Assessing Higher Education Learning Outcomes in Civil Engineering: the OECD AHELO Feasibility Study

Roger Hadgraft, Jacob Pearce[a], Daniel Edwards[a], Julian Fraillon[a], Hamish Coates[a]
RMIT University, Australian Council for Educational Research[a]
Corresponding Author Email: *Roger.Hadgraft@RMIT.edu.au*

BACKGROUND

Higher education is increasingly a global business. At present, university ranking schemes are heavily reliant on research indicators, while students are likely looking for an excellent teaching and learning environment. Aware of this discrepancy, the OECD has funded the AHELO project.

PURPOSE

AHELO was designed to test the feasibility of an international assessment of higher education learning outcomes. The two test disciplines are civil engineering and economics, together with an assessment of generic skills. The civil engineering test will be reported in this paper.

More than a ranking, AHELO is a direct evaluation of student performance. It is intended to provide data on the relevance and quality of teaching and learning in higher education. The test aims to be global and valid across diverse cultures, languages and different types of institutions.

DESIGN/METHOD

The AHELO project was developed between 2008 and 2012. Preliminary work focussed on developing the assessment framework, which builds upon the frameworks used for accreditation across the world (EA, Washington Accord, ABET, etc), with outcomes grouped into Basic/Engineering Sciences (both branch-specific and general), Engineering Processes (Analysis, Design and Practice), and Generic Skills. It was decided that the test would focus on the engineering processes (using constructed response tasks) and basic and engineering sciences (using multiple choice questions to test knowledge of engineering fundamentals).

Sample questions were developed and reviewed by the International Reference Panel in Oct 2010. During 2011, a pilot of the chosen constructed response tasks plus MCQs was run around the world with approximately 10 students from each university from each participating country (more than 300 responses from Australia, Japan, Canada, Colombia and the Slovak Republic). These results and student feedback were used to modify some questions on the basis of Discrimination Factors.

RESULTS

In 2012, a full scale test has been conducted and is reported in this paper. Nine countries have participated with a total of more than 6,000 students. Results are currently being analysed to examine similarities and differences between the participating countries and as indicators of areas that need more attention in engineering curricula.

CONCLUSIONS

AHELO is the first international, standardised test that attempts to measure student outcomes from engineering programs across the world irrespective of language and cultural differences. This is important in an increasingly global engineering job market. The paper reports the procedures taken to ensure a valid and reliable test.

KEYWORDS

Assessment, graduate outcomes, OECD, international

Introduction

Higher education is increasingly a global industry and while rankings of institutions are now a dominant feature of this industry, current international comparisons rely almost exclusively on research metrics. The OECD Assessment of Higher Education Learning Outcomes (AHELO) project seeks to test the feasibility of using international assessments to probe the quality of higher education around the world. As highlighted by the OECD (2012a), *no reliable international data exists on the outcomes of learning and the few studies that do exist are nationally focused. Available rankings [*of universities*] reflect neither the quality of teaching and learning nor the diversity of institutions*.

The aim of the AHELO Feasibility Study has been to assess whether reliable cross-linguistic, cross-cultural and cross-institutional comparisons of higher education learning outcomes are scientifically possible and whether their implementation is feasible. The AHELO Feasibility Study has scientific and practical dimensions. Two research questions underpin the study:

a. Is it scientifically possible to produce cross-linguistic, cross-cultural and cross-institutional valid comparisons of higher education learning outcomes?
b. Is it feasible to implement a valid cross-linguistic, cross-cultural and cross-institutional assessment of higher education learning outcomes?

The provision of common objective data on graduate capabilities has the potential to play a significant role in assisting institutions to monitor and enhance the standard of their educational provision on a global scale. This will be increasingly important to developing countries as they seek to compete in global markets. Institutions need more information on actual learning outcomes to assist with this international positioning.

Although it is too early to make conclusions pertaining to the research questions posed here, it is possible to discuss the processes undertaken to ensure that it will be possible to answer these questions at the conclusion of the AHELO feasibility study. This paper gives a brief introduction to the engineering strand of AHELO and lists the principles that underpin the validity and reliability of the data collection and the overall evaluation of the feasibility of this study. It then provides an overview of the AHELO Engineering Assessment, detailing the phases in its development, outlines the implementation of testing, and describes the analysis processes employed. Elements of this paper are based heavily on a number of key documents prepared for the AHELO Feasibility Study, including the AHELO Engineering Assessment Framework (OECD, 2011), the AHELO Analysis Plan (OECD, 2010a), the AHELO Assessment Design (OECD, 2010b) and the AHELO Engineering Assessment Development Report (OECD, 2011b).

The AHELO Feasibility Study covers three strands (Generic Skills, Economics and Civil Engineering) as well as three contextual dimension survey instruments. The focus of this paper is on the Civil Engineering strand, which Australian universities participated in as part of the international study. AHELO is world-first in attempting to determine the feasibility of assessing the learning outcomes of university graduates across multiple countries and cultures. The study is being conducted for the OECD by an international consortium led by the Australian Council for Educational Research (ACER) and it officially commenced in January 2010 and concludes at the end of 2012.

The AHELO Engineering Strand

Over the past few decades, the profession of engineering and the roles of engineers have changed rapidly. The problems faced by engineers in today’s world are increasingly complex and require engineers to have both strong technical knowledge and skills, and understanding of relevant environmental, social, economic and cultural contexts. In addition, as for other professions, engineers are expected to be good communicators, to work effectively in interdisciplinary teams, to conduct themselves ethically and professionally, and to be able to constantly update and improve their technical and personal skills. These generic skills areas are well covered in the engineering education and professional literature (e.g. Bons & McLay, 2003; Walther, Mann & Radcliffe, 2005; Gill, Mills, Sharp & Franzway, 2005).

Such changing requirements are continuous, but they are also identified formally in reviews that are undertaken periodically by national professional peak bodies. The past decade or so has seen such reviews in the United States (National Academy of Engineering, 2005), the United Kingdom (Royal Academy of Engineering, 2007), and Australia (Institution of Engineers Australia, 1996; King, 2008). The recommendations in such reviews are usually focussed on changing university-level engineering curricula and pedagogy, revising professional accreditation requirements, and intensifying connections to both professional practice and to school education.

The common trend in modernising engineering education is to increase the focus on graduates' competencies in project work, communication, and collaborative skills, and increase their understanding of ethical practice in the contexts in which engineering problems and projects exist (Boles, Murray, Campbell & Iyer, 2006; Walkington, 2001; West & Raper, 2003). Underpinning much of the curriculum redesign and revision are the agreed graduate outcomes as required by national engineering accreditation processes. Over the past decade, these have increasingly been framed in terms of graduates' learning outcomes and competencies, rather than focusing on input measures. Thus, engineering curricula are specified in terms of expected outcomes, rather than subject content. There is also substantial commonality in the statements of these terms as used internationally by bodies concerned with both professional and education accreditation: Washington Accord, 2009; European Network for Accreditation of Engineering Education (ENAEE), 2008; USA Accreditation Board for Engineering and Technology, ABET 2008; Engineers Australia (EA), 2006; UK Quality Assurance Agency (QAA), 2006; and EU Tuning Process (Tuning Project, 2004).

Design and method

While educational processes and outcomes in engineering are relatively well defined, a need remains to produce robust data on learning outcomes and graduates' potential for subsequent success in work and further study. The assessment of engineering capability undertaken as part of the AHELO Feasibility Study provides an opportunity to contribute to a more evidence-based approach to ascertaining quality in higher education. Within the context of the AHELO Feasibility Study, the consortium responsible for the engineering strand, led by the Australian Council for Educational Research (ACER), explored the feasibility of directly measuring learning outcomes in engineering across different cultural, linguistic and institutional contexts.

The emphasis of this project on *feasibility* is of utmost importance in conceptualising the study. The terrain that this work enters is new, and to-date unexplored. As such, the design and methodology of this study has necessarily been a learning process in itself. Establishing the feasibility of this study rests not only in the raw data and outcomes produced in the assessments themselves, but also in the design, implementation, scoring and application of the work.

A number of core standards have been developed through research, consultation and experience to provide foundations for AHELO. They are intended to be relevant, succinct, measurable and enforceable. Where possible and relevant, the AHELO Technical Standards rest on and reference broader standards produced for educational measurement, assessment and evaluation. In particular, readers are referred to various International Test Commission standards (see: ITC, 2000; 2005; 2010; 2011), the IEA standards (Martin, Rust & Adams, 1998) and the OECD PISA standards (OECD, 2010c).

The standards have been developed with major and inter-related principles in mind:

a. Relevance: Technical work must be well positioned within salient educational, policy and practical contexts and must be designed and conducted to enhance the relevance of AHELO.
b. Consistency: Data should be collected in an equivalent fashion in all Higher Education Institutions (HEIs) and systems, using equivalent assessment materials. A comparable sample of the student population should be assessed under test conditions that are as similar as possible.
c. Precision: Data collection and submission practices should leave as little room as possible for variation or error, whether systematic or random. This includes errors that could be caused by variations in testing environments for different groups of students, and errors that could occur during data preparation.
d. Generalisability: Data are collected from specific individuals, in a particular situation, and at a certain point in time. The selection of individuals, test materials, tasks and all other testing conditions

should be highly standardised.

e. Timeliness: In AHELO, it is imperative that all activities are conducted within given timelines in order to meet HEI, national and international schedules, and reporting deadlines.

The Engineering Assessment Framework

The AHELO Engineering Assessment Framework was the guiding document during instrument development. Materials were developed in direct consultation with the key competencies, which are explicated in more detail in the framework itself. The Assessment Framework was based on the AHELO-Tuning document (OECD 2009a), the AHELO Engineering Assessment workshop held at ACER in Melbourne in January 2010, the TECA document (Coates & Radloff, 2008), and broader AHELO technical materials. It was informed by the processes and practices adopted in the PISA literacy surveys (e.g. OECD 2009b), and the combined expertise of consortium staff. Subsequent drafts incorporated review comments from consortium members and Engineering Expert Group members.

The Engineering Assessment Framework defines the domain to be tested. Specifically: first-cycle engineering competency is the demonstrated capacity to solve problems by applying (i) analysis using basic engineering and scientific principles, (ii) engineering design and (iii) engineering practice skills, which correspond to the three components of Engineers Australia's Stage 1 Competency Standard (Engineers Australia, 2011). These skills are supported by generic skills such as communication and teamwork, which are assessed through the generic skills component of AHELO.

An assessment instrument must tap into the different aspects of a test taker's proficiencies. Engineering competency entails applying relevant skills and knowledge in solving problems of interest to an engineer. Recognising that engineering problems occur in a diverse array of situations, a representative sample of engaging contexts for items were chosen to exercise the constituent components of engineering competency. The contexts in which students need to demonstrate their skills and knowledge include both those specific to civil engineering and those more generally applicable across a number of fields of engineering.

Overview of the Engineering Assessment instrument

Development of the Engineering Assessment took place between July and December 2010, focus group qualitative analysis was undertaken in the first half of 2011 and revision of the Engineering Assessment testing took place between July 2011 and January 2012. The development of both the Engineering Assessment Framework and Engineering Assessment was undertaken by a consortium of organisations in Australia, Japan and Italy, led by ACER, and incorporated the expertise of engineering educators and specialists from around the world.

The AHELO Engineering Assessment includes a broad sample of items covering a range of difficulty that enables the strengths and weaknesses of populations and key subgroups to be determined with respect to the components of engineering competency. The aim of the assessment is not to implement another final-year university engineering exam. Rather, the aim is to assess a final-year engineering student's ability to think like an engineer and to display the non-technical competencies that practising engineers must possess.

Each student undertaking the 90 minute assessment answers one constructed response task and 25 multiple choice questions. Item response formats include:

a. multiple choice: simple and complex multiple-choice items (the latter are answered by selecting an option from each list of choices);
b. constructed response: comprising several questions made up of short constructed-response, e.g. numerical or short text; extended constructed response, e.g. creation of flow charts, designs, dot-pointed specifications, and longer written responses.

Drafting and preparation: Constructed Response Tasks

Wide-ranging research into civil engineering contexts was conducted by the ACER, resulting in 12 initial constructed response (CR) tasks, such as a major hydroelectric storage, bridge failure, concrete construction, flood prevention structures and others. These problems were drawn from textbooks, exam papers, concept inventories, civil engineering professional assessments (such as certification examinations, and civil engineering projects worldwide, based on ~~internet~~Internet research.

Materials were also submitted to ACER by consortium partners and stakeholders, following an AHELO Engineering Assessment Workshop held at ACER in Melbourne in January 2010. Based on the resources collected, tasks were drafted and developed to fit the specifications of the Assessment Framework. Initial drafts were interrogated by item developers from the project team in Australia, Japan and Italy in panel sessions where each element of the item was scrutinised and revised according to criteria such as: content validity, clarity and context, format, test takers perspective, and scoring options. The CR tasks focussed on the 'above content' areas of the framework—the engineering processes (analysis, design and practice).

Drafting and preparation: Multiple Choice Items

The multiple choice (MC) items were designed as a set of items that prompt students to demonstrate their competency in Basic and Engineering Sciences. They were included to provide a fast and efficient way to collect data on each student's engineering knowledge, understanding and skills. They complement the CRs in providing an instrument that covers a wide range of *basic engineering knowledge*, along with specific above-content competencies. They also verify the robustness of competencies assessed by the CRs. Since application of basic engineering and scientific principles requires their mastery, results from multiple choice items should indicate whether students have in fact developed the fundamentals that underlie competencies required to analyse and synthesise solutions to complex engineering problems.

Development of the multiple choice items began with licensing examinations developed by the Institution of Professional Engineers Japan and the Japan Society of Civil Engineers. An extensive list of translated items from this source was presented by NIER to the Engineering Expert Group for their review. Forty items were selected and revisions and further developments were advised, with Engineering Expert Group members approving final versions.

Expert Group review and revision

Drafted and revised items were subjected to review and discussion with an international panel of experts in civil engineering. This panel included representatives from across the world with expertise in engineering education and from within industry (OECD, 2011a). During the review process, items were culled and revised. A final set of items was established following this review and prepared for implementation in focus groups among a number of countries. This process included translation and adaptation of test materials.

Focus Groups with students

The drafted assessment materials were tested with university students in Australia, Canada, Japan, Colombia and the Slovak Republic in 2011. In these structured focus groups, students from the target population undertook a sample of the drafted assessment materials and then discussed the items facilitated by a moderator who followed a set of prescribed discussion prompts. Detailed reports from universities involved were collated and coordinated at the national level and sent to the assessment development team to assist in the evaluation of the draft materials. Some psychometric analysis of the focus group results was also undertaken to guide revision and selection of final items.

Preparation of final assessment

Revision based on the focus group results was implemented by the assessment development team in Australia, Japan and Italy. A core set of final items was produced and inserted into the AHELO online test delivery system.

Translation and adaptation

The final step in the development of the AHELO Engineering Assessment involved a complex process of translation and adaptation of the test materials into the various languages of participating countries. In total, nine countries implemented the engineering assessment in the feasibility study: Australia, United Arab Emirates (Abu Dhabi), Canada (Ontario Province), Colombia, Egypt, Japan, Mexico, Russia and the Slovak Republic. These countries represent a diverse range of languages and alphabets.

For each country the test materials were translated and adapted for each national context by cApStAn and verified by engineering experts. cApStAn specialises in translation work of this kind.

Implementation of the assessment

Each country involved in the study invited universities with a Civil Engineering bachelor program to participate in the feasibility study. In general, the AHELO project team suggested about ten universities per country and up to 250 students from each institution. The actual implementation country-to-country varied as a result of levels of institutional participation and the numbers of students varied institution-to-institution due to the size of engineering schools and the ability of universities to encourage participation. Testing of students took place between February and June 2012 across all participating countries. All testing was facilitated via a secure, online assessment platform.

Scoring

Scoring of the assessment responses to MC items was undertaken automatically through the online system. For CR tasks, the process of scoring results was more sophisticated. For the CR tasks, rubrics were provided for every discrete question. Each rubric indicates the number of points available and the required student responses for each score point. Rubrics also indicate example student responses. Rubrics were developed by the assessment development team and revised based on feedback from the Engineering Expert Group as well as from Lead Scorers from each of the participating countries.

Following the scoring rubrics, the marking of AHELO Assessment responses was undertaken in each country under the guidance of the national Lead Scorer. An online scoring platform was used for this process, ensuring scripts remained anonymous and 20 per cent of scripts were automatically double marked. Data was collected from the double marked items and analysed to establish levels of consistency across scorers.

Analyses of results

Similar to the PISA reporting practice (OECD, 2009), results will be reported on a scale constructed using a generalised form of the Rasch model (Rasch, 1960). The form of the Rasch model that is used in this study employs the scores obtained by students to produce estimates for both the difficulty of items and the ability of students on a single real-valued scale.

The model and scaling methods allow the linking of measures of student performance with data collected in a contextual survey instrument completed by participating students so that characteristics such as gender, socioeconomic standing, geographical location and institution attended can be controlled for in analyses. This enables statistical comparisons of population means between students grouped by these background factors. For example, comparisons of performance between participating institutions will be possible.

Reliability and Validity analyses

As a large and innovative study it is necessary to analyse the validity of several facets of the AHELO method and data. These analyses focus on the validity of instruments and applications of use, and on whether results generalise across reporting and contexts.

In addition to item response modelling, other classical analyses are being conducted to generate reliability and validity statistics, and test the efficiency of alternative scoring methods. Item-total statistics will be generated for each national group. A series of reliability generalisability (RG) studies will be conducted to review whether the errors of measurement are stable across contexts. Reliability estimates will be produced for the student and institutional levels.

Item fit to the measurement dimension will be assessed using a range of item statistics. The weighted mean-square statistic (infit), a residual based fit statistic, will be used as a global indicator of item fit. Weighted infit statistics will be reviewed both for item and step parameters. The analysis of item fit and the estimation of item parameters will be carried out with the ACER ConQuest software, designed for fitting both unidimensional and multidimensional item response and latent regression models.

Item response modelling will be used to assess the ‘targeting’ of the test to respondent cohorts. This involves checking whether the distribution of item difficulty maps well against the distribution of respondent capability.

In addition to this, item characteristic curves (ICC) will be generated for every item, which provide a graphical representation of item fit across the range of student abilities for each item (including dichotomous and partial credit items). The functioning of the partial-credit scoring guides will further be reviewed through investigation of the proportion of responses allocated to each response category and the differences in mean abilities of students by response category.

The cross-contextual validity of the test items will also be explored by assessing differential item functioning (DIF) (for groups that have sufficient sample). Specifically, Item Response Theory (IRT) will be used to detect variance of item parameters across contexts. Such variance indicates that groups of students with the same ability have different probabilities of responding correctly to an item. This is commonly referred to as 'item bias' as it indicates that the probability of successful performance is a function of group membership as well as individual ability.

AHELO instruments use a range of item types. Psychometric analyses will check the synergies between these different item types. Analyses will be conducted to explore whether there are any off-dimensional interactions between assessment items and instruments and student or institutional groups. These analyses will take account of item content and difficulty as well as respondent characteristics.

Analyses will be conducted of the generalisability of constructed response task data across national and linguistic contexts. These analyses will review:

a. whether the statistical distribution of scores from subjectively scored CR tasks varies (in terms of effect size units) across national, institutional and disciplinary contexts;
b. the extent to which construct generalisability across institutions, and particularly national/linguistic contexts, implies that scoring rubrics have been interpreted invariantly;
c. inter-rater reliability, using variance decomposition analyses based on cross-rater reliability statistics collected via online systems during fieldwork, and reviewing whether standards were met; and
d. consistency of scoring and scoring outcomes across national contexts, established by cross-scoring translated response tasks or capturing agreement between Lead Scorers.

Discussion

Between February and June 2012, 6,000 students sat the AHELO Feasibility Study Engineering Assessment. These students were spread across nine countries and about 90 universities. Within Australia, eight universities participated and nearly 200 students took part in the assessment. In addition to this, contextual data was collected from all universities involved in the international study.

Scoring of results was undertaken within each participating country based on the scoring processes and rubrics discussed earlier. Australia and Canada worked together in this scoring process to swap student responses and cross-validate scoring. A more detailed study of the reliability between scorers from multiple countries is also underway.

Final analysis and collation of results is currently being undertaken by the AHELO Consortium, with the delivery of results to the OECD expected in late 2012. An international conference hosted by the OECD in March 2013 will provide a platform for the full review of outcomes from the project.

While the student learning outcomes data is an important facet of AHELO, as emphasised in this paper, the processes of the development of the frameworks and assessments are also an extremely important outcome. It is this outcome that we have described and reported on in this paper.

On this basis, key outcomes to highlight include (OECD, 2012b):

- the development and international validation of the Engineering Assessment Framework and the Engineering Assessment instrument;
- The independent translation, adaptation and verification of the AHELO Engineering Assessment to robust international standards, in a transparent process in all participating countries;
- The implementation of the AHELO Engineering Assessment across different countries, languages, contexts and systems.

The OECD concludes at this stage that 'based on the feasibility study findings to date, there are genuine and sound indications that much of AHELO is feasible' (OECD, 2012c).

Importantly, there are other outcomes from the study, based around the practice of assessment in higher education, which have offered the opportunity for reflection as a result of this work. This includes how we treat assessment data in Australian universities and how it might be used in the new standards regime to be administered by the Tertiary Education Quality and Standards Agency (TEQSA).

Recommendations and Conclusion

It is clear from this study that there are a range of statistical tools that can and should be used to provide confidence that the learning outcomes being assessed in Australian engineering courses and programs have both validity and reliability. Likewise, there are sound psychometric assessment practices and measures used in this study (described above in *Reliability and Validity Analyses*) that can and should be used routinely in examination processes.

At the course, program, school and university level, this is important, particularly when addressing quality assurance issues. At a national scale, it might be reasonable for TEQSA to enquire whether these kinds of analyses are routine within each university and across each discipline. The reliability and validity tests and tools mentioned in this paper offer a starting point for the objective measure of learning outcomes and quality assurance standards within higher education institutions.

The AHELO feasibility project has been an opportunity to peer under the bonnet of engineering assessment. In the process, the learning outcomes have been carefully scrutinised on multiple occasions by panels of experts in order to develop focussed assessment tasks that can be completed in 90 minutes. The detailed analysis of items and country performance will be the subject of a future paper. For now, this study offers many insights into the rigorous educational statistical analysis tools that can be used in higher education for measuring these outcomes. Incorporating routine application of these tools in our universities in the future will offer greater quality assurance and facilitate the improvement of teaching and learning for the benefit of students.

References

ABET (2008). *Criteria for accrediting engineering programs: effective for evaluations during the 2008-2009 cycle*. Engineering Accreditation Commission, Retrieved 2 Oct, 2008, from http://www.abet.org/forms.shtml

Boles, W., Murray, M., Campbell, D. & Iyer M. (2006). Engineering the learning experience: Influences and options. *Proceedings of the 17th Annual Conference of the Australasian Association for Engineering Education*, Auckland, 10-13 Dec.

Bons, W. & McLay, A. (2003). Re-Engineering Engineering curricula for tomorrow's engineers: Engineering education for a sustainable future. *Proceedings of the 14th annual conference for Australasian Association for Engineering Education and 9th Annual Women in Engineering Forum*, Melbourne, 29 Sept-1 Oct.

Coates, H. & Radloff, A. (2008). *Tertiary Engineering Capability Assessment: Concept Design*. Canberra: Group of Eight Universities.

Engineers Australia (2011). Stage 1 Competency Standard for Professional Engineer. Retrieved 10 July 2012, from http://www.engineersaustralia.org.au.

European Network for Accreditation of Engineering Education (2008): *EUR-ACE® Framework Standards for the Accreditation of Engineering Programmes*. Retrieved 16 Jul, 2009, from www.enaee.eu

Gill, J., Mills, J., Sharp, R. & Franzway, S. (2005). Education beyond technical competence: gender issues in the working lives of engineers. *Proceedings of the 4th ASEE/AaeE Global Colloquium on Engineering Education*, Sydney, 26-29 Sept.

Institution of Engineers Australia. (1996). *Changing the Culture: Engineering education into the future*. Canberra: Institution of Engineers Australia.

International Test Commission (2000). *International Guidelines for Test Use*. Retrieved 2 Feb, 2012, from http://www.intestcom.org

International Test Commission (2005). *International Guidelines on Computer-based and Internet Delivered Testing*. Retrieved 2 Feb, 2012, from http://www.intestcom.org/guidelines

International Test Commission (2010). *International Test Commission Guidelines for Translating and Adapting Tests*. Retrieved 2 Feb, 2012, from http://www.intestcom.org

International Test Commission (2011). *ITC Guidelines for Quality Control in Scoring, Test Analysis, and Reporting of Test Scores*. Retrieved 2 Feb, 2012, from http://www.intestcom.org

Martin, M.O., Rust, K & Adams, R.J. (1998). *Technical Standards for IEA Studies*. Amsterdam: IEA Secretariat.
National Academy of Engineering (2005). *Educating the Engineer of 2020*. Washington: National Academies Press.
OECD (2009a). *A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering*. Retrieved 12 Feb, 2010, from http://www.oecd.org/dataoecd/46/34/43160507.pdf
OECD (2009b). *PISA 2009 Assessment Framework*. Retrieved 12 Feb, 2010, from http://www.oecd.org/dataoecd/11/40/44455820.pdf
OECD (2010a). *AHELO Feasibility Study Analysis Plan*. Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2010b). *AHELO Assessment Design.* Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2010c). *PISA 2012 Technical Standards*. Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2011a). A Tuning-AHELO Conceptual Framework of Expected/Desired learning Outcomes in Engineering. *OECD Publishing Education Working Papers No. 60*. Retrieved 23 Jul, 2012, from http://www.oecd.org/dataoecd/46/34/43160507.pdf.
OECD (2011b). *Engineering Assessment Development Report*. Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2012a). *Testing student and university performance globally: OECD's AHELO*. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2012b). *AHELO Interim Feasibility Report*, Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
OECD (2012c). *AHELO Newsletter, July 2012*, Paris: OECD. Retrieved 23 Jul, 2012, from http://www.oecd.org/edu/highereducationandadultlearning/ahelodocuments.htm
Quality Assurance Agency (QAA). (2006). Subject benchmark statement: Engineering. Retrieved 2 Oct, 2008, from http://www.qaa.ac.uk/academicinfrastructure/benchmark/statements/Engineering06.pdf
Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen, Danish Institute for Educational Research.
Royal Academy of Engineering (2007). *Educating Engineers for the 21st Century*. London: RAE.
Tuning Project (2004). *Tuning educational structures in Europe: Generic competencies*. Retrieved 30 Apr, 2009, from http://www.tuning.unideusto.org/tuningeu/index.php?option=content&task=view&id=183&Itemid=210Walkington, J. (2001). Designing the Engineering curriculum to cater for generic skills and student diversity. *Australasian Journal of Engineering Education, 9*(2), 127-135.
Walther, J., Mann, L. & Radcliffe, D. (2005). Global Engineering education: Australia and the Bologna Process. *Proceedings of the 4th ASEE/AaeE Global Colloquium on Engineering Education*, Sydney, 26-29 Sept.
Washington Accord (2009). *Graduate Attributes and Professional Competencies*. Retrieved 12 Feb, 2010, from http://www.washingtonaccord.org/IEA-Grad-Attr-Prof-Competencies-v2.pdf
West, M. & Raper, J. (2003). Cultivating generic capabilities to develop future engineers: An examination of 1st year interdisciplinary Engineering projects at the University of Sydney. *Proceedings of the 14th annual conference for Australasian Association for Engineering Education and 9th Annual Women in Engineering Forum*, Melbourne, 29 Sept-1 Oct.

Copyright statement

Copyright © 2012 Roger Hadgraft, Jacob Pearce, Daniel Edwards, Julian Fraillon, Hamish Coates: The authors assign to AAEE and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to AAEE to publish this document in full on the World Wide Web (prime sites and mirrors), on Memory Sticks, and in printed form within the AAEE 2012 conference proceedings. Any other usage is prohibited without the express permission of the authors.