Unclassified **EDU/IMHE/AHELO/GNE(2012)6**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **12-Mar-2012**

English - Or. English

DIRECTORATE FOR EDUCATION
INSTITUTIONAL MANAGEMENT IN HIGHER EDUCATION GOVERNING BOARD

EDU/IMHE/AHELO/GNE(2012)6
Unclassified

Group of National Experts on the AHELO Feasibility Study

INTERIM FEASIBILITY STUDY REPORT - EXECUTIVE SUMMARY

Highlights - the AHELO Feasibility Study: What have we learned so far?

9th meeting of the AHELO GNE

Paris, 19-20 March 2012

This document was prepared by the ACER Consortium. It presents the highlights of document EDU/IMHE/AHELO/GNE(2012)5.

The AHELO GNE is invited to TAKE NOTE of this document.

Contact:
Consortium: ahelo@acer.edu.au
OECD Directorate for Education: Diane.Lalancette@oecd.org

JT03317553

English - Or. English

TABLE OF CONTENTS

HIGHLIGHTS - THE AHELO FEASIBILITY STUDY:

WHAT HAVE WE LEARNED SO FAR?

What is AHELO?

1. The goal of the Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study is to assess whether it is possible to measure and compare at the international level what undergraduate degree students know and can do on graduation, in order to provide better information to higher education institutions, governments and other stakeholders, including students and employers.

2. This study entails an evaluation of the scientific feasibility of carrying out an international assessment of higher education learning outcomes (in generic and subject-specific skills) at the end of a bachelor's degree program, as well as gauging the feasibility of its practical implementation. Work is unfolding in two phases:

- Phase 1 – Initial proof of concept

  The first phase (January 2010 to June 2011) focused on providing an initial proof of concept by devising assessment frameworks and instruments that have sufficient validity in various national, linguistic, cultural and institutional contexts.

  The goal was to get a sense of cross-linguistic and cross-cultural validity by:

  – developing provisional assessment frameworks and testing instruments suitable for an international context for each of the disciplinary strands of work: economics and engineering;

  – adapting an existing instrument for the generic skills strand; and

  – validating those tools through small-scale focus group testing in participating countries.

- Phase 2 – Scientific feasibility and proof of concept

  The second phase (March 2011 to December 2012) is focussing on addressing issues of practical feasibility and further investigation of validity issues by assessing students learning outcomes.

  The goal is to evaluate the scientific and practical feasibility of an AHELO by implementing assessment instruments and context questionnaires in small groups of diverse higher education institutions to explore:

  – the best ways to implicate, involve and motivate leaders, faculty and students to take part into the testing; and

  – the relationships between context and learning outcomes, and the factors leading to enhanced outcomes.

3. With the completion of the AHELO Feasibility Study, the information collected on student performance and the analysis of the results will help to assess whether a full-fledged AHELO study could feasibly be taken forward from both scientific and practical standpoints.

4. Outcomes from the AHELO Feasibility Study will guide the decision to be made by OECD member countries as to whether to launch a full-fledged study in the longer term.

Progress of Work (as of March 2012)

Participating countries

5. Seventeen countries are participating in the AHELO Feasibility Study, with Abu Dhabi joining the Engineering strand in February 2012. Country participation is distributed across three strands of testing: nine countries have been involved in the Generic Skills strand, seven in the Economics strand, and nine in the Engineering strand. All countries participate in the Contextual Dimension, which incorporates student-, faculty- and institution-level instruments (see Annex I for the list of participating countries).

Phase 1 progress of work

6. With respect to the first phase of work, the assessment frameworks in economics and engineering have been developed and validated by expert groups and participating countries. The instruments have also been developed (economics, engineering) or amended for international use (generic skills). Countries have now also adapted and translated the assessment instruments to their own language/cultural background and carried out cognitive labs and/or focus groups to put them to the test with small numbers of students and faculties for qualitative feedback.

Phase 2 progress of work

7. Meanwhile Phase 2 activities have started with the development of the framework and instruments for the Contextual Dimension instruments (seeking background information from institution leaders, faculties and students). The field implementation of the assessment and contextual instruments has commenced. The field testing of those instruments and surveys will involve between one and two hundred students per institution, and about ten higher education institutions per country.

Interim feasibility assessment

Evaluation of the AHELO feasibility study outcomes

8. AHELO's viability hinges on proof that it is feasible to develop instruments that measure learning outcomes on an international scale, and that it is feasible to implement these assessments using methods that are scalable, efficient and secure:

- Scientific feasibility depends on whether it is possible to develop assessments that are perceived as valid in diverse institutional, cultural and linguistic contexts. The study also needs to gauge whether test items perform as expected and test results meet pre-defined psychometric standards of validity and reliability.

- Practical feasibility is being evaluated on the basis of how effective the strategies to secure institutional and student cooperation have been, and to what extent the implementation of the assessments has brought benefits to participating institutions and demonstrated its value for teaching improvement.

9. The main strategy to assess the success of the AHELO Feasibility Study will be to examine evidence from its processes and results, and use this evidence to draw conclusions as to the scientific and practical feasibility of a generalised assessment. To do so, the evaluation of feasibility is based on the AHELO Analysis Plan that outlines the operational criteria and thresholds to be used to define success for each of the technical challenges to be addressed.

10. Although results from Phase 2 field testing are yet to come, some intermediate findings and insights can already be reported from the work completed to date. The findings summarised below are briefly presented in Annex II in response to the specified Analysis Plan questions. These preliminary results are also described in more details in the Revised AHELO Interim Feasibility Report – March 2012 [EDU/IMHE/AHELO/GNE(2012)5].

Preliminary results

11. Drawing on evidence available at this time, responses to the AHELO Analysis Plan indicate that:

a. It is not yet possible to determine whether international consensus has been reached on a Generic Skills Assessment Framework as development work has only recently commenced.

b. While work is well underway it is too soon to determine whether the Generic Skills Assessment provides valid and reliable measurement given the absence of content specification, lack of technical information on constructed response tasks and inclusion part-way through the study of multiple choice questions.

c. International development and validation of the Economics and Engineering Assessment Frameworks has affirmed that it is possible to define discipline-specific learning outcomes internationally.

d. The Economics and Engineering tests, which have been endorsed by domain and national experts, suggest that it is feasible to develop valid discipline assessments to international standards, noting that further evaluation is required during implementation to confirm that the Economics test is pitched 'above content'.

e. The Contextual Dimension Framework and the student, faculty and institution instruments have been developed and validated with stakeholders for the purposes of the AHELO Feasibility Study to reflect an international consensus about the important contexts that shape higher education learning outcomes.

12. The implementation phase of AHELO is just commencing but as with any study of this nature much design and development is incorporated into the production of instruments. Feasibility of implementation insights thus far include that:

a. The AHELO Assessment Design sets out a vision and strategy for success which, to date, has been sustained and scaled. The only major revision, in response to feedback, has involved clarifying and augmenting the Generic Skills strand with multiple choice items and an assessment framework.

b. Work done so far in Economics, Engineering and Contextual Dimension and Project Management strands has worked well practically and has scaled well. National management arrangements have been effective, even despite a range of political, economic and natural problems. Strategies for engaging ministries, institutions, students and stakeholders have proven effective, with much remaining to be done in the implementation phase.

c. In terms of cross-national, cross-cultural and cross-linguistic generalisability, evidence is available to confirm that the test instruments—including constructed responses and multiple choice items—for Generic Skills, Economics and Engineering have been translated, adapted and verified independently to international standards using transparent processes. The three context

instruments have been translated for all languages, and adaptation and verification is complete in all countries.

d. The support from systems and institutions has not only been high, but has outstripped predictions. An increasing number of countries have chosen to participate in AHELO as the study has progressed. Similarly, in the majority of participating countries, National Project Managers reported high levels of interest and have not reported difficulty recruiting institutions.

Looking forward

13. Overall, as the results of the Interim Feasibility Report suggest, significant ground has been made in establishing feasibility. Major outcomes have been achieved, such as international agreement regarding discipline frameworks and instruments, and many challenges have been overcome. This is an important development.

14. Good progress has been made in little more than a year, and in developing instruments and preparing for fieldwork there is certainly no evidence to suggest that an international assessment of higher education learning outcomes is not possible. More boldly and positively, and without discounting the significant work and challenges to be confronted during implementation, there are genuine and sound indications that much of AHELO is feasible.

15. Despite the positive outcomes to date, more work needs to be done to assess the feasibility of an AHELO. Activities scheduled for Phase 2 will allow for a more comprehensive evaluation of the AHELO Feasibility Study looking more specifically whether:

- the study has been implemented in a methodologically rigorous fashion;
- it has been possible to engage faculty and a random sample of student respondents in the study;
- the instruments have been delivered successfully; operations have been managed successfully by countries and institutions;
- people could be trained to score tasks in cross-linguistically and cross-culturally generalisable ways; or
- statistical reports of assessment results are of value to systems and institutions.

16. While progressing towards the end of the feasibility study, the evaluation of the outcomes will be pursued through extensive and careful psychometric analysis of the AHELO feasibility study data, but also with technical reviews by international experts as well as the international conference in early 2013 gathering technical experts, country representatives and different stakeholder groups to discuss whether, and how, to take the results from the feasibility study forward.

ANNEX I – AHELO PARTICIPATING COUNTRIES

Table 1 – AHELO Participating Countries

Strands of Work	Participating Countries
Generic Skills	1. Colombia 2. Egypt 3. Finland 4. Korea 5. Kuwait 6. Mexico 7. Norway 8. Slovak Republic 9. USA (CT, MO and PA)
Economics	1. Belgium (Fl.) 2. Egypt 3. Italy 4. Mexico 5. Netherlands 6. Russian Federation 7. Slovak Republic
Engineering	1. Abu Dhabi 2. Australia 3. Canada (Ontario) 4. Colombia 5. Egypt 6. Japan 7. Mexico 8. Slovak Republic 9. Russia Federation

ANNEX II – SUMMARY RESPONSES TO THE ANALYSIS PLAN QUESTIONS

17. The AHELO Analysis Plan specifies nineteen questions among which eight are specific to instrumentation feasibility considerations and eleven to implementation feasibility considerations. The summary below provides a brief overview of the information collected in response to questions specified in the Analysis Plan, drawing on the evidence available at the time to assess AHELO's feasibility. More detailed information is provided in the Revised AHELO Feasibility Study Interim Report – March 2012 [EDU/IMHE/AHELO/GNE(2012)5].

Feasibility of instrumentation

- *Was the Generic Skills Assessment Framework reflective of an international consensus about the areas that are important to assess? (Question 1)*

  - The original study design for the assessment of generic skills sought to adapt an existing instrument (the United States' Collegiate Learning Assessment) and did not include the development of an international version of an assessment framework. As the work on instrumentation evolved and multiple choice questions were added the need for an international consensus about the generic skills to be measured became apparent, and has contributed to the decision to develop the Generic Skills Assessment Framework, which is not yet completed.

- *Was the instrumentation developed on the basis of the Generic Skills Assessment Framework faithful to the spirit and intent of the framework? (Question 2)*

  - Evaluation of whether an adapted version of the Collegiate Learning Assessment provides valid and reliable measurement of a Generic Skills Assessment Framework has thus far been limited by the absence of a framework and of data on which such assessment might be made. Hence it is too soon to draw an interim conclusion.

- *Was the provisional Economics Assessment Framework reflective of an international consensus about the important learning outcomes in Economics? (Question 3)*

  - The Economics Assessment Framework defines the domain to be tested and specifies the expected learning outcomes for students in the target population. The framework provides an overview of the instrumentation required to measure the competencies, with discussion of issues such as time, language level, item type, scoring, assessment delivery and administration, and reporting.

  - International development and validation of the Economics Assessment Framework suggests it is possible to define discipline-specific learning outcomes internationally.

- *Was the instrumentation developed on the basis of the Economics Assessment Framework faithful to the spirit and intent of the framework? (Question 4)*

- Initial validation of the AHELO Economics Assessment shows it has the potential to operationalise the Economics Assessment Framework well and provide valid, reliable and efficient measurement of target constructs.

- The endorsement of the assessment in Economics by domain experts and national managers in a number of countries suggests that it is possible to develop assessments to international standards in this domain.

- *Was the provisional Engineering Assessment Framework reflective of an international consensus about the important learning outcomes in Engineering? (Question 5)*

  - International development and validation of the Engineering Assessment Framework suggests it is possible to define discipline-specific learning outcomes internationally.

  - A useful index of feasibility is how the Engineering Assessment is received by relevant academic and professional communities. It is important that consensus was reached by the development team which comprised experts from Australia, Japan and several European countries. Feedback from broader stakeholders consulted throughout the development has also been positive including Engineering societies and associations of professional engineers.

- *Was the instrumentation developed on the basis of the Engineering Assessment Framework faithful to the spirit and intent of the framework? (Question 6)*

  - Initial validation of the AHELO Engineering Assessment shows it has the potential to operationalise the Engineering Assessment Framework well and provide valid, reliable and efficient measurement of target constructs.

  - The endorsement of the assessment in Engineering by domain experts and national managers in a number of countries suggests that it is possible to develop assessments to international standards in this domain.

- *Was the provisional Contextual Dimension Assessment Framework reflective of an international consensus about the important contexts that shape higher education learning outcomes? (Question 7)*

  - The Contextual Dimension Assessment Framework is informed by the processes and practices adopted in contextual dimensions used around the world. Development of the framework was undertaken through research and consultation, and by seeking the expert opinion of a range of groups and individuals from across the world.

  - The framework has been developed and validated which, for the purposes of the AHELO Feasibility Study, reflect an international consensus about the important contexts that shape higher education learning outcomes.

- *Was the instrumentation developed on the basis of the Contextual Dimension Assessment Framework faithful to the spirit and intent of the framework? (Question 8)*

  - Three survey instruments have been developed to underpin the Framework: a Student Context Instrument (SCI), a Faculty Context Instrument (FCI) and an Institution Context

Instrument (ICI). In addition, a range of indicators have been specified for collection at the national/system level to provide additional context.

- Widespread consultation based on the AHELO Contextual Dimension instrumentation suggests that the Contextual Dimension Student, Faculty and Institution Context Instruments have the potential to operationalise the Contextual Dimension Assessment Framework well and provide valid, reliable and efficient measurement of target constructs.

Feasibility of implementation

- *Was the AHELO Assessment Design valid and feasible? (Question 9)*

  - The AHELO Assessment Design was written in 2009 to advance the AHELO Consortium's integrated approach for designing and implementing the AHELO Feasibility Study. It has received minor updates as new countries have joined the study. The only major revision has involved clarifying and augmenting the Generic Skills strand with multiple choice items and an assessment framework. Assessing the validity and feasibility of the Assessment Design requires examination of a number of factors such as its scope, the extent to which it underwent changes, and the suitability of its design to provide a basis for full-scale extension. This is not fully possible until implementation is completed in late 2012.

- *Was the study feasible from a practical perspective? (Question 11)*

  - AHELO will only be successful if it is practical. Practical feasibility encompasses a range of measures in relation to cost, timing, and communication processes, the use of technology, and acceptance by institutions, faculty and students. All of these must be scalable if AHELO is to expand to a greater number of disciplines and countries. At this stage it is possible to make preliminary remarks about operations associated with instrumentation and pre-implementation.

  - The AHELO Consortium has been able to provide all deliverables to the OECD within agreed budgets and timeframes despite significant budget reductions, the inclusion of new countries during the study, and the compression of development times.

  - National operations are another crucial facet of practicality. Countries have made significant investments in setting up national infrastructure for AHELO. Overall, up to this point, National Project Managers and Institution Coordinators have been effective and efficient and have made their best efforts to adhere to timelines.

  - A large number of diverse stakeholders are involved in the AHELO Feasibility Study. To achieve the smooth flow of work, a clear, consistent and transparent approach to communication has been essential. This outcome has been achieved through effective use of information technology to communicate and exchange materials.

  - It is not yet possible to report on feedback from institutions about the AHELO Feasibility Study as systematic input will not be sought until 2012.

- *Was the study successfully generalised cross-nationally, cross-culturally, cross-linguistically and cross-institutionally? (Question 12)*

- Although AHELO is not designed like most international education assessments to yield national estimates, it remains critically important to determine the extent to which nations have engaged with the study. Such engagement is difficult to appraise, but much can be read from the scope and scale of ministerial involvement. The scope of country participation in the AHELO Feasibility Study has increased since its inception.

- As institutions are both the main units of analysis and the reporting level in AHELO, it is essential that the study be generalisable in this way. It is imperative that AHELO's assessments reflect and encourage rather than compress or even suppress institutional diversity. At this stage of the AHELO Feasibility Study it is not possible to fully determine whether assessment resources or processes are able to be generalised across institutions.

- Establishing cross-cultural and cross-linguistic generalisability was sought by translation, adaptation and verification processes. Two different approaches were used: one for the Generic Skills constructed response tasks; and another for the Generic Skills multiple choice questions, Contextual Dimension instruments, and Economics and Engineering tests. From a general feasibility perspective based on quality, delivery, transparency and cost there is much to recommend the approach used in the Economics and Engineering assessments. While it required time for some countries to establish national teams, national processes were effective once these were trained.

- *Was it possible to engage systems and institutions in the study? (Question 13)*

  - Engaging national higher education systems and institutions in the AHELO Feasibility Study is vital to its success. The support from systems and institutions has not only been high, but has outstripped predictions on many measures. The enthusiasm and engagement of higher education systems in countries around the world, and from institutions within those systems, indicates the great interest with which the AHELO Feasibility Study is viewed and the broad acceptance of the significance it has for higher education internationally.

  - Similarly, National Project Managers in all countries taking part in the Economics and Engineering strands reported high levels of interest among institutions and did not have difficulty in recruitment. In some countries, the level of interest from institutions was such that more than the required number of institutions was recruited.

18. At this stage it is however too early to evaluate feasibility in ways that enable response to the following questions specified in the AHELO Analysis Plan:

- *Was the study implemented in a methodologically rigorous fashion? (Question 10)*

- *Was it possible to engage faculty respondents in the study? (Question 14)*

- *Was it possible to engage a random sample of student respondents in the study? (Question 15)*

- *Was the instrument delivered successfully? (Question 16)*

- *Were survey operations managed successfully by countries and institutions? (Question 17)*

- *Was it possible to train people in different countries to score tasks in cross-linguistically and cross-culturally generalisable ways? (Question 18)*

- *Were statistical reports of assessment results of value to systems and institutions? (Question 19)*