Assessing and assuring Australian graduate learning outcomes: principles and practices within and across disciplines

Final Report 2014

The University of Sydney (lead institution)
The University of Queensland
RMIT University

Associate Professor Simon Barrie (project leader),
The University of Sydney

Dr Clair Hughes, The University of Queensland
Professor Geoffrey Crisp, RMIT University

Report authors: Simon Barrie, Clair Hughes, Geoffrey Crisp, Anne Bennison

Acknowledgements

The authors wish to acknowledge the support provided for this project by the Office for Learning and Teaching.

They also wish to acknowledge the contributions made to the project by:

Margaret Buckridge as the project’s external evaluator.

Trudy W. Banta, Claire Carney, Christine Ewan, Malcolm Farrow, Mark Freeman, Roger Hadgraft, Romy Lawson, Jesse Marshall, Donherra Walmsley, Beverley Oliver, and Catherine Vandermark as members of the project reference group.

Trudy W. Banta and Royce Sadler as keynote speakers at the AAGLO capital city fora.

James Tracy for setting up and maintaining the project website and graphic design work on project materials.

Academics from Australian Universities for so generously providing time to talk to us about their assessment practices.

Colleagues at The University of Sydney, The University of Queensland and RMIT University for their interest and encouragement.

Colleagues from related OLT projects both past and current for generously sharing outcomes and engaging with the project team.

Aliisa Mylonas for conducting some for the interviews during the data collection phase of the project.

Nicole Shepherd for her expertise and assistance with NVivo software that was used to process interview data.

Rebekah Ramsay for compiling the Endnote library of references that were identified during the ongoing literature review.
List of acronyms used

AAGLO  Assessing and Assuring Graduate Learning Outcomes
ALTC  Australian Learning and Teaching Council Ltd
AQF  Australian Qualifications Framework
ATN  Australian Technology Network
GLO  Graduate Learning Outcome
Go8  Group of Eight universities
HERDSA  Higher Education Research and Development Society of Australasia
IRUA  Intensive Research Universities Australia
ISSOTL  International Society for the Scholarship of Teaching and Learning
LTAS  Learning and Teaching Academic Standards
OLT  Australian Government Office for Learning and Teaching

(Note: The OLT is referred to throughout as the current project funding agency though the ALTC is acknowledged as source of the original grant.

National GAP  The National Graduate Attributes Project (ALTC 2009)
NILOA  National Institute for Learning Outcomes Assessment
QAA  The Quality Assurance Agency for Higher Education
RMIT  RMIT University
RUN  Regional Universities Network
TEQSA  Tertiary Education Quality and Standards Agency
TLO  Threshold Learning Outcome
UQ  The University of Queensland
US  The University of Sydney
WIL  Work integrated Learning
Glossary

**Unit of Study:** refers to a single component of a program, degree course or qualification usually taken over a semester or sometimes a year and referred to in some institutions as a unit or subject or module (adapted from AQF 2012).

**Course:** See Program

**Unit of Study Learning Outcomes:** describes the learning that students are intended to demonstrate upon successful completion of the learning activities in a unit of study.

**Graduate Attributes:** descriptions of the core abilities and values a university community agrees all its graduates should develop as a result of successfully completing their university studies (adapted from Bowden et al 2000). They are an orientating statement of education outcomes used to inform curriculum design and engagement with teaching and learning experiences at a university (Barrie 2009). They are referred to by various terms in Australian universities as graduate capabilities, graduate skills, employability skills.

**Graduate Learning Outcome:** the umbrella term used in this report to encompass the various terms and definitions used to describe Graduate Attributes in Australian universities. This term was used to emphasise the project’s focus on the learning achieved by students at the time of graduation.

**Program:** a program of learning is a degree course curriculum, training package, collection of units of study or structured workplace learning that leads to an award or qualification. (AQF 2012)

**Program Learning Outcomes:** a summary of the set of knowledge, skills and attitudes that a person should have acquired as a result of successfully completing the learning activities in a program. (Adapted from AQF 2012)

**Threshold Learning Outcome (TLO):** threshold standards, expressed as the core learning outcomes that a student of a given discipline must have achieved by the time of graduation defined by the Australian Learning and Teaching Academic Standards project.

(Un)common terminology: In producing this report the team has adopted the definition of ‘Unit of Study’ provided in the Australian Qualifications Framework (2012). However, the term ‘Course’ has been retained where used in earlier project material such as AAGLO Summary Papers or when using direct quotes from interviewees or from the literature.
Executive summary

The Assessing and Assuring Graduate Learning Outcomes (AAGLO) project brought together key perspectives in Australia and internationally, around the shared need for assessment of student learning in Universities to deliver convincing evidence to relevant stakeholder groups of achievement of stated learning outcomes for graduates. The project addressed two central questions:

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

Context: The Assessing and Assuring Graduate Learning Outcomes (AAGLO) project sought to contribute to the Federal government's plans to transform higher education in Australia through the redesign of quality assurance arrangements. The project’s focus was highly relevant given the establishment of both TEQSA and the Standards Panels in Australia in 2011. The focus on Learning Standards that emerged from the Bradley Review (DEEWR 2008) was attributed to community concerns with the consistency of standards across the higher education sector and with quality assurance arrangements that did not sufficiently consider teaching and learning outcomes. Assurance of Learning Standards will emphasise the importance of assessment in relation to agreed external reference points. It highlights the fact that assurance of the standards of graduate learning outcomes rests largely on the ability of assessment practices to deliver convincing evidence of achievement of those outcomes to relevant stakeholders.

Previous work has clearly demonstrated the importance of identifying appropriate learning outcome standards and there has been much work in Australia and internationally recently focusing on developing disciplinary lists of outcomes standards. While many such lists are now available as external reference points for developing statements of outcomes, regardless of the ‘outcomes’ included on the list, evidence of the standard of their achievement requires assessment of students’ (or graduate’s) abilities.

The last twenty years of research in the field of graduate outcomes has consistently emphasised the importance of the disciplinary or contextual expression of these abilities and the role of discipline based assessment to provide convincing evidence of their achievement will be crucial. The variability in assessment practices between disciplines is widely recognised with different disciplines often having different ‘signature’ assessments. There is also typically variability between assessment practices in the same discipline at different universities. Given the intention to support diversity in institutional offerings this is perhaps to be expected. There is also considerable variability in the processes used by different Australian higher education institutions to assure the standards applied in those assessments. The presumed variability in standards between institutions that arises from the current situation is a regular topic for public concern and comment.

Project Outcomes: The AAGLO project identified the assessment tasks in a range of disciplines that academics proposed generate convincing evidence of achievement of graduate learning outcomes. It identified the assurance process trusted by disciplines in relation to those assessments. It analysed the assessment strategies collected to identify the characteristic features of convincing assessment and assurance strategies for Graduate Learning Outcomes (GLOs). These features were summarised as a set of ‘principles’ for use by those interested in designing new assessments or making strategic decisions about which assessments are important in relation to GLOs. The assessment features, along with other key issues were identified from the literature, empirical data collection and consultations with other project teams and the expert reference group and were summarised in a set of ten key issues papers:
**AAGLO Summary 1:** The ALTC AAGLO project and the international standards agenda

**AAGLO Summary 2:** Assurance of graduate learning outcomes through external review

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

**AAGLO Summary 4:** Standardised testing of graduate Learning Outcomes in Higher Education

**AAGLO Summary 5:** Approaches to the assurance of assessment quality

**AAGLO Summary 6:** Policy issues in the effective assessment and assurance of GLOs

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

**AAGLO Summary 8:** e-Assessment issues in the effective assessment and assurance of GLOs

**AAGLO Summary 9:** Whole-of-programme approaches to assessment planning

**AAGLO Summary 10:** The student perspective

All project research publications, project resources and papers are available on the AAGLO project website at [http://www.sydney.edu.au/itl/projects/aaglo](http://www.sydney.edu.au/itl/projects/aaglo)

**Recommendations:** In addition to the resources to support engagement in the assessment and assurance of learning standards a series of key recommendations was proposed for ongoing work in this area:

<table>
<thead>
<tr>
<th>Recommendation 1:</th>
<th>That Universities review their choice of assessment tasks in programs and units of study to ensure these include design features that provide convincing evidence of achievement of GLOs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommendation 2:</strong></td>
<td>That universities provide professional learning opportunities and support for academics to develop effective discipline-based assessment for the assurance of GLOs</td>
</tr>
<tr>
<td><strong>Recommendation 3:</strong></td>
<td>That universities promote whole-of-program approaches to assessment planning through the development of guidelines and supporting systems (including technology).</td>
</tr>
<tr>
<td><strong>Recommendation 4:</strong></td>
<td>That universities examine the implementation of assessment policy with a view to identifying any unintended effects of policy on assessment practice and the extent to which other related policies limit or facilitate appropriate assessment and assurance of GLO.</td>
</tr>
<tr>
<td><strong>Recommendation 5:</strong></td>
<td>That university communities identify and implement effective strategies for appropriate and meaningful student participation in assessment design, implementation and assurance.</td>
</tr>
<tr>
<td><strong>Recommendation 6:</strong></td>
<td>That universities review and document the processes they have in place to assure the quality of tasks and judgements in unit of study and program assessments.</td>
</tr>
<tr>
<td><strong>Recommendation 7:</strong></td>
<td>That the OLT continue to support initiatives that engage academics in productive, evidence-based dialogue around the issue of standards among discipline communities on an institutional or national level.</td>
</tr>
</tbody>
</table>

**Conclusions:** A shift of emphasis from input standards to academic process and outcome standards as the focus of recent regulatory arrangements has provided the impetus for investigation of approaches as diverse as standardised discipline based tests and peer review incorporating extended discussion of student work samples in relation to specific criteria and standards. While a viable approach will require multiple components, this project has underlined the central role of academic-led, discipline based assessment in the assurance of GLOs.

A key feature underpinning the credibility of discipline-based assessment is the appropriate choice of assessment task coupled with credible assurance of the quality of the assessment task and the assurance of the quality of the judgements made on the basis of students’ performance of that task.

Agreement on the ‘purpose’ of higher education and the meaning of ‘learning’ are precursors to assuring a ‘standard’. It is clear that until academic communities (staff, students, broader society) and higher education regulators or ‘Standards Panels’, have a more nuanced and shared understanding of what it is they are assuring as GLOs from a university education, it is unlikely that assessment of those outcomes will be credible.
# Table of Contents

Acknowledgements ..................................................................................................................... 3  
List of acronyms used ................................................................................................................. 4  
Glossary ....................................................................................................................................... 5  
Executive summary ..................................................................................................................... 6  
Tables and Figures .................................................................................................................... 10  
Chapter 1: Introduction ............................................................................................................ 11  
  1.1 Overview ...................................................................................................................... 11  
  1.2 Background to the project ......................................................................................... 11  
  1.3 Scope of the AAGLO project ..................................................................................... 13  
  1.4 Project evaluation ..................................................................................................... 15  
  1.5 Organisation of the report ....................................................................................... 15  
Chapter 2: Overview of AAGLO activities and outcomes ......................................................... 16  
  2.1 Project objectives ....................................................................................................... 16  
  2.2 Project activities ........................................................................................................ 16  
Chapter 3: Assessing and assuring graduate learning outcomes: national and international agendas ................................................................................................................. 19  
  3.1 The AAGLO Project and the International Standards Agenda ............................... 19  
  3.2 National and international approaches .................................................................. 21  
  3.3 National priorities ................................................................................................... 30  
Chapter 4: Interview methodology and findings ...................................................................... 32  
  4.1 Methodology .............................................................................................................. 32  
  4.2 Findings ..................................................................................................................... 35  
Chapter 5 Dissemination ........................................................................................................... 57  
  5.1 AAGLO website ......................................................................................................... 57  
  5.2 AAGLO summary papers ......................................................................................... 57  
  5.3 Capital city fora ......................................................................................................... 57  
  5.4 Publications and presentations .............................................................................. 58  
  5.5 Impact ....................................................................................................................... 59  
Chapter 6 Evaluation ................................................................................................................. 60  
  6.1 Engagement with key stakeholders ....................................................................... 60  
  6.2 Project website ......................................................................................................... 61  
  6.3 Capital city fora ......................................................................................................... 61
6.4 External Evaluation ................................................................. 64

Chapter 7 Conclusions and recommendations ......................................................... 68

7.1 Assessment task design .................................................................................. 68
7.2 Whole-of-program approaches ...................................................................... 69
7.3 Assessment policy .......................................................................................... 75
7.4 The importance of students as partners in assessment .................................... 76
7.5 Assurance of Learning standards ..................................................................... 79

References ............................................................................................................. 80

References specific to AAGLO Summary 3 ............................................................. 88

Appendix A: Conference abstracts ......................................................................... 90

Appendix B: Links to AAGLO Summary Papers 1 – 10 ........................................... 97

Appendix C: AAGLO Interview Schedule ............................................................... 98

Appendix D: Glossary of assessment task types .................................................... 100

Appendix E: Independent evaluator’s report ......................................................... 104
Tables and Figures

Tables

Table 1: Overview of AAGLO project activities and outcomes ................................................ 17
Table 2: Representation of disciplines in data collection ........................................................ 33
Table 3: Units of study by level ................................................................................................ 34
Table 4: Tasks nominated for the assessment of graduate learning outcomes ...................... 36

Figures

Figure 1: Overview of AAGLO project ...................................................................................... 14
Chapter 1: Introduction

1.1 Overview

The Assessing and Assuring Graduate Learning Outcomes (AAGLO) project was funded in 2010 under the Australian Learning and Teaching Council’s (ALTC) Strategic Priority Project Scheme to support productive institutional participation in the Federal government’s plans to transform higher education in Australia through the redesign of quality assurance arrangements.

The AAGLO project brought together the various stakeholder perspectives in Australia and internationally, around their shared need for student assessment in universities to deliver evidence of achievement of stated learning outcomes for graduates. In the past it has been intended for such statements to be used as a basis for curriculum development and renewal. Now however, developments in Australian higher education sector regulatory arrangements indicate that some form of consistent evidencing of graduate learning outcomes will play a significant role in quality assurance processes. A new regulatory authority, the Australian Tertiary Education Quality and Standards Authority (TEQSA), has responsibility for providing independent advice on standards, quality and regulation with reference to the Australian Higher Education Standards Framework which now incorporates benchmarks for learning as well as for teaching quality.

This study used a highly consultative, appreciative inquiry approach that builds on and continues the productive conversations already taking place among disciplinary communities and the previous work undertaken through the Australian Learning and Teaching Council (ALTC) Learning and Teaching Academic Standards (LTAS) project. It incorporated insights from international research and practice on graduate attributes in Australia, and other countries. The objective of the AAGLO project was to identify how assessment can provide convincing evidence of achievement of graduate learning outcomes (GLOs) and explored both the breadth and depth of appropriate assessment and assurance strategies. The focus was on two key questions about assessment quality:

- 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?

A key objective of the project was to raise awareness and to organise active participation in debates concerning the standards agenda by establishing productive dialogue among the key stakeholder groups.

1.2 Background to the project

The AAGLO project builds on a previous investigation into the implementation of graduate outcomes in curriculum and assessment – The National GAP (Barrie, Hughes and Smith 2009). The National GAP initiative brought together key people working in the graduate attributes field as practitioners, researchers or institutional leaders in Australia and overseas to articulate the various issues around the uses of graduate attributes and to provide a framework for universities to engage in a meaningful way with the relationships among their graduate attributes and student learning activities and assessment tasks.

AAGLO sought to refresh this discussion on graduate attributes around the need to assure the quality of the assessment tasks that are used to evidence graduate learning outcomes. AAGLO is one of a number of related OLT 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 and nuanced – ‘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.

The AAGLO project was grounded in the need to support productive institutional participation in the Federal government’s plans to transform or ‘revolutionise’ higher education in Australia through the Australian Qualifications Framework (AQF) and the redesign of quality assurance arrangements ‘that puts standards at the centre of the system’ (Vandermark 2010). The renewed focus on standards that has emerged from the Bradley Review (DEEWR 2009) is attributed to community concerns with the consistency of standards across the higher education sector and with quality assurance arrangements that did not sufficiently consider teaching and learning outcomes, including discipline outcomes. The Tertiary Education Quality and Standards Agency (TEQSA), established in 2011, has audit and regulatory powers including monitoring quality and setting standards within a five-part framework which includes teaching and learning (and academic) standards.

The AAGLO project recognises that future approaches to the assurance of graduate learning outcomes and program standards will emphasise the importance of assessment in relation to external reference points such as the threshold learning outcomes developed through the ALTC LTAS project and the significant role this signals for TEQSA in implementing recommendations (DEEWR 2010, 60) that:

... the Australian Government commission and appropriately fund work on the development of new quality assurance arrangements for higher education as part of the new framework set out in Recommendation 19. This would involve:

• a set of indicators and instruments to directly assess and compare learning outcomes; and
• a set of formal statements of academic standards by discipline along with processes for applying those standards. (Chapter 4.1)

The importance of assessment in providing evidence of standards achievement for each level of authority engaged in assurance processes is also highlighted in questions posed by the LTAS (ALTC 2009, 3-5) project.

• What evidence is there to support efficient and effective assurance of academic standards in Australia?
• How can we be sure that students have achieved national minimum standards and local mandated standards?
• Is there a pattern of non-achievement of core LOs If so, what needs to be fixed in the learning objectives, the curriculum or assessment process?
• Are core learning outcomes mapped to the curriculum and student assessments?
• Are robust systems in place for ensuring the consistency, validity and reliability of assessment for the core set of learning outcomes?
The significance of AAGLO was its direct relationship to the current Australian higher education context and its maintenance of a connection with the work of the ALTC LTAS project, developments in the AQF and the establishment of TEQSA as ongoing points of reference throughout its duration. AAGLO addressed specific issues related to the assessment and assurance of graduate learning outcomes (GLOs) in the Australian context:

- issues already identified through earlier Graduate Attributes initiatives in Australia and the international standards movement
- the relationship of the LTAS TLOs with assessment
- concerns that government arrangements for the assurance of GLOs in Australia will disenfranchise discipline communities; require additional undifferentiated elements to current assessment processes; result in undesirable ‘standardisation’ of assessment with a backwash impact on curriculum; and a perception that assessment for assurance is incompatible with assessment for learning
- the possibility of over-reliance on a narrow range of assessment models and/or capstone assessments conducted in only the final year of study and the consequent missed opportunity to harness the standards agenda to broader program-level curriculum and assessment renewal and the associated challenge of implementing assessment that is consistent with the program level articulation of graduate learning outcomes.
- the need to promote a more active role for students in the assessment process.

1.3 Scope of the AAGLO project

Over an eighteen month period the AAGLO project team investigated issues and practices in the assessment and assurance of graduate learning outcomes through a range of activities including an ongoing literature review, visits to international agencies and institutions, consultation with colleagues through conference round tables. An international reference group comprising ten members who were representative of stakeholder groups was appointed to guide the project. The members of the reference group were:

- Professor Trudy W. Banta (Professor of Higher Education and Senior Advisor to the Chancellor for Academic Planning and Evaluation, Indiana University-Purdue University, Indianapolis, USA)
- Dr Claire Carney (Head of Enhancement, Quality Assurance Agency, Scotland)
- Emeritus Professor Christine Ewan (Higher Education Consultant)
- Mr Malcolm Farrow (CEO, Professions Australia)
- Associate Professor Mark Freeman (Director Office of Learning and Teaching in Economics and Business, The University of Sydney, Sydney, Australia)
- Professor Roger Hadgraft (Director, Engineering Learning Unit, The University of Melbourne, Melbourne, Australia)
- Dr Romy Lawson (University of Technology Sydney, Sydney, Australia)
- Mr Jesse Marshall (2011)/Ms Donherra Walmsley (2012) (President, National Union of Students)
- Professor Beverley Oliver (Pro-Vice-Chancellor (Learning Futures), Deakin University, Geelong, Australia)
- Ms Catherine Vandermark (Branch Manager, Higher Education Quality Branch, Department of Education, Employment and Workplace Relations)
Input from the various members of this group informed selection of keynote speakers for the capital city fora, contributed content to and editing of AAGLO Summary papers, provided advice on methodology and interpretation of analyses and guided project management.

Communication with other project teams working in related areas was also undertaken to maintain currency with the priorities of discipline communities and projects.

A major project activity was a series of telephone interviews undertaken to explore assessment practices in selected disciplines in Australian universities. Interviewees were identified by their colleagues as being actively engaged in assessment and assurance practices in their disciplines. The project sought perspectives on the types of assessment tasks that academics perceived would provide convincing evidence of the attainment of graduate learning outcomes. The AAGLO project team did not seek to prescribe specific task types for individual learning outcomes; rather the project team sought to present academics’ perspectives on which task types they would use to evidence the attainment of graduate learning outcomes within their particular discipline. The importance of discipline approaches to evidencing quality assessment was investigated throughout the project.

The overall process of the project is summarised in Figure 1.
1.4 Project evaluation

The project team incorporated both formal and informal evaluation strategies to inform their ongoing monitoring of progress. In addition, Ms Margaret Buckridge (Griffith University) was commissioned to conduct a formal, independent evaluation of the project. This evaluation was both formative and summative and undertaken through:

- a review of project materials and documentation;
- participation of the evaluator in the May capital city fora (Brisbane, Melbourne and Sydney); and
- participation in meetings with the project team.

An overview of all evaluation activities and a copy of the external evaluation are included in Chapter 6 of this report.

1.5 Organisation of the report

The project report is organised as follows:

- Chapter 2: Overview of AAGLO activities and outcomes
- Chapter 3: Assessing and assuring graduate learning outcomes: national and international agendas
- Chapter 4: Interview research methodology and findings
- Chapter 5: Dissemination activities
- Chapter 6: Evaluation
- Chapter 7: Conclusions and recommendations
- Appendices
Chapter 2: Overview of AAGLO activities and outcomes

2.1 Project objectives

The objective of the AAGLO project was to identify how assessment can provide convincing evidence of achievement of graduate learning outcomes (GLOs). The focus of the research was therefore on two key questions about assessment:

- 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?

In seeking to develop an evidence-based framework of key features of high quality assessment and assurance practices for graduate learning outcomes the project adopted a methodology which also supported its aim of raising awareness and fostering active participation in debates concerning the standards agenda. The project sought to build on the team’s networks and to further extend the many existing productive dialogues taking place within and among key stakeholder groups.

The collaborative appreciative inquiry approach (Cooperrider et al 2000) used by the project actively sought to incorporate diverse international and disciplinary perspectives and insights into effective practice. The gathering of data and the subsequent analysis of assessment tasks and processes collected across a range of disciplines in national and international contexts supported this inclusive approach. The research strategies used enabled effective practice to be identified and made accessible across the sector in the form of principles and examples.

2.2 Project activities

The AAGLO project team undertook a number of activities to investigate the two key research questions. Project activities produced a number of outputs.

The first activity undertaken by the project team was a situational analysis necessitated by the funding of a number of projects by the OLT on various aspects of graduate learning outcomes at the same time as the AAGLO project was funded. This enabled the project team to identify areas of collaboration and to finely tune the project to avoid the duplication of other projects’ activities. A summary of these related projects is available at [http://www.itl.usyd.edu.au/projects/aaglo/summaries.htm](http://www.itl.usyd.edu.au/projects/aaglo/summaries.htm). The project team maintained contact with other projects through invitations to project leaders to join the AAGLO reference group. Less formal inter-project communication strategies included discussion with project leaders and team members, attendance at each other's public events and serendipitous inclusion in the AAGLO project interview processes of academics who were also participants in other projects. A light analysis of the related projects was also undertaken to identify general trends in project activity.

Literature review and consultation with national and international colleagues through a variety of mediums including roundtable presentations at conferences (see Appendix A for conference abstracts) and visits to international centres of excellences have progressively informed project outputs. As a result of these activities, the project team has been able to locate its activities in national and international contexts and to contribute to national debates throughout the duration of the project.

A major activity, an investigation of current assessment and assurance practices at Australian universities, was undertaken through an extensive series of telephone interviews with discipline academics and leaders. Interviewees were selected to represent a range of disciplines and identified with the assistance of relevant LTAS Discipline Scholars. The interviews were exploratory but focussed on collecting data in relation to:
• effective practices in unit of study and program level assessment of graduate learning outcomes
• approaches to assessment quality assurance
• the incorporation of graduate learning outcomes in curriculum and assessment
• influential factors and practices in unit of study and program level leadership, management and governance of assessment.

These interviews constituted the major data collection activity for the AAGLO project and enabled the project team to identify effective practices both within and across disciplines at unit of study, program and institutional level.

Table 1: Overview of AAGLO project activities and outcomes

<table>
<thead>
<tr>
<th>Activities</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situational analysis</td>
<td>“Related projects” identified and documented, communication with project and institutional leaders</td>
</tr>
<tr>
<td>Literature review</td>
<td>Summary papers:</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 1: The ALTC AAGLO project and the international standards agenda</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 2: Assurance of graduate learning outcomes through external review</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 3: Challenges of assessing Graduate Learning Outcomes (GLOs) in work-based contexts</td>
</tr>
<tr>
<td>Consultation with reference group</td>
<td>Outcomes (GLOs) in work-based contexts</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 4: Standardised testing of graduate Learning Outcomes in Higher Education</td>
</tr>
<tr>
<td>Visits to international centres of excellence</td>
<td>AAGLO Summary 5: Approaches to the assurance of assessment quality</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 6: Assessment policy issues in the effective assessment and assurance of GLOs</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 7: Characteristics of tasks effective in assessment of GLOs</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 8: e-Assessment issues in the effective assessment and assurance of GLOs</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 9: Whole-of-programme approaches to assessment planning</td>
</tr>
<tr>
<td></td>
<td>AAGLO Summary 10: The student perspective</td>
</tr>
<tr>
<td>Conference roundtables</td>
<td>Endnote library</td>
</tr>
<tr>
<td></td>
<td>Conference presentations</td>
</tr>
<tr>
<td></td>
<td>Invited presentations</td>
</tr>
<tr>
<td>Participation in national debates</td>
<td>Participation in related OLT project meetings and for a</td>
</tr>
<tr>
<td></td>
<td>Submission in response to government discussion paper on the</td>
</tr>
<tr>
<td></td>
<td>Assessment of Generic Skills</td>
</tr>
<tr>
<td></td>
<td>Co-authorship of “Mapping learning and teaching standards in Australian Higher education: An issues and options paper”</td>
</tr>
<tr>
<td>Telephone interviews with discipline leaders at Australian universities</td>
<td>Overview of current assessment and assurance practice in Australian universities.</td>
</tr>
</tbody>
</table>
As the preliminary scoping of related projects and consultations evolved, and as the literature review was completed and the interviews were conducted and analysed, a series of key issues was identified as topics for a suite of ten AAGLO summary papers (listed in Table 1 and provided in full in Appendix B). The topics and the content for the summary papers reflected key areas of interest emerging from the research and enabled this progressive documentation and dissemination of the AAGLO project’s findings to directly and promptly feed into and support the debate on assessment and assurance of Graduate Learning Outcomes currently taking place in Australia. Table 1 provides an overview of major project activities and outcomes.
Chapter 3: Assessing and assuring graduate learning outcomes: national and international agendas

3.1 The AAGLO Project and the International Standards Agenda

Activities in Australia reflect international trends in the move for national governments to seek evidence of the quality assurance processes for the assessment of graduate learning outcomes. While many of Australia’s issues and concerns are distinctive to its own higher education system, we also have much in common with sectors in other parts of the world and mapping this formed the initial stage of the AAGLO investigation. Outcomes from this and other components of the project investigation were progressively recorded as a series of AAGLO Summary Papers which are available on the project website. The format of these short, generally two-page, summaries has been chosen to engage busy academics in informed discussion on topics of current relevance to the standards agenda without the necessity for extensive pre-reading. They are threaded throughout this report and the complete set is listed in Appendix B.

The first summary locates the AAGLO project within an international context.

---

AAGLO Summary 1: The ALTC AAGLO project and the international standards agenda

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 and nuanced – ‘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 Learning Outcomes in Higher Education) 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 misinterpretation 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.

3.2 National and international approaches

Further exploration of international trends and perspectives was undertaken through early project activities such as the literature review, visits to universities and other institutions such as the OECD and the Council for Aid to Education (CAE), consultation with an international reference group and the engagement of colleagues with similar interests through the conduct of round table discussions at national and international conferences. This established some common themes and approaches to graduate learning outcomes including assurance through external review, the challenges associated with assessment in work-based contexts, and the uses and misuses of standardised testing.
3.2.1 External review

The practice of external review is well-established in the UK as a means of assuring academic standards but still relatively uncommon in the Australian context. However, three ALTC Strategic Priority Projects funded in 2010 and a separate project, the Quality Verification System being trialled by the Group of Eight (Go8) are all investigating various approaches to external peer review. Summary Paper 2 provides an overview of relevant issues and practices.

---

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 (Go8) 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.

There is no single approach to external peer review of standards and though none of the four projects currently underway has finalised its reporting, they are sufficiently different in the processes they have adopted to offer contrasting perspectives on productive ways forward. In particular, discipline projects in History and Accounting are broad in scope and have a strong focus on professional learning through activities including attention to assessment design. By contrast, the QVS and the ‘Sector-wide (moderation) model ...’ (Krause et al) projects prioritise a ‘light touch’ approach. The challenge for institutions and regulatory authorities will be to find a middle ground that balances the feasibility of light touch approaches with the professional learning and constructive impact on practice resulting from opportunities for evidence-based discussion of student work samples with reference to external standards such as the ALTC LTAS Threshold Learning Outcomes.
3.2.2 Work-based contexts

Work-based contexts present particular opportunities for developing and demonstrating graduate learning outcomes, but they also present particular challenges to formal assessment and the achievement of consistency of standards.

Assessing and assuring Australian graduate learning outcomes

3.2.2 Work-based contexts

Work-based contexts present particular opportunities for developing and demonstrating graduate learning outcomes, but they also present particular challenges to formal assessment and the achievement of consistency of standards.

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 [7]
- 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 [4]
- 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 [9]

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 [10] [8]
- 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) [11] [12]
- to require demonstrations of learning beyond the context-specific, behaviours performed routinely in the workplace [11] [13]
- 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 [14] [15]
- 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 [11]
- to develop feasible options for resourcing assessment undertaken in diverse locations and under conditions over which the institution has varying levels of control.

5. Learning activities
The development of ‘knowing students’ able to undertake active participation in assessment arrangements that sustain complex learning [16] creates challenges:
- to equip students with the ability to engage productively in processes that transform the work experience into the conscious development of
Assessing and assuring Australian graduate learning outcomes such as:

- Dialogue about learning
- Self- and peer assessment
- Seeking and applying feedback
- Group and individual reflection
- Generation or recognition of evidence of graduate learning outcomes
- Representation and tracking of achievements

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 (p. 37).

3.2.3 Standardised testing

Much of what we know about the uses and misuses of standardised testing have been learned from the school sector where the practice is far more common than in higher education. While there is nothing inherently wrong with the notion of standardised testing there are sufficient examples of the consequences of misuse – misleading institutional rankings, ‘teaching to the test’, restricted curriculum, unethical behaviour and so on - to have created a high level of scepticism throughout much of the sector. This was evident in the strength of resistance to the Federal government’s proposal to adopt a version of the Collegiate Learning Assessment (CLA) for the assessment of generic skills. The proposal was abandoned, at least temporarily, as a consequence of the sector’s adverse reaction.

The issue of standardised testing was addressed by both keynotes speakers in their presentations at a series of National fora conducted in 2012 to disseminate AAGLO findings located at <http://www.itl.usyd.edu.au/projects/aaglo/summaries.htm>. Professor Trudy Banta cited examples from the USA that indicated the limited capacity of generic standardised tests to account for institutional impact, the difficulty of motivating students to produce representative work and the unfairness of tying results to funding when, by their nature, standardised tests advantage or disadvantage students in particular disciplines. Professor Royce Sadler argued the invalidity of generic standardised testing as generic attributes such as critical thinking or analysis manifested themselves in vastly different ways in different disciplines and concluded that standardised testing had “Not been shown (to my knowledge) to produce improvements in real educational outcomes in any country, educational system or sector, in 60 years of attempts.”

December 2011

AAGLO Summary 4:
Challenges of assessing Graduate Learning Outcomes (GLOs) in work-based
Assessing and assuring Australian graduate learning outcomes contexts

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) 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:

1. What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes? and,

2. 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 sanction 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 To monitor, support or advocate</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

---

1 Graduate learning outcomes: the intended learning outcomes of a particular university program as specified by an institution or discipline.
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 an 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).

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, socio-economic 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 disincenivtise 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).

---

Standardised testing continues to be used widely in the school sector in Australia and continues to provide evidence of how controversial this practice can be.

### 3.3 National priorities

Eleven projects and four fellowships funded by the OLT were identified as being active in related areas during the period of the AAGLO project. Together they represent a point in time overview of Australian standards-related activity. They can be considered indicative of national priorities as submissions were led by high-profile institutional or discipline leaders, an approach favoured by the OLT as most likely to result in effective dissemination and implementation in practice, and selected for funding by panels of peers.

The projects have can be clustered in relation to:

1. articulation of discipline or major threshold learning outcomes (minimum standards)
2. resources to support discipline-level curriculum and assessment development/renewal based on discipline-specific learning outcomes
3. resources to support institutional/program level curriculum and assessment mapping based on program learning outcomes
4. assessment enhancement to generate credible evidence of student achievement of standards
5. assurance of standards (minimum and/or full range) through external peer review.

As the literature and examples cited in this section demonstrate, countries in many parts of the world have similar agendas when it comes to the assessment and assurance of graduate learning outcomes and academic standards.
Chapter 4: Interview methodology and findings

The major project activity was a series of telephone interviews with academics from Australian universities. This chapter outlines the methodology adopted and selected findings. The interviews generated a large amount of data which have been stored in a format that allows ongoing analysis and interrogation. Only findings of direct relevance to the interview questions are reported here.

4.1 Methodology

4.1.1 Selection of participants

Interview participants were selected to ensure representation of each of the demonstration discipline areas that were the focus of the ALTC Learning and Teaching Academic Standards (LTAS) project and also to ensure that those interviewed represented a range of university types and locations.

The disciplines chosen were:

- Business - Business, Management, and Economics
- Chemistry – Science
- Drama and performance studies - Creative and Performing Arts
- Engineering - Engineering and ICT
- History - Arts, Social Sciences, and Humanities
- Law - Law
- Veterinary Science- Health, Medicine and Veterinary Science

A purposive sampling strategy was used to compile a list of potential participants with an understanding of and active role in disciplinary learning, teaching and assessment. Purposive sampling is often associated with qualitative research methodology and involves a strategic decision by the researcher as to the ability of the data source to provide expert or informed or specific information or feedback about the topic or issue under consideration (Teddie and Yu 2007). The initial starting point was the list of academics who had contributed to LTAS projects in the selected disciplines, for example as members of project teams or of the discipline reference group. The initial list of potential participants for each discipline was sent to the respective Discipline Scholar along with information about the AAGLO project and a request to recommend additions and/or substitutions. This consultation with LTAS Discipline Scholars enabled the lists to be refined and added to, as appropriate.

The inclusion of participants in senior positions such as Dean and Associate Dean enabled the collection of information on issues, policies, practices, and resources at the institutional, program as well as unit of study level.

4.1.2 Recruitment of participants

Following ethics committee approval for the research, invitations to participate in an interview, along with a participant information sheet and consent form were sent via email to ten academics in each discipline at a cross-section of Australian universities. As only seven Australian universities offer Veterinary Science, invitations were sent to an academic at each of these institutions.
In order to ensure adequate representation from each discipline, academics who were unable to participate in the interview process were invited to nominate a substitute from their own institution. Additional invitations were sent to nominated substitutes which resulted in a total of eight-five invitations being sent to academics in 33 universities. The interview team followed up invitations with reminder emails and/or phone calls and forty-eight academics accepted the invitation to participate in an interview. Table 2 lists the disciplinary spread of invitations sent and interviews completed.

Table 2: Representation of disciplines in data collection

<table>
<thead>
<tr>
<th>Discipline</th>
<th>No. of invitations sent</th>
<th>No. of interviews completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Chemistry</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Drama and Performance Studies (Drama)</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Engineering</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>History</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Law</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Veterinary Science (VetScience)</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>85</strong></td>
<td><strong>48</strong></td>
</tr>
</tbody>
</table>

4.1.3 Demographic information

Academics who agreed to be interviewed represented 26 universities across all states, the Australian Capital Territory, and Northern Territory. The interview sample included staff from each of the Group of Eight (Go8), staff from three universities in the Australian Technology Network (ATN), staff from five of the Innovative Research Universities of Australia (IRUA), and from four of the Regional Universities Network (RUN). Although the majority of universities were large (>20,000 students) metropolitan universities, the sample included academics from seven small universities and six universities located outside capital cities.

The 30 male and 18 female participants held positions at a range of levels of responsibility within their institutions. Many had multiple, overlapping responsibilities – these included senior leadership responsibilities such as Dean or Associate Dean (15 respondents), program (12) or unit of study (36) coordination responsibilities, and teaching responsibilities in one or more units of study (41).

Reflecting the purposive sampling approach adopted, many interview participants were, or previously had been, involved in disciplinary initiatives relating to assessment and standards. Projects represented among the participants included the LTAS projects (17) or other national OLT projects (10) and three LTAS Discipline Scholars. In addition four participants had participated as a reviewer in the Group of Eight (Go8) Quality Verification System (QVS) project, two had participated on other external review initiatives and four participants had been or were currently members of disciplinary accreditation panels. Many participants had published accounts of their teaching practice and provided these as additional information to the AAGLO project team.

4.1.4 Interview process

Telephone interviews were conducted by either one of the members of the AAGLO project team, the project manager, or an additional interviewer. All interviewers were experienced researchers with considerable assessment expertise. Interview questions sought information about assessment and assurance processes at unit of study, program, and
institutional level. Participants were asked to frame their responses in the context of a unit of study they taught. An interview schedule was developed by the project team based on the outcomes of an earlier case study (Hughes 2012), to ensure that all interviews covered essentially the same ground (Appendix C).

Information sent to participants proposed an indicative interview time of 30 minutes but most interviews were considerably longer with some extending to an hour or more. Not all interviews addressed all items on the interview schedule as some participants were unable to comment on all items, one interview was unexpectedly shortened and unable to be resumed because of participant circumstances and one senior administrator with no unit of study responsibilities provided responses on selected questions only. All responses were included in the analysis.

Extensive notes were taken during the interviews but they were neither recorded nor transcribed as the focus was on collecting data on breadth of practice rather than detailed linguistic or content analysis.

4.1.5 Unit of study contextual information

Participants were asked to base their initial responses on a unit of study that they either taught or with which they were familiar. These units of study were at both undergraduate and postgraduate level (see Table 3) with the majority (35) being a compulsory part of the program to which the unit of study contributed. Three were final year capstones, four involved work placements or internships and two involved participation in cross-institutional competitions.

Table 3: Units of study by level

<table>
<thead>
<tr>
<th>Level</th>
<th>Business</th>
<th>Chemistry</th>
<th>Drama</th>
<th>Engineering</th>
<th>History</th>
<th>Law</th>
<th>Vet Science</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Level 2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Level 3</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Level 4+</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Masters</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>44</td>
</tr>
</tbody>
</table>

4.1.6 Analysis

The notes taken during each interview were uploaded to NVivoVersion 9 QSR International, 2010) for data analysis undertaken with the assistance of a consultant with expertise in using the program. NVivo was selected to provide secure storage of the data and to enable ongoing interrogation as further questions of interest arise.

An earlier study conducted by Hughes (2012) provided a starting point for coding qualitative data. Preliminary coding was undertaken by the project team and the project manager in collaboration with the NVivo Version 9 expert (the analysis team). As analysis progressed, one participant spoke about a doctorate program, two spoke generally about second and third year courses and one had no direct involvement with any courses.
an inductive approach was taken to the development of final coding categories, where categories were derived from a close reading of the data (Mason, 2002). Each member of the analysis team used the data to generate a list of categories for responses to all questions. During initial meetings of the analysis team agreement was reached on the main categories, and all the interview notes were then coded to these categories. A final check of coding was conducted in order to assure accuracy and trustworthiness.

Use of the NVivo program allowed the analysis team to interrogate the data for basic information such as the key demographic characteristics of the research participants - discipline or position held in the university (e.g. senior administrator or lecturer) or number of group projects nominated by participants. The program also enabled the analysis team to look for connections in the data, such as associating assessment tasks with particular disciplines or particular types of graduate learning outcomes (GLOs).

4.1.7 Ethical considerations

Ethics approval for data collection was obtained from all three universities participating in the AAGLO project - The University of Sydney, The University of Queensland (UQ), and RMIT University.

Data were collected, coded and stored according to the conditions stipulated in the ethics approval submission. Examples included in project outcomes have been de-identified and any practices included in reports were edited to remove identifying features prior to their inclusion in any published work. Where examples have been identified, this is with permission of academics and appropriate acknowledgement has been made. In such instances academics were given opportunities to approve the wording of examples used and to request revisions if needed.

A final ethics completion report was provided to The University of Sydney as the lead organisation for the research.

4.2 Findings

Interview findings are reported in relation to:

- Tasks used for the assessment of GLOs
- GLO priorities in current assessment practice
- Assurance of task and judgement quality
- Monitoring of student progress
- Factors influencing quality

4.2.1 Tasks used for the assessment of GLOs

Participants identified a broad range of task types as providing students with opportunities to demonstrate progress towards or achievement of the significant learning outcomes expected on graduation. Though the selected tasks were broadly representative of those in regular use in Australian universities (Appendix D) those most frequently mentioned were reports, critical reviews or essays, oral presentations and various forms of tutorial or rehearsal activities.

As could be anticipated, particular disciplines demonstrated a preference for particular or ‘signature’ forms of assessment (Table 4). For example, industry or consultancy type reports were favoured by business, chemistry and engineering; research essays or critical reviews were mostly associated with history and law; performance with drama and so on.
Table 4: Tasks nominated for the assessment of graduate learning outcomes

<table>
<thead>
<tr>
<th>Task Type</th>
<th>Businesss</th>
<th>Chemistry</th>
<th>Drama</th>
<th>English</th>
<th>History</th>
<th>Law</th>
<th>Vet Science</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Crit. Review/essay</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Oral presentation</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Tutorial/rehearsal</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Reflective piece</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Examinations</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Performance</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Work placement</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Working demo</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Multicomponent task</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>17</td>
</tr>
</tbody>
</table>

However the strength of these associations can owe as much to disciplinary traditions as to pedagogical considerations and it is not difficult to imagine instances in which it would be quite appropriate to assign non-traditional tasks for student completion provided alignment with learning objectives was maintained. Tasks that could be considered ‘generic’ in that they were applicable across a range of disciplines included assessed peer reviews, case histories, video clips, posters, workshops developed for peers, personal development portfolios, (school) work programs, diary entries, ministerial briefings and conversation scripts. These largely constitute the ‘Other’ category in Table 4 as they were generally nominated by only a single participant.

The frequency with which some task types were mentioned caused some surprise among participants at the AAGLO National fora series. For example, examinations did not feature largely in the interview discussions, but this was not because participants did not use this form of assessment – many did - but because they preferred to discuss other types of tasks in relation to the generation of evidence of graduate learning outcomes. Conversely, group tasks were used more frequently than expected. Approximately a third of all nominated tasks (16) was a relatively high proportion for this sample size given the widely-acknowledged difficulty of addressing some group work issues such as social loafing and equitable assessment.

4.2.2 Task characteristics

Though the tasks discussed in interview conformed to many of the guidelines for effective assessment practice contained in the literature, it was possible to identify some features or factors that help to explain why participants had nominated them as being particularly appropriate for the assessment of graduate learning outcomes.
In the first instance, tasks developed student learning rather than only providing opportunities to demonstrate learning that had already occurred. ‘Assessment for learning’ (Samball et al 2012) was evident where students were required to continue developing their understanding during the process of task completion by engaging in research, data collection and analysis, collaboration or reflection.

Many tasks were relevant to real world professional practice. Examples included the production of industry-type reports in engineering, chemistry or business through engagement in authentic processes such as negotiation with clients and other team members, working within budget restrictions in producing outcomes and producing texts suited to targeted audiences. Engaging students in practices distinctive to a discipline was a further way of ensuring task authenticity so even when students engaged in similar sounding tasks such as ‘research’, the form of research and sources of data – literature review, laboratory analysis, analysis of primary sources – reflected the practices of a specific discipline.

Task authenticity was effective in engaging students in their learning but engagement was also planned through the corporation of novel tasks into unit of study assessments. Diary entries, ministerial briefings and conversational scripts were some examples of less traditional tasks.

Though initially asked to nominate a single task, a significant number of participants (17) argued that adequate understanding of this task would be achieved only if it could be contextualised in relation to one or more of the other components of the unit of study assessment plan. What we have termed ‘multicomponent’ tasks (in Table 4) required students to develop responses to a set of related tasks, and often the employment of different communication modes: for example, a written project proposal with feedback provided and integrated into a subsequent written report which was then communicated as an oral presentation. This type of multicomponent task, most common in business, engineering and law, was categorised according to patterns and strength of task relationships identified in an earlier pilot project (Hughes 2012).

It was evident too that the design and management of group work were given careful attention. Tasks promoted student interaction as they were generally beyond the scope of what could be achieved by a single student working alone. They also required group consensus on original perspectives, decisions or a recommended course of action. Some interviewees spoke specifically of how well-suited such tasks were to the development of group skills, the monitoring of group progress and the assessment of group products and processes.

Reflective tasks were used in combination with tasks such as group projects, work placements and performance in order to raise student awareness of the learning that had occurred and to discourage merely passive learning. Students were also required to take an active role in their learning through opportunities for flexibility and choice and even to sometimes select, develop and demonstrate learning outcomes by exhibiting independence and self-regulation.

The characteristics of effective tasks are further discussed and illustrated in AAGLO Summary 7.
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
Assessing and assuring Australian graduate learning outcomes

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.

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.

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
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

BUSINESS

9. 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 (Tangalooma 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

HISTORY
11. Students research an aspect of world history and submit a draft essay for peer review. They provide feedback on two draft essays and then have an opportunity to revise their own essay prior to final submission. Their essay and the two peer reviews are assessed.

Paul Turnbull, The University of Queensland

12. Students work in groups of 3-4 to prepare a museum exhibition based on a tutorial theme with a real-life context, such as the debates about conscription in WW1 or the acceptance of asylum seekers. The group decides an exhibition format (or various formats for different sections of the exhibition) to engage an audience - game, role play, objects, images, contemporary dance, etc. with the stipulation that text be used sparingly and that the presentation promote active learning.

Paul Sendziuk, The University of Adelaide

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

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 helped 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 car park 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*

---

**e-Assessment**

The increasing availability of educational technology has greatly expanded the options for assessment design. It was therefore of interest that though approaches such as ePortfolios or online quizzes were mentioned by a small number of academics, interviews provided no examples of electronic media being used as the main platform for assessment of GLOs. Online technology was used in support of one project’s moderation processes and several participants spoke of their hopes for forms of technology that would assist in the management of other assessment-related processes such as mapping where and how learning outcomes were addressed throughout a program and tracking student progress in their development. However, there was very little mention of the role of technology in providing opportunities for students to generate evidence of learning. e-Assessment still appears to be an approach whose adoption is dominated by enthusiasts with limited take-up by mainstream academics.

While effective e-Assessment design conforms to general assessment principles, it also involves additional considerations particular to the media used. AAGLO Summary 8 addresses recent developments in the application of technology to assessment and also outlines standards associated with effective delivery.

---

**AAGLO Summary 8:**

**e-Assessment in the effective assessment and assurance of GLOs**

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

1. **What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes (GLOs)?**
2. **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 and Molk-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 appropriate task type, whether that task is completed online or offline.

**4.2.3 GLO priorities in current assessment practice**

The types of learning assessed through nominated tasks privileged some learning outcomes while neglecting others. Various cognitive outcomes such as critical thinking, analysis, problem-solving and research skills formed the largest category of learning outcomes addressed through identified tasks. Communication skills, mainly written but also including oral were the next largest group of learning outcomes targeted. Teamwork, task, project and self-management, were also fairly prominent amongst the learning outcomes discussed. Though this was relatively unsurprising given the high proportion of group tasks that were
nominated, outcomes such as teamwork were however not an assessment focus for all group tasks.

By contrast, outcomes such as ethical development, intercultural competence and social responsibility that are increasingly emphasised in more recent initiatives such as the LTAS project were rarely addressed. Neglect of this type of ‘wicked’ or ‘fuzzy’ (Knight and Page 2007) outcome is widely attributed to the significant assessment challenges they pose and is discussed in some detail in the assessment literature (Hughes and Barrie 2010). The persistence of this neglect can also be explained in part by quality assurance processes which focus on superficial policy compliance, generally at the level of the individual unit of study (AAGLO Summary 6) and also to approaches to program design that allow considerable autonomy to unit of study coordinators in the selection of the learning objectives to be demonstrated through assessment. Over a third of participants reported personal knowledge and experience as the basis of unit of study learning objectives and, while over half mentioned institutional graduate attributes, the selection of these for individual units of study was not always guided or monitored through coordination or assurance processes.

Programs that were subject to external accreditation were an exception here as professional accreditation was a high-stakes process with severe consequences for institutions unable to demonstrate systematic mapping of individual units against accreditation requirements. While in the past these processes had focused largely on inputs such as coverage of essential content and topics, participants reported a growing emphasis on the review of samples of student work by accreditation panels when seeking evidence of the quality and standards of graduate learning outcomes.

As several examples in AAGLO Summary 7 demonstrate, minor adjustments to existing tasks extended the range of GLOs that could be incorporated into assessment plans. For example business students were required to produce reports for nominated audiences, both technical and non-technical, instead of generic academic reports on unit topics. Engineering students complemented formal, industry-type reports with team reflections and commentaries on group processes as well as portfolios of individual achievement. Law students directed and reported on individual learning undertaken during an internship. As well as expanding the scope of the learning assessed through individual tasks, approaches such as these enhanced student awareness of their own learning processes and helped to establish foundations for lifelong learning.

4.2.4 Assurance of task and judgement quality

Pre and post implementation processes
Processes in place to assure the quality of both assessment tasks and assessment judgements occurred both before and after tasks were completed by students or judgements were completed by assessors. The processes are outlined in AAGLO Summary 5.

AAGLO Summary 5:
Approaches to the assurance of assessment quality

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 nontreatening 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.

Comprehensive quality assurance systems incorporated activities at several strategic points in the assessment process:

- approval processes for new or significantly revised assessment tasks that generated feedback beyond policy compliance and which were consequential in that the application of feedback was monitored
- task review and evaluation processes which involved all with assessment responsibilities in the consideration of relevant data - samples of student work, student responses to institutional satisfaction surveys, grade distributions and personal reflections
- induction processes during which assessors considered samples of student work from previous units of study in order to develop shared understandings of or calibrate the standards to be applied to assessment judgements and to ensure consistency of information provided to students by all teaching staff
- moderation activities to engage assessors in discussion of student work standards in relation to consistent and clear reference points.

Opportunities for sharing perspectives, practices and problems in discussion with peers was universally acknowledged by all interviewees as the origin of any significant assessment enhancements that they had experienced or observed.

External quality processes
Several participants were able to report on their experience of accreditation and (to a lesser extent) external review quality assurance processes. Normally assessment was only one component in the broad approach to quality assurance taken by professional accreditation panels but could vary greatly in the extent to which it was emphasised from panel to panel, even within a discipline. In the majority of cases assessment had only a minor role and sometimes no role at all. Seven participants reported significant attention to the review of the standard of student work demonstrated in requested samples assessed at ‘pass’, ‘medium’ and ‘high distinction’ levels while one panel had requested samples assessed as ‘minimum pass’ standard only. Where accreditation was thorough it was held to exert considerable influence on program standards.

*Accreditation standards are the glue that holds the program together. Engineers Australia (EA) looks at all assessment and inputs for accreditation every 4 to 5 years.*
Looking at student work is part of the process – each course submits examples of student work. It’s a massive process. EA upholds standards so there is no need to TEQSA to drive standards as well. Professional accreditation processes are a good way to maintain standards.

The interview sample serendipitously included a small number of participants in concurrent external review processes. Two of these - an OLT ‘Sector-wide (moderation) model’ (Krause et al) and the Quality Verification System pilot conducted by the Group of Eight universities were large-scale activities undertaken across several disciplines and institutions. Another two were smaller, discipline-specific OLT-funded projects with a broader remit that allowed backward mapping from peer review activities to address areas such as task design to enhance alignment with specific disciplinary threshold learning outcomes. Though these projects are yet to report their conclusions, at this stage it appears that promising ways forward for peer review of standards will combine appropriate aspects of both – ‘light-touch’ feasibility and opportunities for face-to-face discussion of student work with reference to national disciplinary TLOs- while avoiding the risks of perfunctory compliance, inconsistent standards and impractical resourcing implications.

4.2.5 Monitoring of student progress

There were very few examples of monitoring or recording individual student development of GLOs progressively throughout the years of a program. Common practice was the representation of student outcomes as grades only with cumulative outcomes aggregated for the updating of Grade Point Average (GPA) scores.

For some the mapping of inputs was a proxy for mapping cohort progress throughout a program on the assumption that if graduate learning outcomes were covered across individual units of study and if assessment was aligned with these learning outcomes, then students who passed all assessment tasks could be considered to have achieved all the graduate learning outcomes. Some participants were uncertain of the strength of this line of logic, including those in institutions with standardised grade cut-offs such as 50% “Pass” grades, a ‘standard’ they felt had little credibility as a guarantee of achievement of learning objectives or GLOs.

Few units of study were classified as ‘capstones’ but most (37) assessments discussed were from compulsory units of study. Being able to specify compulsory units of study facilitated institutional capacity to track student progress in the development and demonstration of GLOs. This was an advantage for professional programs with few electives. By contrast, generalist programs with multiple pathways posed additional challenges to tracking student progress. This was particularly evident where few units of study were compulsory program requirements, where prerequisites had been removed and where units of study were labelled only as ‘Introductory’ or ‘Advanced’ resulting in cohorts consisting of students at various stages of program progression. This was a significant problem for coordinators in planning systematic development and assessment of GLOs across the years of a program.

It was evident that many institutions had or were considering the importance of some form of progressive documentation of student development or achievement of program graduate learning outcomes and more than a quarter of interviewees reported that their institution was currently investigating possibilities. Some take-up of ePortfolios was reported but there were also reservations about the value of this form of documentation of progress unless consistent practice could be guaranteed throughout the years of a program. A common theme was the desire for technology that could be used in support of mapping processes and monitoring student development and achievement of graduate learning outcomes across the years of a program.

4.2.6 Factors influencing quality

Enhancement factors
Interviewees were generally able to describe an example of an assessment quality enhancement from their own experience and were unanimous in attributing this enhancement to opportunities to exchange ideas and practice in discussion with peers. Examples included participation in external peer review processes and also institutional support.

*Participation in “Achievement Matters” [OLT project] has led to improved task design and development of criteria rubrics.*

*Assessment practice is supported by access to academic development. In this case it is available within the faculty so assistance is immediate.*

Consequential and rigorous approval processes and inclusive, evidence-based evaluation practices were also cited as contributing to effective assessment planning. There were additional examples of year-level approaches, generally where a first-year coordinator was successful in working with coordinators to provide for GLO development through systematic curriculum and assessment planning across all concurrent units of study and to encourage all teaching staff to raise student awareness of the relationships among their units.

**Limiting factors**

Though interviewees sometimes needed prompting to recall an experience of assessment enhancement, they were more forthcoming when asked what limited or restricted their practice. Concerns with the nature and clarity of graduate learning outcomes were evident in several responses as this had implications for the ease with which they could be assessed and assured.

*Knowledge is easiest to assess and there are many resources for communication but reflection and metacognition are harder to assess as students feel we are assessing opinion rather than learning.*

*Some of the TLOs are challenging to assess, particularly lifelong learning.*

Learning outcomes that were so generic as to be meaningless in the context of specific disciplines or those that implied low standards were perceived as problematic.

*The GAs are very generic there is not a tight mapping of them to course activity so very little attention is paid to them.*

*The TLOs are too basic, probably because they have to apply across a sector that includes both TAFE and research intensive universities. They are there for minimum standards that a research intensive university would want to be exceeding.*

Program-wide consistency and institutional structures and practices could also be limitations.

*Also the extent to which they are useable depends largely on institutional organisation. How closely institutional organisation provides program that closely match the TLOs.*

Increasingly heavy workloads came out as a major issue, particularly for those experiencing growth in class sizes. In addition, research productivity was more likely to attract professional recognition and reward which made assessment a less attractive alternative when allocating scarce resources of time and effort. Participants in several of the national dissemination fora also pointed out that activities associated with assessment enhancement such as committee membership, meetings, project initiation and implementation were considered ‘extra curricula’ and therefore rarely included in formal workload allocations.
Personal qualities such as staff openness to and capacity for change were perceived as restrictive to assessment enthusiasts or leaders when attempting to drive change with their colleagues.

A limitation to assessment practice is people’s imagination. A person likes doing what is familiar but this doesn’t always achieve the desired LOs

For change to happen there needs to be cultural change. There is broad acceptance with some academics engaged more than others. Some still like their own patch.

Resourcing was also high on the list of limiting factors. The assessment of graduate learning outcomes is often reliant on authentic performance tasks which can be resource-intensive. However resourcing levels resulting in unfavourable staff-student ratios had made the provision of these opportunities much more challenging. Limited resourcing could also affect the amount of consideration sessional staff such as tutors were able to give to student work and to the provision of feedback. Resourcing levels had also restricted the inclusion of sessional or multi-campus staff in the task development, calibration and moderation activities required to assure learning outcomes at appropriate standards.

We fail to grapple with the reality of the casualisation of workforce and the variable standards for teaching and learning.

Policies, in particular those related to assessment and quality assurance, could also have a restrictive impact on practice though this was obviously not the intention of those responsible for policy development.

Requiring student grades to follow a bell curve distribution is a limitation as students are required to complete an application process and this course is always oversubscribed the students tend to be above average.

Policy issues are addressed in more detail in Summary paper 6.

AAGLO Summary 6: Approaches to the assurance of assessment quality

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 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.
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 and assessment design 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>Supplemeting 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>Issue</td>
<td>Background</td>
<td>Why this is an issue</td>
<td>How the issue can be addressed</td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>----------------------</td>
<td>--------------------------------</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>
Chapter 5 Dissemination

Dissemination was a priority from the establishment of the project and activities undertaken were designed to incorporate dissemination purposes wherever possible. In addition, the team was responsive to any serendipitous opportunities that arose.

The establishment of TEQSA has intensified interest in this area in Australia. This has meant that the project has been conducted in a rapidly changing environment of heightened interest around standards both in Australia and internationally. This environment along with the number of parallel projects funded by the ALTC/OLT has created additional opportunities for collaboration and dissemination. Maintaining ongoing communication with the leaders of these projects and LTAS Discipline Scholars has therefore been a dissemination priority.

5.1 AAGLO website

A project website was established at <http://www.itl.usyd.edu.au/projects/aaglo/> and served as a dissemination point for various project outputs including the AAGLO Summary Papers, the conference presentations and the capital city fora outputs.

Early dissemination activities were undertaken to identify the activities of relevant groups such as other OLT project teams in order to communicate the project plans and to establish productive relationships. Consultations with the international reference group created wider awareness of the AAGLO project, as did the series of round table discussions conducted at international conferences. Reference group members provided useful insights as well as commentary for a number of the AAGLO Summary papers.

The conduct of an ongoing literature review and a scholarly approach to the development of the AAGLO Summary Papers resulted in the construction of an extensive Endnote library. This contained 266 relevant references indexed to the major areas of assessing GLOs, as well as quality assurance and standards for student outcomes.

5.2 AAGLO summary papers

A series of AAGLO summary papers was developed on topics that emerged as significant issues in the sector around assessment and assurance of GLOs. These papers (listed in Table 1) were designed to be succinct and to raise awareness of the national and international perspectives on the particular issues including related international trends, external review, standardised testing, and the characteristics of effective tasks and assurance processes.

5.3 Capital city fora

The AAGLO project held a series of five fora in May and July 2012 to disseminate project findings and provide an opportunity for academics to engage with colleagues in discussing the assurance of learning standards in Australia’s emerging quality framework.

Professor Trudy W. Banta, Professor in Higher Education and Senior Advisor to the Chancellor for Academic Planning and Evaluation at Indiana University-Purdue University Indiana was the keynote speaker for the fora in Brisbane, Melbourne and Sydney. Professor Royce Sadler, Senior Assessment Scholar at The University of Queensland was the keynote speaker in Perth and Adelaide. Having two keynote speakers expanded the range of perspectives able to be presented and allowed for alternative views on some issues such as portfolio assessment and the use of rubrics. Both keynote presentations were recorded and are available along with handouts on the project website <http://www.itl.usyd.edu.au/projects/aaglo/2012_fora.htm>.

Assessing and assuring Australian graduate learning outcomes 57
Professor Banta provided a review of discipline-based assessment tasks and their ability to provide convincing evidence of the attainment of graduate learning outcomes. She explored the issues around academics collaborating on setting their assessment tasks in order to facilitate informal benchmarking and to reduce workload and share good practice. Professor Banta described the use of standardized testing in the USA and some of the inherent limitations in this approach to assuring the quality of learning outcomes across a program or unit of study. A particular interesting point highlighted during the presentation was that the most significant correlation for student performance on standardised tests was their performance on previous standardised tests and not their performance associated with current learning. The importance of intrinsic motivation for students to learn and perform well was highlighted as was the necessity for teaching staff to spend time developing meaningful assessment tasks that engage students in a variety of task types.

Professor Sadler reminded the audience that standards can mean different things in different disciplines. Standards when applied to a pair of safety glasses can be defined by very specific criteria, but learning and teaching standards are rarely absolute and rarely shared across disciplines. There is not a common vocabulary yet to define assessment standards.

The capital city fora provided an opportunity to gather further insights into the use of different assessment task types to evidence the attainment of graduate learning outcomes. Participants expressed concern that there were many concurrent activities taking place in the sector around learning and teaching standards and quality assurance and that it was difficult for the discipline academic to know what was expected of them in terms of evidencing assessment outcomes to those outside of their discipline. Academics frequently know how to evidence learning outcomes to colleagues within their discipline, but it is sometimes difficult to find a common language across disciplines.

5.4 Publications and presentations

The project team also used opportunities afforded by national and international conferences in mid-2011 to obtain peer review on draft documentation (Assessment in Higher Education Conference, Carlisle, 2011 and the Higher Education Research and Development Society of Australasia Conference, Gold Coast, 2011). A later opportunity (International Society for the Scholarship of Teaching and Learning, Milwaukee, 2011) enabled the project team to explore the implications of institutional adoption of discipline-specific learning outcomes for the purposes of assurance and comparison of standards and invite the sharing of participant experience of the international standards agenda in the UK, areas of Europe and America. Appendix A lists the abstracts for major, peer-reviewed conference presentations.


Hughes, C., Barrie, S., Crisp, G. and Bennison, A. (2011, October). Assessing and assuring discipline-specific graduate learning outcomes. Roundtable conducted at the conference of the International Society for the Scholarship of Teaching and Learning (ISSOTL), Milwaukee, WI, USA.


5.5 Impact

AAGLO project activities have led to a number of invitations to provide information about the project and its findings to particular organisational groups within the higher education community. Presentations have been made to organisations including the Council of Australian Directors of Academic Development (CADAD), the Australian Business Dean’s Council Teaching and Learning Network, and the Learning and Teaching Academic Standards Discipline Scholars. Members of the project team have also been invited to present papers at the University of the Sunshine Coast Teaching and Learning Conference, the Graduate Attribute Showcase (Faculty of Nursing and Midwifery, University of Technology Sydney, September 2012), the 3rd Annual Informa Learning and Teaching Forum (Brisbane September 2012), and the Assessing and Reporting Teaching and Learning Outcomes Conference (Melbourne, March 2013).

The AAGLO Project has established links with international scholars and organisations that have both informed the project and provided a means for disseminating project outcomes. Evidence of international interest in the project resulted in an account of the AAGLO project being published in the newsletter of the Quality Assurance Agency, Scotland and links to the AAGLO project website being provided on The University of Winchester website. Discussions with representatives from Alverno College are underway in relation to a proposal for two collaborative two-day workshops (Sydney and Brisbane) for May 2013.

Several potential collaborations with international organisations have also been identified including one with the Knowledge Institute of Thailand.
Chapter 6 Evaluation

Evaluation activities were integral to project design and wherever appropriate, evaluation activities were integrated with development and/or dissemination activities and incorporated both formal and informal approaches.

The AAGLO project team employed a range of evaluation strategies to enable ongoing review and refinement of the project. In addition to the formal processes described below and consultation with the reference group, regular face-to-face and telephone meetings enabled the project team to monitor project progress in relation to milestones and processes.

6.1 Engagement with key stakeholders

Senior leaders in Australian tertiary institutions were involved in the AAGLO project through representation on the project reference group, participation in the data collection phase of the project and participation in the AAGLO capital city fora. Deputy Vice-Chancellors (Academic) at all Australian universities were invited to attend the AAGLO fora, with subsidies being provided for those outside the metropolitan areas of Brisbane, Melbourne, Sydney, Perth and Adelaide (the cities in which the fora were held). The ways in which stakeholders have engaged with the project have provided many opportunities for both formal and informal evaluation and also, to some extent, constitutes an evaluation of the worth of the project itself.

Disciplinary leaders and academics were involved in the project in a number of ways. Two leaders of other projects in this area and two of the LTAS Discipline Scholars were members of the project reference group. All the LTAS Discipline Scholars collaborated with the project team in selecting example disciplines and the academics from their discipline to be invited to interview in the data collection phase of the project. Forty-eight academics, representing all university types (Go8, ATN, IRUA, RUN), sizes, and locations, contributed data to the project.

Recognition of the AAGLO project relevance has been demonstrated through invitations to make presentations to the Council of Australian Directors of Academic Development (CADAD), the LTAS Discipline Scholars and the Australian Business Deans’ Teaching and Learning Network.

Project events including conference presentations and capital city fora have been well attended. Presentations at national (HERDSA 2012, 3rd Annual Informa Teaching and Learning Forum 2012) and international conferences (STLHE 2012, ICED 2012, EARLI-SIG Assessment and Evaluation Conference 2012) have all attracted in excess of forty participants. Combined attendance at the capital city fora was over one hundred and fifty, with participants representing twenty-seven universities and five other organisations and ranging from lecturers to Deputy Vice-Chancellors (Academic).

Students (through the current President of the National Union of Students) and employer groups (though Professions Australia) were represented on the reference group.

Five of the seven disciplines that were the focus of the data collection phase of the project were subject to accreditation. Accreditation processes were discussed throughout the interview process from the perspective of academics. From the data collected it was apparent that some accreditation processes such as those associated with Engineering are increasingly concerned with assessment practices and learning standards in universities but others do not, as yet, focus on the types of tasks that are used to provide evidence of student achievement of graduate learning outcomes or with the quality assurance of assessment tasks and judgements.
6.2 Project website

The project website was used to document project progress and to progressively disseminate outcomes such as resources and summary papers. Use of materials disseminated via the project website (hosted by The University of Sydney) was monitored between 28 August 2011 and 28 August 2012. Data collected indicates that the website was a useful means of disseminating information about the project.

The AAGLO home page was viewed more than 700 times, with the Project Outline viewed more than 500 times and the registration page for the AAGLO fora viewed almost 700 times.

By the completion of the last capital city forum, the AAGLO Publications page was viewed more than 400 times with individual AAGLO Summary Papers each viewed more than 60 times.

6.3 Capital city fora

A number of strategies were used by the project team to enable both formative and summative evaluation of the effectiveness of the capital city fora. Formative evaluation was achieved through observation and recording of participant responses to interactive activities and a brief survey completed by participants at the conclusion of each forum. The data obtained in this manner was used to inform planning for subsequent fora and enabled continual refinements to be made.

General comments about the fora included the following:

I found the discussions at the T & L perspective (rather than a discipline level) quite interesting. It was good to engage with a range of academics.

Sydney Forum

As a person from the private sector, it was invaluable to connect with the AAGLO project. By such means the universities serve the wider (educational) community.

Sydney Forum

Each aspect of the workshop was extremely useful and engaging and really helped to clarify my thinking and practice. Thank you.

Perth Forum

Discussion around how to address ensuring that we have provided GLOs (around transcript expansion, standardised testing, explicit GLOs and moderation). Pros and cons to be thought of was useful as well as information on assessment tasks (although more time on this would have been useful).

Adelaide Forum

Overall participant feedback was positive and indicated that participants found that the most valuable aspects of the fora were:

- the presentation of findings of the AAGLO project
  Prelim findings of research. Needed to spend more time on this, e.g. involve participants in discussion.
  Brisbane Forum

- The findings of the project were a useful reminder of the key issues.
  Melbourne Forum

- Hearing about the research – lots of ideas and avenues to pursue, helpful to have a bigger picture.
Assessing and assuring Australian graduate learning outcomes

Seeing/discussing the results of the project and the divergent assessment practices and their relationship to course LO and GLO.

- the international perspective presented by Professor Trudy Banta (Brisbane, Melbourne and Sydney)

Trudy’s keynote was a fantastically well-informed “historical” view which can help us engage in the political discourse surrounding standards.

Hearing Trudy’s take on standardised testing. Hearing Trudy’s ideas about ‘1st look’ and ‘2nd look assessments’. Seeing the potential for improvement of assessment as part of the assurance agenda.

- the keynote presentation by Professor Royce Sadler (Perth and Adelaide)

Loved Royce’s thought provoking talk.

Royce provided a great overview of where we are at and the challenges we face.

The demonstration (by Royce Sadler) of the futility of standardised tests to identify whether Graduate Outcomes have been achieved, and the promise in the explicit discipline based inclusion of LOs in criteria for teaching and assessment.

- opportunities to discuss the issues around assessment and assurance of graduate and to hear what is happening in other discipline and institutions

Interaction with colleagues and context-setting of assessment and graduate skills/attributes.

Discussing with colleagues around issues and strategies – wide-ranging and in-depth. The workshop format worked to facilitate this very effectively.

Enhanced awareness of what is happening in other institutions. More aware of the challenges the sector is facing in terms of good assessment practice.

Forum to hear viewpoints from other disciplines and to realise we all share the same dilemmas.

The wonderful opportunities to discuss, unravel, argue, comment upon issues of a assessment that affect the learning of students (and the teaching processes aligned with learning). Excellent professional development. Thank you.

Hearing the results of the data collection on assessment tasks. Finding out what is going on across disciplines will be useful when talking about quality assurance within disciplines.
Ability to share practice with others (and discover similar needs/difficulties).

- opportunities to network with colleagues from other institutions

Good networking opportunity although it would have been good to have deliberate mixing between institutions.

Conversations with colleagues whom I rarely get to speak to – people of different backgrounds and roles.

Some suggestions for improvement raised by participants that that project team endeavoured to address in subsequent fora were centred around providing:

- more opportunities for interaction between participants

More time for discussion of the issues raised, especially for consideration of the external policy context and how AAGLO and similar projects’ findings might fit. A whole day would have been good and I would very much like to attend such an event.

- more structured workshop activities

Scenario discussion could have been more structured and longer.

- more detailed examples of task types that can be used to generate evidence of achievement of graduate learning outcomes

- more discussion on assurance of graduate learning outcomes across the years of a program.

Participants also made suggestions for future work in this area.

Where to next: How to engage the academic community (e.g. Alverno) into the rich conversations ... as a positive, not a compliance exercise. Discipline Scholars are a useful model.

In future I’d like to know more about how students’ development of GLOs is supported and recorded across the YEARS of a program. (emphasis in original)

Forum has triggered more ideas that need discussion.

I am interested in finding out what students have learnt/developed by asking students. Perhaps this was outside the scope. It is assurance rather than assessment.

‘Whole-of-program’ approaches in a generalist degree – are there approaches out there that work?
6.4 External Evaluation

Margaret Buckridge (Griffith Institute for Higher Education, Griffith University) was engaged to undertake formative and summative evaluation of the AAGLO project. This evaluation was undertaken through regular contact with the project team, attendance at an early planning meeting, provision of feedback on project processes and outputs, and attendance at three of the capital city fora. Her Evaluation Report is presented in full below.

**Evaluation Report, August 2012**

**Introduction**

The following report presents an independent evaluation of the ALTC-funded project *Assessing and Assuring Graduate Learning Outcomes*.

The project was located within the current debates in the higher education sector about how to assure the quality of student learning within the sector. Specifically, there were three goals:

1. to identify effective strategies for assessing graduate learning outcomes in a way that genuinely tests the integration of discipline knowledge with more generic capabilities,
2. to identify the continuing processes which provide assurance that the assessment tasks are valid for the purpose, and that student achievement on those tasks is judged by appropriate standards, and
3. to raise awareness and to organise active participation in debates concerning the standards agenda by establishing productive dialogue among key stakeholder groups.

The project team has sought to implement these goals via various activities which run the gamut from a quite micro focus (particular assessment tasks) to a much more macro focus (analysing the options at sector level for assuring the quality, or standard, of graduates’ learning).

- It carried out an interview study in seven discipline areas, approaching key practitioners in those areas to talk about assessment tasks and assessment regimes
- It developed a series of summary papers in relation to the key issues (readily available on the project website)
- It developed an Endnote Library of the relevant scholarly literature
- It held forums in five capital cities to report, progressively, on the findings emerging from the interviews and also to enable discussion of the various strategies for assuring quality that were currently in play
- In conjunction with these major activities, the team also engaged in conference presentations, international interactions, submissions to government, and collaboration on a major Issues and Options paper.

My role as evaluator has been both formative and summative. I have had regular contact with the team, particularly with its Queensland members (Clair Hughes and Anne Bennison). I attended an early planning meeting of the full team, saw drafts of the Summary Papers, was present for three of the five capital city forums and was part of the subsequent de-briefing of these events.

The sources of evaluative information used in this report include: my own observations, comments from some members of the project’s reference group, a consideration of the feedback sheets collected from attendees at each of the five forums, and notes taken by me as I circulated at the forums which I attended.

In making the following evaluative comments, I shall look both at the conduct and at the achievement of the project.
The Conduct of the Project
The project was well-organised and well-managed. Although each of the three academic members of the project team faced new or additional employment demands in the course of the project, focus and timelines were well-maintained via a relationship with the project manager. The project activities were largely consistent with what was initially proposed by the team.

- There was a considerable amount of scholarly review and national and international liaison in order to clarify the ground of the project. This was particularly important given the number of projects, both concurrent and recent, which had the potential to overlap with this work. By virtue of this work, the project positioned itself in a timely and strategic way at the centre of the standards debate. This resulted, as planned, in a series of summary papers which picked up the key issues in relation to the standards agenda, making the debates accessible to an audience who may not have either policy or assessment expertise.

- The discipline focus of the interview study was seen as crucial given the emphasis in recent scholarship on discipline-specific manifestations of broader graduate capabilities. Even so, there was no intention to replicate the work of the LTAS projects, but rather to gather accounts of practice and process in order to build guidelines and to illustrate them with examples. In the event, for various logistical reasons, there were slightly fewer disciplines and slightly fewer interviewees than originally planned. There is no suggestion, however, that the interview data were compromised or limited by this – the data were copious and rich, albeit with perhaps less innovation and variety than anticipated, and the team used NVivo to analyse and clarify what they had gathered.

- The capital city forums were designed both to introduce the topic of the standards agenda and to reveal the findings from the interviews, but also to give participants a chance to ask questions and share views. These were well designed for the purpose. Each forum featured a provocative guest speaker, and project team members used their networks to ensure that a range of institutions and of institutional roles was represented in the audience.

- The reference group for the project was chosen, sensibly, in such a way as to have representation from all relevant stake-holder groups. Although there was input from the reference group early in the project, there was considerably less as the project moved into its implementation stages. One member, who thought the project very valuable, felt that the reference group might have been able to add clout to its dissemination process. It is clear, however, that one of the risks incurred when people are appointed in a representative way is that the actual people change – at least five of the representative members, in particular, had changed status and employment by the time I attempted to consult them as part of the evaluation process.

As already indicated, the scope of this project was ambitious. The team was seeking discipline-specific classroom-level information about effective assessment tasks, and at the same time seeking to use this information to compile guidelines and exemplars which would contribute to a deeper understanding of how best to assure the quality and standards of learning at the sector level. Given this scope, it was important that the activities proposed for the conduct of the project should hold together the focus on practice with the focus on policy. This weaving together was not just a matter of carrying out the planned activities, but also of consciously using them to make the connections between practice and policy clear. Many of the responses to the forums indicate that this was successfully achieved.

The Project’s Achievement
This has been a significant project, partly because of its timeliness. The project started out with three goals, and achievement on each of them has been substantial.
Contribution to sector debate
At the macro level, one of the project’s goals was to raise awareness around the standards debate in higher education. I believe that this has been an area of outstanding achievement for the project. This is a continuing, fraught debate both nationally and internationally, although it is clear that not all discipline-level academics are aware of the terms of this debate. This project has canvassed, via its activities, the various options that are in play for ensuring that Graduate Learning Outcomes meet the standard expected. The forums have seen discussion of the main options: standardized testing, rubrics, portfolios, external examiners, and consensus moderation.

For many of the attendees, Trudy Banta’s presentation, based on her own wealth of experience, provided a first opportunity to understand the issues surrounding standardized testing at a much deeper level; similarly Royce Sadler’s presentation broke open the challenge of how to assess whether GLOs were actually being achieved, demonstrating ‘the futility of standardized tests to identify this’ (forum feedback sheet), and canvassing the potential of more local, discipline-based consensus. Forum attendees were clearly appreciative of the overview provided; many also mentioned the usefulness of ‘touching base’ with colleagues in other institutions in relation to these issues. In later forums, where participants were directed more explicitly to the Summary Papers, feedback on these suggested that they would take their place as valuable resources in local and institutional policy-making. The papers represent, at this point in time, one of the most worthwhile legacies of this project. (Two or three more papers remain to be published on the website at the time of writing.)

Contribution to practice
At the more micro level, the project had a couple of goals (1 & 2 as indicated at the beginning of this report), gathering information about tasks and processes via a discipline-based interview study. Many Australian academics are still new to the idea of genuinely integrating discipline knowledge-and skills outcomes with more generic graduate learning outcomes. The purpose of this data, therefore, was to use it to shed light on the kinds of assessment tasks which can assess these outcomes and to derive guidelines and examples to assist thinking and practice in this area. It was also to identify processes, both pre and post assessment, which would ensure that standards were monitored and maintained. The project team was conscious that a number of disciplines had developed outcome statements and resources via the LTAS project. The idea was not to replicate this, but rather to explore tasks and strategies as they were actually being implemented.

The findings from this exploration were progressively presented at the forums and are presented at greater length in Summary Paper 8 (still in draft form). Forum attendees were avid for these findings. There was great interest in the kinds of tasks which, in the judgment of discipline academics, were capable of carrying the weight of integrating substantive knowledge and generic outcomes. A number of feedback sheets suggested that attendees would like to have seen even more examples. (In fact, they were made aware at the forums that examples of both tasks and processes were, or would be, available in the Summary Papers.) It is clear that academics are keen to ensure that the ultimate Standards solution is discipline-based, and that the policies governing the monitoring and maintenance of those standards should be local to the discipline, albeit transparent and accountable. The forums did excellent work in helping to clarify and coalesce some of this thinking.

The project’s products
I have already indicated above that the Summary Papers make a strong and timely contribution to sector thinking in this area, particularly the first four papers (on the standards agenda, external review, work-based learning standardized testing). It should be noted that the subsequent papers, which focus more on the findings of
the discipline study, help to flesh out a positive approach to implementation of the standards agenda. Summary Paper No.6 is especially noteworthy for the clarity of its analysis of policy issues and policy solutions. Papers that are still in preparation as the project moves towards its conclusion will address e-assessment, student perspectives and whole-of-program strategies.

The other product that should be mentioned is the Endnote Library, a valuable and motivating resource for discipline academics who need some scholarly support to get into the Scholarship of Teaching in the area of assessment. One of the reference group members praised this particularly, suggesting that many academics were operating without a good conceptual grasp of assessment.

Conclusion
This has been a project where both conduct and achievement have been strong. While it is not clear nationally what formal mechanisms for ensuring standards will ultimately be adopted, the project has provided a scholarly overview of the options and has explored the meaning of assessing and assuring GLOs (Graduate Learning Outcomes) at discipline level. It has produced resources both for discussion and for the improvement of practice. Some of the responses to the project have made clear that there is also necessary further work to be done. There are broadly-based concerns that best practice is incompatible with already over-stretched workloads and with an increasing use of casual staff, as well as concerns that academic staff do not ‘own’ this agenda, but rather see it as part of a growing managerialism. Clearly, these are big challenges. Although the resources developed by the project go some way towards addressing these problems, it probably remains the case that further dissemination and development are still required. It is for this reason that it is very pleasing to note that the project has obtained permission to spend unused budget (itself a testament to the efficiency of the project’s management) to fund a couple of workshops in conjunction with an Alverno College team. Alverno has long been acknowledged as modeling excellent practice in the area of integrated curriculum and assessment – these workshops will be a welcome addition to the further dissemination of the project’s work.
Chapter 7 Conclusions and recommendations

The AAGLO project investigation has been broad in scope in seeking to explore the landscape rather than the fine detail of practice in issues related to the assessment and assurance of graduate learning outcomes. Project findings, many of which have been progressively documented in a series of ten AAGLO Summary papers, have been informed by consulting a range of national and international sources. This has enabled the identification of emerging issues and highlighted recommendations for future priorities.

7.1 Assessment task design

One of the key questions posed by AAGLO was:

What types of assessment tasks are most likely to provide convincing evidence of student achievement of or progress towards graduate learning outcomes?

This investigation has provided discipline practitioners perspectives on the types of tasks considered effective in the assessment of GLOs, and in general these were consistent with disciplinary and professional traditions. While these tasks varied between disciplines, there were more abstract features or characteristics of effective assessment of GLOs that were common across disciplines. It was also apparent that many commonly used assessment tasks were notably absent from the range of assessment tasks reported as effectively evidencing achievement of GLOs, for example MCQ and written exams. Interview respondents cited resource constraints as one of many factors driving the selection of assessment tasks other than their ability to demonstrate achievement of GLOs. However the emergence of additional assurance of learning requirements as a focus of regulatory requirements means that their capacity to do so should be paramount in underpinning the choice of tasks. It has also been evident that discipline assessment traditions are often most effective in providing evidence for only some domains of GLOs – content, cognitive and communication capabilities - but are less effective in the assessment of outcomes related to attitude, values and metacognitive abilities which feature largely in statements of graduate attributes and recent articulations of disciplinary TLOs, for example those produced through the LTAS project.

Examples from our interviews and discussions with other project leaders demonstrate that traditional tasks handled in traditional ways may limit their capacity to assure GLOs and that careful attention to task design features or characteristics is equally if not more important than the selection of task type in assessment planning. Two superficially similar tasks can differ significantly through different approaches to the characteristics outlined in AAGLO Summary 7. For example, a standard group work task to undertake a piece of research and communicate it in a written report may generate evidence of only research and written communication abilities but fail to generate evidence of individual self-regulation, autonomy or team work skills. However, setting students a realistic problem to be solved, assigning a life-like professional role such as consultant, suggesting a client as audience of the report, requiring written commentary on group processes, an oral presentation and individual reflections on learning, and the moderation of product marks for individuals through application of an anonymous peer review process, can greatly expand the range of evidence generated without changing the type of task itself. If undertaken in an Accounting program, for example, an enhanced group task along the lines of this example would provide opportunities for students to generate a range of evidence specific to LTAS TLOs:

- **Knowledge**: Integrate theoretical and technical accounting knowledge which includes a selection of auditing and assurance, finance, economics, quantitative methods, information systems, commercial law, corporation, law and taxation law
- **Communication and Teamwork**: Justify and communicate accounting advice and ideas in straightforward collaborative contexts, involving both accountants and non-accountants
• *Self-management*: Reflect on performance feedback to identify and action learning opportunities and self-improvements.

*Threshold Learning Outcomes for Accounting, ALTC 2010, 10*

Assessment design is a skilful undertaking involving not only the selection and incorporation of relevant features in the design of individual tasks but also creating the relationships among tasks that contribute to unit coherence and that facilitate the provision and application of feedback on performance. Unit and program coordinators need the support of appropriate professional learning opportunities to develop appropriate skills.

Assessment design is a priority of a number of past and current OLT projects and the previous ALTC also coordinated events linked to national assessment conferences and initiated the production of summaries of national activity and good practice guides. Despite this, assessment practice in many areas is still in need of attention.

**Recommendation 1:**

That universities review their choice of assessment tasks in programs and units of study to ensure these include design features that provide convincing evidence of achievement of graduate learning outcomes.

**Recommendation 2:**

That universities provide professional learning opportunities and support for academics to develop effective discipline based assessment for the assurance of GLOs.

### 7.2 Whole-of-program approaches

Regardless of the quality of individual tasks and unit assessment plans, systematic development and demonstration of GLOs requires a whole-of-program approach to assessment planning. AAGLO Summary 9 outlines the rationale for whole-or-program approaches, the forms it can take, implementation steps and associated challenges.

**AAGLO Summary 9: Approaches to the assurance of assessment quality**

**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*\(^4\)* 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.

\(^4\)The term ‘assessment’ is used throughout this project to refer to the making of judgements about student achievement.
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 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 (Bath et al 2004) based on collegiality and evidence-based discussion (Brennan 2012)
   - ensuring that retrofitting outcomes to an existing curriculum is a driver of curriculum renewal rather than a perfunctory administrative procedure (Bloxham 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.

3. 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)

4. 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 (Bloxham 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).

- becoming familiar with institutional regulations regarding workload, hurdle or Pass/Fail assessment tasks, the maintenance of academic integrity, self and peer assessment (Bloxham 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 (Bloxham 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 (Bloxham 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.
The coordination of a whole-of program approach to assessment planning is a complex undertaking with implications for resourcing, infrastructure and quality assurance practices: this explains why there are few examples outside accredited programs where accreditation drives the process, and why even within accredited programs the emphasis can be on coordination of subject matter more than the development of integrated and complex learning outcomes.

A whole-of-program approach does not necessarily equate to a single unit-level assessment at the end of a program of study – for example a capstone assessment, although this is a common element of a whole-of-program approach.

There is no single approach to ‘Whole-of-program’ assessment. The approach will vary significantly between the more consistent pathways in focussed professional degrees and the more varied pathways inherent in generalist degrees. The nature of whole-of-program approaches to assessment will be particularly challenging in the latter. Mapping of curriculum initiatives and systems has the potential to be extended to support whole-of-program approaches to assessment however these are typically perceived as bureaucratic and rarely experienced by academics as supporting assessment practice at present.

**Recommendation 3:**

That universities promote whole-of-program approaches to assessment planning through the development of guidelines and supporting systems (including technology).

### 7.3 Assessment policy

Curriculum and assessment renewal may at times need the removal of restrictions as much as the addition of support. Academics often attributed an inability to incorporate specific assessment tasks or approaches into their practice to the restrictions of institutional assessment policy. The understandable concerns of many policy developers or working parties in attempting to protect students from unsound assessment practices can often have the unintended effect of also ‘protecting’ them from enhanced practice. The unintentional consequences of well-meaning assessment policy were discussed in Summary 6.

Even policy free of restrictive components does not guarantee effective practice as policy ‘drift’ or non-compliance had been observed among participants. Good assessment policy needs the support of meaningful quality assurance policies and procedures for faithful and sustained implementation. The OLT has in the past supported investigations into assessment policy (Duck et al 2011) but there is room for further investigation of the reciprocal implications for assessment and other related policies and practices such as quality assurance, performance review and institutional data collection.

**Recommendation 4:**

That institutions examine the implementation of Assessment policy with a view to identifying any unintended effects of policy on assessment practice and the extent to which other related policies limit or facilitate appropriate assessment and assurance of GLO.
7.4 The importance of students as partners in assessment

Students’ first-hand experience of assessment, both good and bad, is often conveyed quite strongly in their responses to institutional surveys. Unfortunately this is often the only form of engagement lecturers have with students in an area of prime importance to both. Summary 7 included examples of more active roles for students in assessment and also referred to the advantages of closer partnerships in developing students’ assessment literacy as a means of encouraging their formation as lifelong learners. Summary 10 has been produced in partnership with a member of the project reference group, the president of the National Student Union, to highlight areas of importance to students.

September 2012

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, Rust, Donovan and Handly, 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.

The student voice can have a powerful impact on academic practice and access to the student perspective on the assessment of GLOs could be beneficial in stimulating change. Students in many Scottish universities for example are invited as equal members of high level committees such as school and program reviews panels, have a leadership role in the management of teaching awards and receive training in requisite skills such as communication and negotiation which can contribute co-curricula evidence of GLO achievement.

Recommendation 5:
That university communities identify and implement effective strategies for appropriate and meaningful student participation in assessment design, implementation and assurance.
7.5 Assurance of Learning standards

Assurance of assessment tasks and of the judgements made using those tasks contribute to the assurance of learning standards. There are many activities involved in such assurance for example: through unit of study level induction and moderation activities, program-wide comparisons through a variety of mechanisms sometimes criterion based but largely through the comparison of grade distributions; and among institutions and countries through external review and benchmarking. A shift of emphasis from input standards to academic process and outcome standards as the focus of recent regulatory arrangements has provided the impetus for investigation of approaches as diverse as standardised discipline based tests and peer review incorporating extended discussion of student work samples in relation to specific criteria and standards. A viable approach will require many elements however this project has underlined the central role of academic-led, discipline based assessment in the assurance of graduate learning outcomes.

A key feature underpinning the credibility of discipline based assessment is the assurance of the quality of the assessment task and the assurance of the quality of the judgements made on the basis of students’ performance of that task. This assurance of task design and judgement was required for both ‘whole-of-program’ or unit of study assessments and unit of study assessment during a unit of study, if these are to contribute (as they can) to the development and assurance of GLOs.

Recommendation 6:
That universities review and document the processes they have in place to assure the quality of tasks and judgements in unit of study and program assessments.

The purpose of higher education and the meaning of learning are a precursor to assuring a ‘standard’. A consistent theme in both the research on graduate learning outcomes, and the research on assessment criteria over the past decade relates to the difficulty in unambiguously defining learning ‘outcomes’ for the purposes of assessment. Consistent with the previous ALTC funded study on graduate attributes, the National GAP project (Barrie, Hughes and Smith 2009), the AAGLO project found strong evidence for the need for attention to the conceptualisations of outcomes. It also found strong evidence for the need for better support for curriculum and assessment development. Both projects also concluded that attention to these factors alone is not enough to ensure either effective implementation of graduate learning outcome or the generation of credible evidence of their achievement.

Many of the factors reported as influencing the assessment and the assurance of graduate learning outcomes - the values implicit in quality assurance and promotion processes, the traditions of the disciplines, the policy environment, the attitudes of staff and students, the competing rewards of research - are cultural factors (Umbach 2007). The extent to which such factors are consistent with the desired GLO assessment and assurance practices is a strong determinant of where academics see it as logical to invest their time and other resources.

Above all, it is clear that until academic communities (staff, students, broader society) have a more nuanced and shared understanding of what it is they are assuring as graduate learning outcomes from a university education, it is unlikely that assessment of those outcomes will be credible.

Recommendation 7:
That the OLT continue to support initiatives that engage academics in productive, evidence-based dialogue around the issue of standards among discipline communities on an institutional or national level.
References


AHELO Project website [www.oecd.org/edu/ahelo](http://www.oecd.org/edu/ahelo)


British Standards Institution (BSI) website http://www.bsigroup.com

Collegiate Learning Assessment (CLA) website http://www.collegiatelearningassessment.org/


Framework Reference Model for Assessment (FREMA) website http://www.frema.ecs.soton.ac.uk/overview


Hughes, C. (2012). A case study of assessment of graduate learning outcomes at the programme, course and task level. *Assessment and evaluation in Higher Education* (iFirst).


IMS Global Learning Consortium website [http://www.imsglobal.org/question/qtiv1p2/imsqti_overviewv1p2.html](http://www.imsglobal.org/question/qtiv1p2/imsqti_overviewv1p2.html)


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


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


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

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

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


References specific to AAGLO Summary 3


7 Barrie, S., Hughes, C. & Smith, C. (2009).*The National Graduate Attributes Project:*


Appendix A:

Conference abstracts


Symposium Chair: Sue Bloxham

Higher education assessment is a massive, global, edifice involving students, academics and administrators as millions of assignments, exams, performances and other tasks are assessed and graded annually. At the heart of this unstoppable behemoth are institutions’ and subject disciplines’ academic standards. Confidence in these standards has been questioned in recent years in several in several HE systems (Brown, 2010; Coates, 2010, Vanderslice, 2004) prompted by a growing diversity and complexity in higher education (Woolf & Cooper, 1999). There are claims of grade inflation and lack of parity of standards as Universities and academics largely retain the right to set and maintain their own standards. So, is assessment safe in our hands?

This symposium brings together three papers which investigate this question from theoretical, policy and practice perspectives. The first paper examines the failure of existing theory to provide a workable model for standards in use. It explores the lack of alignment between empirical evidence regarding academic judgement processes and the assessment information provided to students. The second paper provides an insight into approaches to policy implementation regarding standards. It explores features of a range of nationally-funded projects designed to enhance the understanding and use of standards in Australia’s evolving regulatory environment. The final paper explores professional judgement at the practice level. It reports on an empirical study regarding how different members of the academic profession come to know academic standards as operationalised in their grading of student work.

The symposium debate will draw on commonalities in the papers regarding ways forward for policy in rebuilding confidence in assessment, for example using communities, dialogue and exemplars. It will provide the opportunity to debate how standards might be more effectively understood and used in the academy at different levels and with different audiences.

Paper 1.- Sue Bloxham

“You can see the quality in front of your eyes”: developing a policy agenda for academic standards in use.

This paper considers academic standards in assessment; an imperative at a time of political and reputational pressure in the standards arena. Assumptions underpinning standards are discussed drawing on contrasting theoretical perspectives; the techno-rational perspective which underpins quality assurance of assessment and the hermeneutic account which emphasises the interpretive nature of academic judgement.

Neither epistemological approach provides a satisfactory basis for grading policy which is illustrated by the contrast between the assumptions underpinning quality assurance of standards (reliable, consistent, transparent) and the everyday practice of grading (shifting, socially constructed and tacitly held standards). Yet millions of assessments are graded unproblematically each year. The paper will consider how these contrasting perspectives are negotiated by lecturers in the act of grading work as a basis for developing a policy agenda for academic standards in use.

It draws on a mixed-method study of faculty from four disciplines and 3 universities. ‘Think aloud’ protocols and interviews were utilised to access both actual and espoused data regarding grading practices. The findings illuminate how lecturers are implicitly combining different epistemologies in negotiating work practices. Whilst recognising that much of what
they do is internalised and tacit, assessors gain a sense of security about their marking judgements from various quality assurance documents and procedures. The act of creating and discussing, for example, assessment criteria provides a clear opportunity for detailed dialogue about standards. There is a strong influence of ‘techno-rational’ artefacts on interpretive practice as many quality assurance processes for assessment become normative.

However, the findings continue to reinforce the local and co-constructed nature of standards albeit within a framework of dialogue about explicit statements of standards. The implications for developing appropriate policy regarding assessment standards, consensus, accountability and information for students will be discussed.

Paper 2.- Clair Hughes

“Current Australian approaches to assessing and assuring academic standards”

An intensification of interest in the quality of teaching and learning in higher education has led to international trends in defining and articulating the standards of academic achievement (Harris 2009). While commonalities can be observed among the ways the standards agenda has been given expression in many parts of the world- for example the development of quality frameworks and the emergence or refocusing of regulatory bodies – distinctly individual national priorities, activities and support structures have also emerged.

An investigation of standards-related projects supported by the (recently abolished) Australian Learning and Teaching Council (ALTC) was undertaken as the situational analysis stage of the Assessing and Assuring Graduate Learning Outcomes (AAGLO) project (Barrie, Hughes and Crisp 2011). Though the issue of ‘standards’ had been a central or peripheral concern of many ALTC projects since its 2004 inception (Oliver 2011), the Learning and Teaching Academic Standards (LTAS) project (ALTC 2011a) was the first ‘standards’ project coordinated on a national scale.

This paper briefly introduces the LTAS project and Australia’s evolving higher education regulatory arrangements (TEQSA 2011) in order to contextualise an analysis of the large and diverse range of activity the ALTC has subsequently encouraged. The analysis of 11 ALTC grant projects and four fellowships that are currently addressing academic standards focuses on factors such as discipline, scope, sponsoring or leading body, the nature of outcomes and approaches to the conduct of the project and to the assurance of standards. The analysis makes sense of the multiple approaches and perspectives being delivered through this range of activity and adds a complementary perspective to the standards discussions in the Australian higher education context.

Paper 3.- Margaret Price

“Learning to mark: exemplars, dialogue and participation”

The problems of shifting from norm-referenced assessment to criterion-referenced assessment have been identified by several scholars in recent years. These important critiques touch on a number of areas, but neglect a key question about how assessors learn to accomplish what Shay calls a ‘socially situated interpretive act’ (2004). Criterion-referenced assessment is premised on clear standards applied consistently by all assessors in a community. However the limitations in making standards explicit and the consequent problems in sharing common understanding of standards means that the criterion-referenced assessment remains ‘faulty’ especially where there are multiple assessors working together. New assessors or incomers to a local community of markers have to quickly try to assimilate the standards in use in order to belong. Research which does exist tends to focus on salaried, full-time academics. This overlooks the heterogeneity of the academic labour force in HE, and the substantial contribution made by contract (hourly-paid) lecturers, particularly in applied disciplines such as Business and Management.
This study explores how nine newcomers to a UK Business School - including salaried and contract staff - attempt to understand local assessment practice. We use a situated learning lens to analyse their diary entries and interviews about their experiences of learning to mark. Drawing newcomers coming to understand local assessment practice. We argue for the fundamental importance of dialogue about exemplars and other aspects of practice – both to develop inter-subject understandings of assessment ‘anchor points’, and also to foster among newcomers (and especially contract lecturers) a great sense of participating in, and being valued by, the local assessment community.


There is considerable activity in higher education systems around the globe, focused on articulating, assessing and assuring the learning outcomes of university graduates. This is predominantly driven by a desire on the part of governments to increasingly regulate, define and assure teaching and learning in universities (Harris 2009). While the proclaimed interest in greater transparency of academic standards is seen by many as being a positive thing, it does involve the setting of new boundaries around university learning and around academic communities, and perhaps the breaking of others.

In Australia the desire to prescribe the boundaries of university learning is perhaps most clearly manifest in the development of new national statements of teaching and learning standards for all university degree classifications as part of the national Australian Qualifications Framework (AQF) and in the proposal by the government’s new university regulatory body Tertiary Education Quality Standards Agency (TEQSA) to monitor universities in terms of the achievement of student learning outcome standards at the graduation. The government’s Australian Learning and Teaching Council (ALTC) have supported ‘boundary setting’ for university learning through a project to articulate ‘threshold learning outcomes’ across a number of discipline areas (ALTC 2011). This, alongside the proliferation of non-university providers in the tertiary education sector, has also prompted a renewed interest by universities in defining and claiming their educational ‘territory’ with employers through statements of graduates’ employable skills (Oliver 2011) and the graduate attributes universities have long claimed to describe the ‘value-add’ of higher education (Barrie 2012). In Australia such efforts have led to widespread, though perhaps somewhat uncritical, acceptance of the need by higher education institutions and discipline communities to define and implement agreed ‘boundary’ statements of university learning outcomes.

What has been conspicuously absent is a critical scholarly consideration of the consequences of such boundary setting. Where debate has occurred, it has been largely focused on how to best measure such graduate learning outcomes without a consideration of the consequences of such actions. This has included a tendency in some quarters to simplify how boundaries are prescribed and to retreat from the articulation of more complex or even open-ended learning outcomes (Barnett 2000) that are more challenging to measure. It has led to a consideration of more standardized curricula and of measuring student learning in higher education using standardized tests and the likely introduction of a national test such as the Collegiate Learning Assessment. It has also however prompted significant interest in improving academics’ assessments of student learning to provide more credible and consistent evidence of graduates’ achievements of graduate learning outcomes.
Drawing on the project team’s review of current national policy in Australia, data from interviews with academic staff and a review of national learning standards projects in Australia, the scope and nature of national efforts to set boundaries for university learning in Australia will be presented. An analysis, completed by the research team and a panel of international experts who form the project steering group, will be presented identifying the consequences of such boundary setting. The research presentation will provide a basis for further discussion with session participants on national and institutional ‘boundary setting’ for university learning in their own countries and universities.


Assessing and Assuring Graduate Learning Outcomes (AAGLO) is an ALTC project that is reviewing current approaches to assessment standards in the light of national moves to assure quality and standards for students’ learning outcomes.

This short paper will first provide an overview of the interconnected DEEWR projects and government initiatives in the area of teaching and learning standards. Against the context of those connections it will present an analysis of key themes emerging from a series of interviews with 47 academics in a range of universities from the disciplines of Business, History, Law, Drama and Performance Studies, Veterinary Science and Chemistry. The interviewees were invited to participate because they had been recently involved in activities that attempted to define standards and what might constitute evidence for appropriate assessment practices for graduate learning outcomes. The interviews sought to identify examples of effective disciplinary and program practice for the assessment and assurance of standards as they apply to graduate learning outcomes. The project team sought clarification of what evidence would academics expect to see in terms of the type of assessment task used to probe the development of skills and capabilities that could be related to graduate learning outcomes?

Connections between assessment practices in different discipline domains were apparent as certain features emerge from the analysis of the interviews; these features can be related to core assessment types and further elaborated in the form of assessment task characteristics that might be related to evidence of standards. These features include the perceived authenticity of major summative assessment tasks, the involvement of students in defining the assessment tasks, effective designs for tasks that are undertaken as a group and a focus on defining tasks that can span multiple units (subjects) rather than continually dividing assessment tasks into smaller and smaller portions.


**Rationale/objectives**

Activity in higher education learning and teaching across the globe indicates widespread interest in greater transparency of academic standards and the promotion of constructive dialogue within and across academic disciplines, institutions and countries (Harris 2009). The ‘standards agenda’ in Australia has been supported by the Australian Learning and Teaching Council (ALTC) through coordination of a project to articulate threshold learning outcomes across a number of discipline areas (ALTC 2011a, 2011b) and funding of disciplinary or generic assessment and standards-related projects (Oliver 2011). Preliminary findings from a 2010 ALTC-funded project - “Assessing and assuring graduate learning outcomes” (AAGLO) - are reported in this presentation by a cross-institutional team with experience in educational development (Barrie et al 2011).
Summary of work
Identification of examples of effective practice in the assessment of graduate learning outcomes was a project priority. This involved telephone interviews with representatives of seven of the disciplines included in the ALTC threshold learning outcomes project and the collection of assessment artefacts.

Outcomes
The interviews yielded data relating to matters such as assessment task descriptions, group task arrangements, the student role, moderation of standards and approaches to quality assurance. Preliminary analysis (incorporating the application of NVivo software to both qualitative and quantitative data) identified a broad range of approaches to assuring the quality of assessment tasks and assessment judgements in Australian universities. While there were distinct disciplinary traditions and practices, there were also many similarities among the different disciplinary approaches which indicated the value of cross-disciplinary sharing of practice. The data also suggested a number of factors associated with effective practice in the assessment and assurance of graduate learning outcomes.

- authentic tasks – relevance to real-life or professional practice
- worthwhile and productive tasks - a blurred distinction between learning and assessment activity
- an active role for students - provision for student engagement ranged from choice of topic or presentation format, to self or peer-assessment and reflection on achievement, and student responsibility for directing, evaluating and documenting their own learning
- a whole-of-program approach to moderation and quality assurance - enhancement appeared strongly related to the extent that these activities provided opportunities for dialogue among academics.

Conclusion and Discussion
In conclusion, though many accounts of assessment paint a dispiriting picture of current practice (Knight 2002; Price et al 2010) the findings of this Australian project are consistent with Boud’s propositions (2010) for assessment strategies for enhancing learning for the next generation of students and also provide rich examples of practice to illustrate their application across a range of disciplines. The project also identified priority areas for future attention including greater uptake of the factors associated with effective practice and an increase in emphasis on pre-emptive approaches to assessment quality (e.g. rigorous course approval) and to shared understandings of assessment judgements (e.g. Sadler’s ‘calibration’).

Hughes, C., Barrie, S., Crisp, G. and Bennison, A. (2011, October). Assessing and assuring discipline-specific graduate learning outcomes. Roundtable conducted at the conference of the International Society for the Scholarship of Teaching and Learning (ISSOTL), Milwaukee, WI, USA.

This roundtable is intended to draw on international experience to inform a project funded by the Australian Learning and Teaching Council (ALTC) - “Assessing and assuring graduate learning outcomes – principles and practices from within and across the disciplines”. The project is grounded in the need to support productive institutional participation in government plans to transform or ‘revolutionise’ higher education in Australia through the Australian Quality Framework, the establishment of a new regulatory body and the redesign of quality assurance arrangements ‘that puts standards at the centre of the system’ (Vandermark 2010).

As the standards agenda in Australia has much in common with similar projects in other parts of the world (e.g. the European Tuning Project and offshoots in North and South America, QAA subject benchmark statements in the UK, the OECD AHELO project), this roundtable provides an opportunity for an international exchange of experience in addressing emerging issues such as:
1. What constitutes evidence of the achievement of graduate learning outcomes? How can it be obtained? What are the related assessment issues (Hughes and Barrie 2010: Jackson 2002: Yorke 2002)?

2. What undesirable consequences are associated with the assessment and assurance of graduate learning outcomes? How can they be avoided? (Banta 2007: Ewell 2009)

3. How can graduate learning outcomes be assured at the program level? (Barrie, Hughes and Smith 2009)

4. What is an appropriate role for students in the assessment of graduate learning outcomes? How can this be achieved? (Mentkowski 2006)

5. What has or could be the contribution of a Scholarship of Assessment (Rust 2007, 2011) in addressing issues associated with the assessment and assurance of graduate learning outcomes.


This roundtable will invite participants to explore potential assessment implications of institutional adoption of disciplinary graduate learning outcomes such as those developed through the ALTC Learning and Teaching Academic Standards (LTAS) project. Assurance of graduate outcomes – like any other aspect of assurance in higher education - is going to be a multifaceted undertaking, with assessment of relevant student learning a key element. However, while there is a growing recognition in many quarters of the need to assure graduate learning outcomes– there is less clarity about how University assessment can deliver evidence to contribute to this assurance exercise.

The roundtable discussion forms part of the consultation process for a project the ALTC has funded to investigate issues and options for “Assessing and Assuring Graduate Learning Outcomes” (AAGLO). The project will identify and disseminate effective practice and make it accessible across the sector through a highly consultative approach that builds on and continues the productive conversations already taking place among disciplinary communities and the previous work undertaken through the LTAS project.

The roundtable discussion will focus on:

- participant experience in the implementation of disciplinary learning outcomes or standards
- exchange of effective practice in assessment and assurance (and pitfalls to be avoided)
- approaches most likely to gain the maximum benefit for all stakeholders.

Roundtable participants will also be invited to register interest in being included in ongoing consultation activities planned for the life of the AAGLO project.


Projects in many parts of the world have developed descriptions of common frameworks such as subject benchmarking statements to provide points of reference for curriculum design and to facilitate comparisons of student outcomes (see for example Tuning Project
A similar project now underway in Australia is facilitated by the Australian Learning and Teaching Council (ALTC 2010) and involves coordination of discipline communities’ definition of learning outcomes for their graduates.

The strongest evidence of implementation of graduate learning outcomes is provided by assessment but the assessment of overarching program outcomes or graduate attributes has proved problematic. Research literature identifies a number of implementation issues, many associated with assessment – e.g. challenges in developing academics’ skills in curriculum and assessment design, creating supportive institutional infrastructure (e.g. mapping tools), developing a judgement consensus within and across discipline communities (Jackson 2002) and the persistence of disciplinary ‘signature’ (Bond 2007).

Given that the Australian ALTC project is closely informed by the international experience, it is reasonable to anticipate that many of the assessment issues that have emerged will also have relevance in the Australian context. The facilitator of this discussion is a member of an Australian project team investigating the international experience in order to develop principles of assessment and assurance of graduate learning outcomes applicable in the Australian context. Through the practice exchange session she hopes to facilitate and learn from the exchange of experience and insights among session participants.
Appendix B:

Links to AAGLO Summary Papers 1 – 10

**AAGLO Summary 1:** The ALTC AAGLO project and the international standards agenda

**AAGLO Summary 2:** Assurance of graduate learning outcomes through external review

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

**AAGLO Summary 4:** Standardised testing of graduate Learning Outcomes in Higher Education

**AAGLO Summary 5:** Approaches to the assurance of assessment quality

**AAGLO Summary 6:** Policy issues in the effective assessment and assurance of GLOs

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

**AAGLO Summary 8:** e-Assessment issues in the effective assessment and assurance of GLOs

**AAGLO Summary 9:** Whole-of-programme approaches to assessment planning

**AAGLO Summary 10:** The student perspective
Appendix C:

AAGLO Interview Schedule

Demographic information

- Name
- Title
- Institution
- Discipline
- Gender
- Relevant institutional roles (e.g. senior management, coordination, teaching)
- Relevant disciplinary roles (e.g., leadership or membership in disciplinary project activity, participation in accreditation or external review initiatives)

PART A – QUESTIONS FOR ALL INTERVIEWEES

1. Contextual Information
   - name of course (whose assessment is to be discussed)
   - name of program (to which this course contributes)
   - program level
   - compulsory course
   - other background course notes

2. Graduate Learning Outcomes
   - Are you the coordinator of the program or course to be discussed?
   - What guided the selection of learning objectives of this course?
   - How are course learning objectives linked to program learning outcomes?

3. Assessment task
   - Describe a significant assessment task for this course....
     - that allows students to show progress towards the learning outcomes they should achieve by the end of their program?
     - that allows students to demonstrate achievement of the graduate learning outcomes of his program?
   - Was this an individual or group task?
   - Is the task contextualised within a real life or lifelike situation or scenario?
   - Is the task written, spoke, visual or a combination of 2 or more modes?
   - What is the relationship among the different assessment tasks in this course?
   - Why did you choose this task? What makes it particularly effective in assessing graduate learning outcomes? How does it link to other assessment in the course?

4. Assessment judgements
   - What qualities or criteria do you look for in assessing student performance on this task?
   - How are these documented?
   - How closely does the wording of assessment criteria reflect the wording of course learning objectives?
• If group tasks
  - do students receive a group or individual assessment?
  - are student involved in making assessment judgements?
• Is there a way of documenting student development of GLOs progressively throughout the years of the program?

5. The student role
• Does this task offer opportunities for students to become conscious of their progress or achievement of course or program learning outcomes? If so, how?

6. Assuring the quality of assessment tasks and the standards of student work
• What internal processes are in place to assure the quality of assessment tasks?
• Is the program subject to external accreditation?
• Are assessment judgements moderated through external review?
• If so, what is the role of assessment in these processes?
• What processes are in place to assure the standard of student work and comparability/consistency of assessment judgements?
• Can you describe any examples of internal or external processes have led to an improvement in the quality of assessment design or assessment judgements?

7. Final Comments
• What factors influence the way you assess graduate learning outcomes?
• Any other comments?

PART B - ADDITIONAL QUESTIONS FOR ACADEMIC MEMBERS OF ACCREDITATION PANELS

• What type of assessment information do you consider as a member of an accreditation panel?
• What type of assessment information do you find the most convincing evidence of the standard of student learning?
• What is your basis for decisions on standards?
• Have you come across any examples of effective assessment practice or processes to ensure the quality of assessment or standards? If yes, can you describe them?

PART C – ADDITIONAL QUESTIONS FOR REVIEWERS PARTICIPATING IN EXTERNAL REVIEW PROCESSES (e.g. G8 QUALITY VERIFICATION SYSTEM (QVS) PROJECT)

• What assessment information was most useful in helping you undertake the review process?
• Would it have been helpful to look at additional material from earlier or later levels in the program?
• What provided the most convincing evidence of the appropriateness of standards?
• What is your basis for decisions on standards?
• Have you come across any examples of effective assessment practice or processes to ensure the quality of assessment? Please describe
• Would widespread take up of discipline standards influence external review processes or outcomes?
### Appendix D:

**Glossary of assessment task types**

Reproduced with permission from Melanie Fleming, Lecturer in Higher Education, Academic Development Unit. TLC, UNE, last modified July, 2009


<table>
<thead>
<tr>
<th>Assessment Task Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>Students prepare a concise description (usually of the type that accompanies a research paper suitable for publication) in a prescribed format with an appropriate reference list.</td>
</tr>
<tr>
<td>Annotated Bibliography</td>
<td>Students complete a literature search on a topic and document their results. The located journal articles, chapters, books or other items are usually listed in alphabetical order or grouped by subject category. Each entry in the bibliography contains a description indicating its relevance or significance in relation to the topic.</td>
</tr>
<tr>
<td>Blog (or weblog)</td>
<td>Students document the development of their thought processes on a topic using an online journaling tool.</td>
</tr>
<tr>
<td>Case Study</td>
<td>Students undertake a detailed investigation of a person, group or organisation. They are involved in analysing and interpreting a wide range of variables to determine their impact on the individual or group in question (Boyce, 1993).</td>
</tr>
<tr>
<td>Critique</td>
<td>Students provide constructive criticism of a professional work (e.g. a performance, published journal article or work in progress). Evaluation of the work’s effectiveness includes making judgments about the appropriateness of the choices made by the original author or creator of the work.</td>
</tr>
<tr>
<td>Debate</td>
<td>Two teams of students adopt opposing views on a topic and using formal means argue their case. In a true formal debate the winner is determined by the quality of the arguments made by the team.</td>
</tr>
<tr>
<td>Demonstration</td>
<td>Students demonstrate their ability to perform practical skills, (for example laboratory procedures) efficiently and accurately, or their ability to carry out a procedure such as a patient examination.</td>
</tr>
<tr>
<td>Design.Drawing.Plan.Sketch</td>
<td>Students produce sketches or drawings by hand or using computer software.</td>
</tr>
<tr>
<td>Discussion posts (online)</td>
<td>Students contribute comments in the form of posts to topics that form part of an online discussion forum. Topics may be chosen by the lecturer or led by students themselves</td>
</tr>
<tr>
<td>Essay</td>
<td>Students prepare an extended prose response to a question, problem or issue set by a lecturer (Habeshaw, Gibbs and Habeshaw, 1993).</td>
</tr>
<tr>
<td>Assessment Task Type</td>
<td>Definition</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Examination (invigilated)</td>
<td>Students complete a “traditional” examination consisting of an “unseen paper with limited choice” of questions for students (Habeshaw, Gibbs and Habeshaw, p 63). There is no time length implied in this definition as it is used here although examinations of this type are usually between 2 and 3 hours in length. It may contain a mix of multiple choice questions, short answer, extended essay or problem solving and calculation questions. The students may or may not have seen the examination questions before. In this context the examination takes place at the end of the semester and in a formalised examination setting.</td>
</tr>
<tr>
<td>Examination (take home)</td>
<td>Student answer the same types of questions as a traditional examination, however, they are usually asked to answer complex questions of an extended nature with access to key information sources (for example textbooks, data tables and so on).</td>
</tr>
<tr>
<td>Exhibition</td>
<td>Students display their completed work; for example completed project work or poster displays (Habeshaw, Gibbs and Habeshaw, 1993). It can be done in a group situation or individually; students may or may not be present at the time to answer questions from those assessing the work.</td>
</tr>
<tr>
<td>Field Notes/Report</td>
<td>Students complete an excursion e.g. a field trip to a location for the purpose of conducting a detailed primary or first-hand examination. Often practical knowledge is gained from this type of assessment.</td>
</tr>
<tr>
<td>Interview</td>
<td>Students conduct a formal meeting with someone for the purpose of eliciting information from them.</td>
</tr>
<tr>
<td>Journal</td>
<td>A personalized account of what a student has learned as part of the course of study (Habeshaw, Gibbs and Habeshaw, 1993). includes reflections on learning/processes.</td>
</tr>
<tr>
<td>Laboratory/Practical</td>
<td>Develops technical and professional level skills in undertaking experimental or research work. Students learn the process of conducting scientific enquiries.</td>
</tr>
<tr>
<td>Literature Review</td>
<td>Students prepare a systematic, explicit and reproducible review identifying and interpreting the existing body of knowledge of recorded work by researchers, scholars and practitioners (Fink, 1998).</td>
</tr>
<tr>
<td>Log/Workbook</td>
<td>Students provide an objective record of tasks completed, goals met, observations made during a student’s course of study (Habeshaw, Gibbs and Habeshaw, 1993).</td>
</tr>
<tr>
<td>Mid Semester Examination</td>
<td>Is usually an abridged version of a traditional examination (defined previously), although it is usually shorter in duration and carries less weighting as a piece of summative assessment.</td>
</tr>
<tr>
<td>Minutes</td>
<td>Students maintain official records meetings, team decisions and actions to be taken.</td>
</tr>
<tr>
<td>Model/Artefact</td>
<td>Students produce a representation of a system (for example an object made to scale) that allows them to explore the properties of that system.</td>
</tr>
<tr>
<td>Modelling (theoretical)</td>
<td>Students construct a theoretical model of an event or phenomena of interest. It is usually computer based and may also allow prediction of future likely events (especially in the case of quantitative models).</td>
</tr>
<tr>
<td>Oral Examination / Viva Voce</td>
<td>Students verbally explain a concept or problem, to one or more examiners. This is a more formal assessment of both verbal communication skills and content knowledge. There is less student scope regarding the structure of the oral or viva voce as this is largely determined by the examiner/s (Brown, Bull and Pendlebury, 1997).</td>
</tr>
<tr>
<td>Assessment Task Type</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Peer Review</td>
<td>Students make judgments and critical evaluations of the work of their peers</td>
</tr>
<tr>
<td>Performance</td>
<td>Students perform a play, interpretation of a work, original composition or recital. The assessment task may be submitted as a live or recorded production and can be completed as an individual or group task.</td>
</tr>
<tr>
<td>Poster presentation</td>
<td>Students demonstrate their synthesis of the outcomes of research or learning in the form of a self-explanatory poster. The poster makes use of text and graphics to make a visual presentation. Typically, the poster is presented in the professional context of a conference or symposium.</td>
</tr>
<tr>
<td>Portfolio</td>
<td>Students have the responsibility for selecting from amongst a range of pieces of work completed throughout the semester those which they believe are their best. They should present a case justifying inclusion of the pieces (see Biggs p. 189 for more information).</td>
</tr>
<tr>
<td>Practicum/Clinical Placement</td>
<td>Students obtain and demonstrate valuable professional skills and competencies relevant to their area of professional practice. These “field experiences” include opportunities for students to observe and participate in work-based activities.</td>
</tr>
<tr>
<td>Presentation (individual or group)</td>
<td>Students may complete this task individually or in a group. Presentations share many of the same characteristics as a viva or oral examination, however students usually prepare visual materials to accompany and assist their explanation or discussion; students may involve and or question members of the audience.</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>Students solve problems of limited scope using appropriate rules or techniques</td>
</tr>
<tr>
<td>Project</td>
<td>Students complete an in-depth investigation of real world significance. It usually has very specific milestones or deliverables, including start and end dates that are determined prior to the investigation commencing. Students may choose the topic themselves, or it may be set by the lecturer or completed in collaboration with industry.</td>
</tr>
<tr>
<td>Quiz/Test</td>
<td>Students may complete similar tasks to those set in more formal examinations (as defined earlier); in practice the questions are usually a combination of multiple choice or short answer. This is usually a low stakes item of assessment that carries minimal weighting in terms of summative assessment.</td>
</tr>
<tr>
<td>Reflection</td>
<td>Students are required to evaluate their own learning in a way that moves beyond habitual behaviours to critical reflection (Kember, McKay, Sinclair, Kam and Wong, 2008).</td>
</tr>
<tr>
<td>Report</td>
<td>Students complete a more structured form of written assessment such as a scientific report or research paper of the type similar or identical in format to a published article.</td>
</tr>
<tr>
<td>Self Assessment</td>
<td>Students assess their own knowledge, work, skill or ability according to a set of criteria that they may develop themselves or in consultation with the lecturer, they may provide a critique of an earlier piece of assessment and indicate what they have learned since then or how they would do something differently (Brown, Bull and Pendlebury, 1997).</td>
</tr>
<tr>
<td>Simulation</td>
<td>Students participate in a simulation of a workplace or environment to examine how it works. The simulation could be computer based or experiential.</td>
</tr>
<tr>
<td>Student negotiated assessment</td>
<td>Students negotiate the topic and/or the nature or scope of the assessment in consultation with the lecturer or supervisor. This is</td>
</tr>
</tbody>
</table>

Assessing and assuring Australian graduate learning outcomes
<table>
<thead>
<tr>
<th>Assessment Task Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research Paper</td>
<td>Students present the results of their investigation into a selected topic according to the style required in the particular discipline. This may include an abstract, but usually includes sections corresponding to an introduction, data and methodology, results and discussion/conclusion.</td>
</tr>
<tr>
<td>Thesis</td>
<td>The student conducts an extended piece of research designed to set up and defend an intellectual position taken by its author (in this case the student). It is usually on a topic chosen by the student (in consultation with the lecturer in the case of undergraduate students or supervisor in the case of postgraduate students) and subject to external examination.</td>
</tr>
<tr>
<td>Workshop</td>
<td>Students participate in an educational seminar or series of meetings emphasizing interaction and exchange of information among a usually small number of participants. It can involve a brief intensive course of study and discussion by students and is very practical in nature.</td>
</tr>
</tbody>
</table>

References:


Appendix E:

Independent evaluator's report

Assessing and Assuring Graduate Learning Outcomes (AAGLO)
Evaluation Report, August, 2012

Introduction
The following report presents an independent evaluation of the ALTC-funded project
Assessing and Assuring Graduate Learning Outcomes.

The project was located within the current debates in the higher education sector about
how to assure the quality of student learning within the sector. Specifically, there were
three goals:

1. to identify effective strategies for assessing graduate learning outcomes in a way
that genuinely tests the integration of discipline knowledge with more generic capabilities,
2. to identify the continuing processes which provide assurance that the assessment
tasks are valid for the purpose, and that student achievement on those tasks is judged by
appropriate standards, and
3. to raise awareness and to organise active participation in debates concerning the
standards agenda by establishing productive dialogue among key stakeholder groups.

The project team has sought to implement these goals via various activities which run the
gamut from a quite micro focus (particular assessment tasks) to a much more macro focus
(analysing the options at sector level for assuring the quality, or standard, of graduates’
learning).

- It carried out an interview study in seven discipline areas, approaching key
  practitioners in those areas to talk about assessment tasks and assessment regimes
- It developed a series of summary papers in relation to the key issues (readily
  available on the project website)
- It developed an Endnote Library of the relevant scholarly literature
- It held forums in five capital cities to report, progressively, on the findings emerging
  from the interviews and also to enable discussion of the various strategies for
  assuring quality that were currently in play
- In conjunction with these major activities, the team also engaged in conference
  presentations, international interactions, submissions to government, and
  collaboration on a major Issues and Options paper.

My role as evaluator has been both formative and summative. I have had regular contact
with the team, particularly with its Queensland members (Clair Hughes and Anne Bennison).
I attended an early planning meeting of the full team, saw drafts of the Summary Papers,
was present for three of the five capital city forums and was part of the subsequent de-
briefing of these events.

The sources of evaluative information used in this report include: my own observations,
comments from some members of the project’s reference group, a consideration of the
feedback sheets collected from attendees at each of the five forums, and notes taken by me
as I circulated at the forums which I attended.
In making the following evaluative comments, I shall look both at the conduct and at the achievement of the project.

**The Conduct of the Project**

The project was well-organised and well-managed. Although each of the three academic members of the project team faced new or additional employment demands in the course of the project, focus and timelines were well-maintained via a relationship with the project manager. The project activities were largely consistent with what was initially proposed by the team.

- There was a considerable amount of scholarly review and national and international liaison in order to clarify the ground of the project. This was particularly important given the number of projects, both concurrent and recent, which had the potential to overlap with this work. By virtue of this work, the project positioned itself in a timely and strategic way at the centre of the standards debate. This resulted, as planned, in a series of summary papers which picked up the key issues in relation to the standards agenda, making the debates accessible to an audience who may not have either policy or assessment expertise.

- The discipline focus of the interview study was seen as crucial given the emphasis in recent scholarship on discipline-specific manifestations of broader graduate capabilities. Even so, there was no intention to replicate the work of the LTAS projects, but rather to gather accounts of practice and process in order to build guidelines and to illustrate them with examples. In the event, for various logistical reasons, there were slightly fewer disciplines and slightly fewer interviewees than originally planned. There is no suggestion, however, that the interview data were compromised or limited by this – the data were copious and rich, albeit with perhaps less innovation and variety than anticipated, and the team used NVivo to analyse and clarify what they had gathered.

- The capital city forums were designed both to introduce the topic of the standards agenda and to reveal the findings from the interviews, but also to give participants a chance to ask questions and share views. These were well designed for the purpose. Each forum featured a provocative guest speaker, and project team members used their networks to ensure that a range of institutions and of institutional roles was represented in the audience.

- The reference group for the project was chosen, sensibly, in such a way as to have representation from all relevant stake-holder groups. Although there was input from the reference group early in the project, there was considerably less as the project moved into its implementation stages. One member, who thought the project very valuable, felt that the reference group might have been able to add clout to its dissemination process. It is clear, however, that one of the risks incurred when people are appointed in a representative way is that the actual people change – at least five of the representative members, in particular, had changed status and employment by the time I attempted to consult them as part of the evaluation process.

As already indicated, the scope of this project was ambitious. The team was seeking discipline-specific classroom-level information about effective assessment tasks, and at the same time seeking to use this information to compile guidelines and exemplars which would
contribute to a deeper understanding of how best to assure the quality and standards of learning at the sector level. Given this scope, it was important that the activities proposed for the conduct of the project should hold together the focus on practice with the focus on policy. This weaving together was not just a matter of carrying out the planned activities, but also of consciously using them to make the connections between practice and policy clear. Many of the responses to the forums indicate that this was successfully achieved.

The Project’s Achievement

This has been a significant project, partly because of its timeliness. The project started out with three goals, and achievement on each of them has been substantial.

Contribution to sector debate

At the macro level, one of the project’s goals was to raise awareness around the standards debate in higher education. I believe that this has been an area of outstanding achievement for the project. This is a continuing, fraught debate both nationally and internationally, although it is clear that not all discipline-level academics are aware of the terms of this debate. This project has canvassed, via its activities, the various options that are in play for ensuring that Graduate Learning Outcomes meet the standard expected. The forums have seen discussion of the main options: standardized testing, rubrics, portfolios, external examiners, and consensus moderation.

For many of the attendees, Trudy Banta’s presentation, based on her own wealth of experience, provided a first opportunity to understand the issues surrounding standardized testing at a much deeper level; similarly Royce Sadler’s presentation broke open the challenge of how to assess whether GLOs were actually being achieved, demonstrating ‘the futility of standardized tests to identify this’ (forum feedback sheet), and canvassing the potential of more local, discipline-based consensus. Forum attendees were clearly appreciative of the overview provided; many also mentioned the usefulness of ‘touching base’ with colleagues in other institutions in relation to these issues. In later forums, where participants were directed more explicitly to the Summary Papers, feedback on these suggested that they would take their place as valuable resources in local and institutional policy-making. The papers represent, at this point in time, one of the most worthwhile legacies of this project. (Two or three more papers remain to be published on the website at the time of writing.)

Contribution to practice

At the more micro level, the project had a couple of goals (1 & 2 as indicated at the beginning of this report), gathering information about tasks and processes via a discipline-based interview study. Many Australian academics are still new to the idea of genuinely integrating discipline knowledge-and skills outcomes with more generic graduate learning outcomes. The purpose of this data, therefore, was to use it to shed light on the kinds of assessment tasks which can assess these outcomes and to derive guidelines and examples to assist thinking and practice in this area. It was also to identify processes, both pre and post assessment, which would ensure that standards were monitored and maintained. The project team was conscious that a number of disciplines had developed outcome statements and resources via the LTAS project. The idea was not to replicate this, but rather to explore tasks and strategies as they were actually being implemented.

The findings from this exploration were progressively presented at the forums and are presented at greater length in Summary Paper 8 (still in draft form). Forum attendees were
avid for these findings. There was great interest in the kinds of tasks which, in the judgment of discipline academics, were capable of carrying the weight of integrating substantive knowledge and generic outcomes. A number of feedback sheets suggested that attendees would like to have seen even more examples. (In fact, they were made aware at the forums that examples of both tasks and processes were, or would be, available in the Summary Papers.) It is clear that academics are keen to ensure that the ultimate Standards solution is discipline-based, and that the policies governing the monitoring and maintenance of those standards should be local to the discipline, albeit transparent and accountable. The forums did excellent work in helping to clarify and coalesce some of this thinking.

**The project’s products**

I have already indicated above that the Summary Papers make a strong and timely contribution to sector thinking in this area, particularly the first four papers (on the standards agenda, external review, work-based learning standardized testing). It should be noted that the subsequent papers, which focus more on the findings of the discipline study, help to flesh out a positive approach to implementation of the standards agenda. Summary Paper No.6 is especially noteworthy for the clarity of its analysis of policy issues and policy solutions. Papers that are still in preparation as the project moves towards its conclusion will address e-assessment, student perspectives and whole-of-program strategies.

The other product that should be mentioned is the Endnote Library, a valuable and motivating resource for discipline academics who need some scholarly support to get into the Scholarship of Teaching in the area of assessment. One of the reference group members praised this particularly, suggesting that many academics were operating without a good conceptual grasp of assessment.

**Conclusion**

This has been a project where both conduct and achievement have been strong. While it is not clear nationally what formal mechanisms for ensuring standards will ultimately be adopted, the project has provided a scholarly overview of the options and has explored the meaning of assessing and assuring GLOs (Graduate Learning Outcomes) at discipline level. It has produced resources both for discussion and for the improvement of practice. Some of the responses to the project have made clear that there is also necessary further work to be done. There are broadly-based concerns that best practice is incompatible with already over-stretched workloads and with an increasing use of casual staff, as well as concerns that academic staff do not ‘own’ this agenda, but rather see it as part of a growing managerialism. Clearly, these are big challenges. Although the resources developed by the project go some way towards addressing these problems, it probably remains the case that further dissemination and development are still required. It is for this reason that it is very pleasing to note that the project has obtained permission to spend unused budget (itself a testament to the efficiency of the project’s management) to fund a couple of workshops in conjunction with an Alverno College team. Alverno has long been acknowledged as modelling excellent practice in the area of integrated curriculum and assessment – these workshops will be a welcome addition to the further dissemination of the project’s work.

Margaret Buckridge
Griffith Institute for Higher Education
Griffith University