to legal guidelines and demonstrate forensic/legal, archaeological and osteological understandings. Reports must also be expressed in terms accessible to a non-technical audience (a jury). Groups are supported through provision of appropriate information and skill development and the availability of class time to undertake some of the required activities. A Peer Assessment Factor used to calculate each person’s individual score for the assessment helps motivate group members to carry their weight and helped hard workers feel that their efforts were acknowledged.

Patrick Faulkner and Glenys McGowan, The University of Queensland

**DRAMA**

13. In one semester, first year students are involved in three performance tasks worth 80% of final marks; an ensemble performance, a group devised performance and a script based performance. Each task involves progressive creative input from the students towards three forms of theatre performance and through participation in three different types of rehearsal processes. The lecturer works as a director or a facilitator. Two assessors are involved in the assessment of practical work.

Theatre and Drama Program, La Trobe University

14. Groups of students research a theatre style - either contemporary or heritage - (e.g. Chinese opera) to locate it within its historical or cultural theoretical context and which they relate to the school drama syllabus. They then prepare a professional development workshop to introduce the style to a group of beginning teachers, their peers. The workshop follows a conference format and a handout must be provided.

Madonna Stinson and Julie Dunn, Griffith University

**ARCHAEOLOGY**

15. Students work in groups to produce a Forensic Expert Witness Report based on their actual excavation and laboratory analysis of material from a mock crime scene. This task encourages group interdependence by being beyond the scope of what a single student could achieve by working in isolation. Written reports must conform to legal guidelines and demonstrate forensic/legal, archaeological and osteological understandings. Reports must also be expressed in terms accessible to a non-technical audience (a jury). Groups are supported through provision of appropriate information and skill development and the availability of class time to undertake some of the required activities. A Peer Assessment Factor used to calculate each person’s individual score for the assessment helps motivate group members to carry their weight and helped hard workers feel that their efforts were acknowledged.

Patrick Faulkner and Glenys McGowan, The University of Queensland

16. A group project requires students to undertake a hypothetical development commission from a government agency seeking advice regarding a carpark relocation with cultural heritage implications. In response to a letter, supposedly from the agency, students investigate legislation and safety issues relevant to this specific context. Some time into the task students receive new information that may (or may not) require changes to their preliminary findings and conclusions. Each group’s Cultural Heritage Management Plan is reported through a presentation with each team member allocated responsibility for one of five compulsory stages. A sixth team member does not present but has responsibility for overall presentation.

Annie Ross, The University of Queensland
AAGLO Summary 8:  
e-Assessment issues in the effective assessment and assurance of GLOs

The AAGLO project and e-assessment
The focus of the AAGLO - Assessing and Assuring Graduate Learning Outcomes project is the investigation of two key questions:

- What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes (GLOs)? and,
- What processes best assure the quality of assessment of graduate learning outcomes?

Though the pedagogical principles for e-assessment are the same as for any other form of assessment, the growing availability of accessible technology has expanded the options for the assessment of graduate learning outcomes.

e-Assessment and GLOs
The Report on Summative E-Assessment Quality (REAQ) was published in the UK by JISC and examined what academics thought would constitute quality in e-assessment. Academics indicated that high quality was associated with:

- Psychometrics (reliability, validity);
- Pedagogy (mapping to intended learning outcomes);
- Practical issues (security, accessibility).

The REAQ Report emphasised that e-assessment was not envisioned solely as a method of convenience for testing large student cohorts, but rather that the medium provided an additional method for designing tasks that were authentic to the learning outcomes being tested. Academics recognised that more sophisticated task types needed to be developed to evidence achievement of or progress towards many GLOs. Quantitative methods of determining validity and reliability were seen as crucial to quality assurance, but the available methods for determining validity and reliability tend to rely on task types that privilege selected response questions (for example MCQs). Academics expressed concern that additional qualitative methods for measuring quality needed to be available for general use in the online environment.

There is an extensive literature in the discipline of measuring validity and reliability, including Classical Test Theory and Item Response Theory (particularly the Rasch Model) (Baker, 2001; McAlpine, 2002; Wang, 2004). These techniques are applicable to e-assessments using selected response formats, but are less well developed for assuring quality of GLOs that require constructed response formats (Crisp, 2012). Academics tend to rely on accumulated discipline-based history about what constitutes an acceptable standard, rather than use quantitative statistical principles. The key validation tool for the majority of current assessments tends to be academic acumen rather than quantitative evidence (Knight, 2007).

There are recent trends towards using a variety of e-assessment tasks to gather more diverse evidence to inform assessment decision making. For example, the use of simulations, role-plays, scenarios, e-portfolios, blogs and serious games can provide assessors with evidence on aspects of performance that are not easy to capture using conventional assessment approaches and may be used to provide evidence of the demonstration of more complex GLOs (Crisp, 2012). The assessment of role-plays, simulations and scenarios would normally involve divergent, constructed responses where students document their reflections on their actions within a virtual environment. These virtual activities tend to assess a student’s affective capabilities rather than their acquisition of discipline content; they also tend to involve complex real world problems that do not have a prescribed solution. Role-plays, simulations and scenarios allow students to explore the complexities of an issue and the need to consider multiple stakeholder perspectives when proposing a solution to a problem.

Allowing students to manipulate data, to examine the consequences of their responses and to make informed decisions about potential solutions are all consistent with assessing advanced skill development in students, as described by the higher levels of the SOLO taxonomy (Biggs & Tang, 2007). Simulations and sophisticated digital tools allow students to construct multistructural and relational responses to questions.

E-assessment can offer new opportunities to assess 21st century skills through the design of tasks that require Web 2.0 creative activities; interactive tasks that include branching and decision points such as role plays and scenario based activities; and through the use of global communication tools, including blogs, wikis and discussions boards. Academics can use experiential and task-based assessments that include the appropriate use of virtual worlds to capture evidence of performance.
rather than the recall of information; authentic assessment activities in virtual should involve criteria relevant to student performance, rather than the recall or manipulation of content knowledge in isolation from context (Richardson & Molka-Danielsen, 2009; de Freitas & Neumann, 2009).

e-Assessment standards

Many issues specific to e-assessment are technical in terms of task compatibility across different computer devices or operating systems and staff development for academics to design tasks appropriate to the online medium. e-Assessment standards, as they exist today, tend to be designed around the construction and delivery of summative high stakes exams and the sharing of content across different operating systems, rather than about pedagogical quality. The British Standards Institute (BSI) has published a series of standards for e-learning and e-assessment:

- BS ISO/IEC 23988:2007: a code of practice for the use of information technology (IT) in the delivery of assessments;
- BS 8426: a code of practice for e-support in e-learning systems;
- BS 8419-1 and -2: Interoperability between metadata systems used for learning, education and training.

BS ISO/IEC 23988:2007 makes recommendations for the use of ICT to deliver valid, fair and secure assessments and to collect and score participants’ responses for high stakes examinations.

A global e-learning standard using XML (Extensible Markup Language, defines a set of rules for human-readable and machine-readable documents) has been published by IMS (IMS QTI). The IMS QTI (Question and Test Interoperability) specifications use standardised XML code to define how to represent assessment content and results, how these can be stored and exchanged between systems, including common learning management systems. The specifications facilitate assessments being authored and delivered on multiple platforms without needing to be rewritten. These specifications and standards tend to be more relevant to vendors of learning management systems, publishers or organisations producing e-assessment content, rather than academics.

e-Assessment guidelines

The International Test Commission published its Computer-based and Internet delivered testing Guidelines; they are designed predominantly for commercial or professional associations designing, delivering and validating e-assessments (ITC, 2005).

The Scottish Qualifications Authority has produced general guidelines for e-assessment that are designed for managers of assessment centres as well as practitioners who are looking to create and deliver e-assessments (SQA, 2007). These guidelines tend to be tips and suggestions related to good assessment practices and link to much of the existing literature.

FREMA is the Framework Reference Model for Assessment project and defines how the components of assessment, for example the services associated with the construction, delivery and recording of e-assessments, interact in an online environment (FREMA). What is interesting about the FREMA model is it defines an ontology for the e-assessment domain as well as a set of e-assessment concept maps describing the entities and processes involved in e-assessment.

The University of Dundee Policy and Procedures for Computer-Aided Assessment have been in use since 2002 and regularly revised. Although these procedures are written for a specific institution, they are generally useful because they highlight the importance of planning the pre and post components of assessment sessions, including an analysis of the quality of the questions and how well the questions related to the learning outcomes.

In Australia, the Australian Flexible Learning Framework and the National Quality Council have published a detailed guide on e-assessment for the VET sector (AFLF, 2011). The guidelines are intended to be used by AQTF auditors as a reference when evaluating the e-assessment offerings of institutions.

Conclusion

As academics become more familiar with the use of evidence-centred assessment design principles we will witness a more coherent alignment between the learning outcomes articulated for a course, the learning activities set by the academic and the assessment tasks completed by the student (Shaffer et al, 2009). This more coherent alignment will allow GLOs to be assessed by the most task type, whether that task is completed online or offline.

AAGLO Project Team

Simon Barrie (The University of Sydney);
Clair Hughes (The University of Queensland);
Geoffrey Crisp (RMIT University);
Anne Bennison (The University of Queensland)
References


BSI. http://www.bsigroup.com


FREMA http://www.frema.ecs.soton.ac.uk/overview

IMS QTI http://www.imsglobal.org/question/qtiv1p2/imsqti_overviewv1p2.html


REAQ http://www.jisc.ac.uk/media/documents/projects/reaq_finalreport.pdf


AAGLO Summary 9:
Whole-of-program approaches to assessment planning

The AAGLO project
The focus of the AAGLO - Assessing and Assuring Graduate Learning Outcomes project is the investigation of two key questions.

• What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes (GLOs)? and,

• What processes best assure the quality of assessment of graduate learning outcomes?

This investigation has incorporated a range of strategies including an extensive literature review, consultations with an international reference group, visits to institutions and interviews with Australian academics from a representative selection of disciplines.

A whole-of-program approach to assessment
A recurring and dominant theme of project findings from these activities has been the essential contribution of whole-of-program assessment approaches to the development and assurance of GLOs. This summary paper focuses on the importance of whole-of-program approaches to assessment, the options available and the challenges involved in implementation.

Though there is no definitive expression or definition of whole-of-program assessment planning, it is generally taken to mean that there is an overarching program framework to guide the development of assessment in individual courses, units or modules to ensure the systematic generation of evidence of program-level learning outcomes. Boud et al (2010) argue a position in which ‘...assessment for learning is placed at the centre of subject and program design’, and elaborate this as:

The development of a full range of graduate attributes requires a systematic approach to assessment that builds and enhances those attributes through tasks that are diverse, complementary to each other and embedded strategically throughout a program of study. Integrated whole-of-program curriculum design needs to incorporate assessment and feedback as well as learning outcomes and teaching and learning activities.

Why whole-of-program approaches to assessment? Why now?
University traditions and cultures have allowed academics considerable autonomy in the development of courses that focus more on individual research interests than on program-level learning outcomes. As a consequence, whole-of-program approaches in any form have not been the norm in designing or evaluating curriculum and assessment and ‘few faculty members teach to collectively owned goals’ (Bass 2012, 24).

While assessment is now more likely to be subject to forms of internal approval processes, such assurance processes are generally operationalised at the level of the course and tend to be concerned more with policy compliance – number and weighting of tasks – than the contribution of individual course assessment plans to overall program coherence. Accreditation processes applicable to professional programs provide tighter accountability for program design, but until recently their focus has been more on content coverage than on the credibility of evidence of student achievement of a comprehensive range of graduate learning outcomes.

There is now an international movement away from a fairly ad hoc to a more structured approach to curriculum and assessment planning. The strength of this movement can be attributed to a confluence of factors:

• a questioning of the coherence of modular degree programmes (Knight and Yorke 2003)
• changing expectations for the outcomes of a 21st century university experience
• global trends to encourage transparency and comparability
• the assessment implications of non-traditional, complex and integrated learning outcomes
• the emergence of new regulatory environments in the higher education sector.

While individual courses are able to provide students with rich and memorable assessment experiences and opportunities to demonstrate development in GLOs, the assumption that students are individually able to assimilate a series of worthwhile experiences into a coherent educational outcome has become increasingly open to challenge. Trends towards modularisation create a fragmented experience (Gibbs 2009) which requires students to assemble an assortment of courses.

1 The term ‘assessment’ is used throughout this project to refer to the making of judgements about student achievement.
to acquire a degree. Such certifications generally imply only the completion of a fixed number of disconnected fragments with little internal coherence or connected learning (Bass 2012). The acquisition of an award in such circumstances can impede the development of a stable professional ‘identity’ by influencing students to adopt chameleon-like behaviours to meet the inconsistent and sometimes contradictory demands of a succession of courses (Graaf 2009).

Expectations of a university graduate for the 21st century have undergone significant evolution from those considered appropriate for earlier times. Program aspirations for graduates have been extended beyond traditional knowledge acquisition. Cognitive and communication skills are increasingly likely to incorporate additional outcomes related to self and social development and the forms of integrated learning essential for active and informed participation in society. The graduate attribute movement which emerged in a number of countries including Australia (DEETYA 1998) was an example of concerted efforts to assure the quality of student learning outcomes. Implementation however proved problematic and while pockets of excellence demonstrated the potential of graduate attributes to influence program renewal, overall impact tended to be patchy and compliance perfunctory and without significant impact on either curriculum or assessment (Barrie, Hughes and Smith 2009).

A more recent initiative is the global standards agenda with the goal of increasing the transparency of student learning outcomes and of promoting comparability of standards within and across countries and encouraging student mobility. This agenda has spawned related activities in many countries: QAA Benchmarking in the UK: Tuning in Europe, Latin America and the USA and the LTAS project in Australia. A common feature of these activities is the articulation of benchmark or threshold learning outcomes (TLOs) to serve as the basis for a range of curriculum and assessment renewal activities and the assurance of standards.

The scope and complexity of learning articulated as appropriate for 21st century graduates has implications for both program and course-level assessment planning. Bloxham and Boyd (2007) summarise these as:

- Complex outcomes that represent an integrated combination of achievements are unlikely to be assessable within a single course
- Repeated demonstrations in differed contexts are required to assure reliability
- Students need to demonstrate progression in the quality or standard of tasks repeated across the years of a program
- Different methods are needed to achieve an appropriate balance of emphasis on formative and summative assessment throughout a program (e.g. group work, self and peer assessment)
- Program-level assessment planning reduces inefficient duplication of student support.

In Australia the inclusion of learning as well as teaching standards in quality frameworks to be monitored by the newly established Tertiary Education Quality and Standards Authority (TEQSA) has meant that whole-of-program approaches to curriculum and assessment are increasingly essential to the assurance of program coherence and the provision of credible evidence of GLOs.

As a result, projects in Australia (AAGLO and related initiatives supported by the ALTC and now OLT) and elsewhere (e.g. the Programme Assessment Strategies [PASS] and Transforming Student Experience Through Assessment [TESTA]) projects in the UK: the National Institute for Learning Outcomes Assessment [NILOA] in the USA) are focussed on the identification of practices consistent with the realisation of effective whole-of-program approaches to assessment planning and with the production of guidelines and illustrative examples.

**Different forms or approaches**

There is however no single approach to whole-of-program assessment planning. Gibbs and Dunbar-Goddet’s (2009) ‘Oxbridge’ model with its reliance on heavily weighted examinations at or near program completion illustrates a relatively rare whole-of-program assessment environment that is particular to very traditional institutions. More moderate options can be described with reference to two key characteristics - ‘the extent to which the assessment covers all the specified program learning outcomes’ and ‘weighting of the assessment in the final qualification’ (PASS Project 2012, 4). This enables the representation of five distinct approaches to program-focused assessment that the PASS project has identified (shaded in Fig. 1).

![Figure 1: Different forms of Program-focused assessment](adapted from the PASS project 2012, p. 4 with permission)

Figure 1 can also serve to locate or represent the features of additional examples such as those identified by AAGLO interviewees (2012) or Gibbs (2009).

AAGLO 1: Assessment is coordinated across all courses in a first-year program, all staff teaching into the program are aware of the assessment plans for all first year courses and use this awareness in...
supporting students to make sense of and manage their assessment requirements.

AAGLO 2: Group project assessments are embedded throughout the years of an UG program. Professional development for all staff involved in these courses is designed to ensure systematic progress in team skills and to calibrate staff understanding of assessment standards to be applied.

Gibbs (2009): A geology degree program conducts six fieldwork trips over three years with each successive trip incorporating new learning outcomes, integrating past learning outcomes and increasing learning expectations. Summative assessment is conducted only once in a final fieldwork mapping exercise.

Implementation issues

Approaches to whole-of program assessment planning
Guidelines for whole-of-program curriculum and assessment agree on the need for clear development principles and effective implementation support mechanisms for the approach to generate credible evidence of achievement of graduate learning outcomes (Moon 2002: O’Neill 2009).

The development of effective implementation mechanisms has engaged the efforts of academics in Australia and in other parts of the world. Approaches range from minor retrospective adjustments to existing program and course arrangements to a complete renewal of current structures and practices. For example, implications for course and program design identified through the TESTA (2012) project included:

- programs with longer, bigger and fewer courses to reduce the number of summative assessments and free up resources for more formative assessment
- limited task variety to support progression and more effective application of feedback
- (re)orientation of assessment to program rather than course level outcomes
- establishment of standards through self and peer assessment used in conjunction with exemplars
- rebalancing the focus of assessment throughout a program through incorporation of more pass/fail requirements in early years followed by heavily weighted integrative or capstone summative tasks towards the end of a program.

Adoption of recommendations such as these imply large-scale upheaval of program and support structures but many guidelines for whole-of-program assessment planning indicate manageable adjustments to existing arrangements rather than radical overhauls. Guidelines vary in detail, but generally have in common some form of mapping (QAA 2011a, Jenkins 1997) and incorporate a number of the elements included in Bloxham and Boyd’s (2007) six-step approach to developing an effective programme assessment strategy:

1. Articulating graduate/program learning outcomes
2. Mapping learning outcomes across a program
3. Making assessment fit for purpose
4. Providing for progression
5. Matching assessment and learning activities
6. Identifying constraints on assessment.

Challenges in whole-of program assessment planning
The authors of many of these guidelines acknowledge that research and experience have shown that this apparently straightforward sequence of steps can be far from simple when it comes to implementation. The six steps (Bloxham and Boyd 2007) frame the following summary of challenges to whole-of-program assessment planning that have been identified through the AAGLO project, various guidelines and assessment literature.

1. Challenges in developing (or adopting) graduate/program learning outcomes
   - achieving a shared conceptualisation of program graduate learning outcomes among all with program responsibilities (Barrie 2006: Knight 2000)
   - incorporating, and where necessary, integrating a range of external reference points such as qualification frameworks, national disciplinary outcomes and the requirements of professional, statutory or regulatory bodies (Bloxham and Boyd 2007)
   - establishing a common language, determining a reasonable number of outcomes, linking these to overall institutional mission or goals and articulation in a form that provides the foundation for an assessment plan (Hatfield 2009)
   - ensuring that program learning outcomes have an impact on assessment practice rather than on assessment documentation only (Moon 2002)
   - determining the appropriate application of basic or threshold learning outcomes to selective or high entry programs (AAGLO 2012)
   - operationalising programme level outcomes in course level curricula and assessment plans (QAA 2011b)

2. Challenges in mapping learning outcomes across a program
   - determining an appropriate mapping approach and tool that engages staff meaningfully the process (O’Neill 2009)
   - managing rigorous but complex mapping procedures (Sumison and Goodfellow 2004)
   - using the identification of program gaps and duplications as a basis for the assurance and improvement of graduate learning outcomes
4. Challenges in matching assessment and learning activities
   • ensuring that retrofitting outcomes to an existing curriculum is a driver of curriculum renewal rather than a perfunctory administrative procedure (Blosham and Boyd 2007; Hatfield 2009)
   • identifying key stages or courses suited for the collection of evidence of student progress towards or achievement of graduate learning outcomes.

5. Challenges in making assessment fit for purpose
   • establishing openness to the need to go beyond traditional disciplinary methods that prove inadequate for the assessment of complex learning outcomes (Hughes and Barrie 2010)
   • recognising that some outcomes (‘wicked’ competencies) cannot be reliably assessed and devising alternative ways of making information of student achievement available to stakeholders (Knight 2000: Knight and Page 2007)
   • keeping a clear focus on what is important - rather than designing overly ambitious or comprehensive assessment plans – to determine what can be and needs to be reliably assessed and free up resources to make these high stakes assessment as reliable as possible (Knight 2000)
   • providing appropriate professional development to enhance capacity of academic staff to provide effective assessment (Hatfield 2009)

6. Challenges in providing for progression
   • making best use of compulsory courses or capstone modules, especially when managing progression in programs where pathways allow choice and cross-disciplinary options (Gibbs 2006)
   • increasing expectations across the program through establishing year level guidelines for frequently assessed learning such as communications skills or identifying situations in which learning that is developed in the early years of a program is to be integrated through the completion of complex tasks in later years (Blosham and Boyd 2007)
   • ensuring stability of program assessment plans through appropriate processes for approval of modifications to assessment tasks (QAA 2011a)
   • providing opportunities for academic staff to calibrate or develop a shared understanding of standards appropriate to particular program levels in order to inform their assessment judgements (Sadler under review)
   • devising an authentic and efficient way of documenting student progress such as the utilisation of portfolios (Hatfield 2009), a practice refined over many years at Alverno College (Loaker 2000, Mentkowski 2006)

5. Challenges in matching assessment and learning activities
   • creating a process model that clearly links or aligns program learning outcomes with teaching, learning and assessment processes (Biggs, 1996: Jackson 2000: QAA 2011a)

6. Identifying constraints on assessment.
   • becoming familiar with institutional regulations regarding workload, hurdle or Pass/Fail assessment tasks, the maintenance of academic integrity, self and peer assessment (Blosham and Boyd 2007)
   • managing constraints resulting from the unintended consequences of well-intentioned policy (AAGLO Summary 6)
   • addressing sources of ambiguity in high-stakes assessment (e.g. reporting progress in aggregated forms such as marks or grades only, information overload ) (Knight 2000, Rust 2011 )
   • managing deadlines and student workload across a number of parallel courses or modules (Blosham and Boyd 2007)
   • managing staff workload through an appropriate balance of formative and summative (high-stakes) assessment (Knight 2000) and incorporating efficiencies such as in-class presentations (Blosham and Boyd 2007).

Conclusion
As for all complex tasks, ‘forewarned is forearmed’ when it comes to meeting the challenges of whole-of-program assessment. The attention the academic world is paying to these challenges has resulted in the emergence of new challenges often in the form of ‘efficiencies’ such as standardised testing and software mapping technology. However, an awareness that there are no simple solutions to the complex challenge of using assessment to enhance the student program experience and generate credible evidence of their learning outcomes (Jackson 2000) should assist in the creation of realistic processes and realistic expectations.

AAGLO Project Team
Simon Barrie (The University of Sydney): Clair Hughes (The University of Queensland): Geoffrey Crisp (RMIT University): Anne Bennison (The University of Queensland)
References


Transforming Student Experience Through Assessment (TESTA) Project. The University of Winchester. Retrieved from http://www.testa.ac.uk/
AAGLO Summary 10: The student perspective

The focus of the AAGLO - Assessing and Assuring Graduate Learning Outcomes project is the investigation of two key questions:

- What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes?
- What processes best assure the quality of assessment of graduate learning outcomes?

The student perspective on assessment and assurance of graduate learning outcomes

Previous national projects (The National GAP) have underlined the importance of engaging students in universities’ efforts to foster the development of relevant graduate learning outcomes. The data collected during AAGLO project reinforced this in relation to several aspects of assessment including the communication to students of expectations and standards and ensuring that students were aware of the rigor of their course’s assessments as demonstrated through assessment task and judgment quality assurance processes.

Much of the assessment carried out by universities is claimed to be done with the best interests of students at heart. In addition, universities make many claims about what students need, want or understand in relation to the development of Graduate Learning Outcomes (GLOs). However, students themselves sometimes observe that the conversations universities have with them in relation to their learning needs do not always do justice to what students actually say or want. In part this might reflect the limited nature of ‘conversations’ mediated by simplistic student surveys or the ‘one-off’ nature of an explanation about ‘intended graduate learning outcomes’ as opposed to an ongoing dialogue about student learning that takes place throughout a program of study.

In seeking to actively listen to and engage with students, the AAGLO project steering group included the President of the National Union of Students (NUS) in Australia. In 2011 the President contributed to the writing of AAGLO Summary 4 on Standardised testing. The current president has led the preparation of this AAGLO Issues paper.

Universities need to do more to help students understand the sort of learning that is possible and important through higher education

It was noted in AAGLO Summary 3 that ‘agreement on the nature of graduate learning outcomes is far from universal’. This is particularly true for students. Whilst students enter university expecting to learn and leave as graduates possessing a set of skills, many students lack a concrete understanding of exactly what they will take from their university degree.

Many students see assessment as a necessary evil, something that they have to do in order to pass a unit of study that will in turn enable them to graduate with a degree in their chosen field, rather than a key component of the learning process. Final exam-only units of study in particular create this mind set. Often, students do not make the connection between an assessment and the learning outcomes and skills that it develops. This can be seen, for example, in many students’ attitude towards group work. Rather than acknowledging that group work is a norm in the work place, and that balancing different abilities, knowledge, and levels of commitment to a project is a valuable skill to learn, students often complain about their fellow group members, failing to realise that learning to manage and work with a diverse team is as important a component of the task as the product of the task itself.

Assessments must be effectively aligned with the graduate learning outcomes of a particular course or discipline, and these links must be made clear to the students in language that they understand.

Students and staff need to think about assessment as part of learning, not an ‘add-on’

There needs to be a dramatic shift in the way in which assessment is viewed, both by students and staff. It needs to be understood that assessment is formative - a part of the learning process rather than simply the culmination or evaluation of it.

This change of attitude must be led by staff. Curriculum and assessment need to be structured in a way that promotes assessment for learning. Assessments should be designed in a way that will encourage individual thought and development of skills, rather than encourage memorisation.

Relevant and regularly updated assessment tasks where the links to graduate outcomes are clearly explained in
plain English are essential. While students are generally aware of the content which is being tested in their assessments, they are unlikely to have explicitly thought about the way in which this relates to higher level skills. Purely ‘fact-based’ assessments, such as weekly quizzes, fatigue students and drive memorisation rather than deep understanding of a subject area and critical thought.

Tasks that build upon previous tasks and assumed skills can be very effective in giving students the confidence that they are improving, provided that they are given meaningful and timely feedback on the preceding assessments. One of the most common problems with assessment at universities is that in many cases, students will hand in a second task before getting back their first assessment. This is very problematic in terms of assessment for learning.

One-on-one engagement with a tutor or lecturer is one of the most effective feedback practices, particularly in later years as the complexity of tasks and level of competency displayed in a variety of areas increases. This is difficult to do in the current context of funding constraints and large class sizes, but tutors scheduling time to discuss assessments individually with their students is invaluable.

In cases where such a discussion is not possible, prompt and comprehensive feedback which is clearly aligned with a set of objectives is most helpful. For example, simply assigning an essay a mark of 65 does not give a student enough information about the strengths and weaknesses of their writing as a "communication" skill to be mastered. Ensuring that marking is done in a way which enables students to understand how a marker arrived at their conclusion is important for students' ability to improve their skills.

What sorts of assessment practices are most useful in providing students with evidence that they can use in representing their learning to prospective employers?

Given that there are few employers who ask for a folio of work outside design/architecture/creative professions, it would be rare for the vast majority of students to be in a situation where they could use their assessments to prove their achievement of a particular learning outcome. As such, further investigation of assessment verification procedures would be the best way of providing evidence to employers of achievement of graduate outcomes through assessment.

In the act of applying for a job – taking the initiative, finding out about the opportunity, writing a cover letter, developing a CV, demonstrating skills and undertaking interviews etc., graduates are demonstrating their competence in many areas. There is a question of the burden of proof and upon whom that should be conferred - whilst employers should feel certain that, for example, an Arts graduate will be able to write and speak effectively, if there is a particular skill set which is highly prioritised at that workplace, it is the responsibility of the employer to set up their hiring processes in a way which enables them to make a fair and accurate assessment of any candidate's aptitude. It is not the role of universities or university assessment to create graduates perfectly moulded for the workplace, their role is to create the next generation of critical thinkers and innovators.

The extra-curricular program provided by a university has a huge effect on the kind of graduates which it produces

Many skills which would be considered Graduate Learning Outcomes are actually learnt outside the classroom. For example, it is much easier to demonstrate an ability to work in a team to an employer through a consistent involvement in extra-curricular activities rather than through group-work assessments in class. This is the unanswered question in the discussions about GLOs - to what extent do extra-curricular activities contribute to GLOs, and can universities take ownership of them or the responsibility to assess this learning?

Should universities be expected to provide opportunities for students to develop particular GLOs through clubs and societies and student organisations? Universities with a reputation for producing graduates who go on to be leaders in their fields often foster student leadership in a vibrant campus environment with plenty of opportunities for students to develop their skills through extra-curricular activities.

Engagement with staff at Australian universities during both the AAGLO project and the National GAP project consistently reinforced how valuable and important it is to hear and respond to the perspective of students in universities’ efforts to foster the development of graduate learning outcomes (Barrie, Hughes, Smith, 2010).

Research on assessment of student learning also points to the importance of involving students as active partners in university assessment (Price, M., Rust, C., Donovan, B. & Handly, K., 2013).

This paper is designed to provide one starting point for such a conversation however the challenge remains for universities to find ways to continue those conversations by embedding them in institutional assessment practices

AAGLO Project Team

Simon Barrie The University of Sydney
Clair Hughes The University of Queensland
Geoff Crisp RMIT University
Anne Bennison The University of Queensland

References
