



ROADMAP FOR THE OECD ASSESSMENT OF HIGHER EDUCATION LEARNING OUTCOMES (AHELO) FEASIBILITY STUDY

2nd Version - 11 June 2009

Background context	2
AHELO goals and challenges	2
Strategic goals	2
Challenges to address in AHELO	3
Why the OECD?	3
The AHELO approach and its rationale	4
A multi-dimensional definition of quality	4
HEIs as units of analysis	5
A target population of students at the end of their undergraduate degree	6
Reporting at institutional level to permit a wide substantive coverage with limited response burden	6
The AHELO feasibility study	6
Expected outcomes	6
Several strands of work reflecting several aspects of quality	7
Maximising insight from a broad geographic and linguistic coverage in a manageable way	9
Countries involved	9
A transparent and collaborative process	10
A cooperative funding mechanism	11
Proposed activities in the different strands of work	12
Four strands of work	12
Main activities for field-testing of assessment and survey instruments	15

ROADMAP FOR THE OECD ASSESSMENT OF HIGHER EDUCATION LEARNING OUTCOMES (AHELO) FEASIBILITY STUDY

2nd Version - 15 June 2009

The purpose of this paper is to describe the roadmap for work to be undertaken as part of the Feasibility Study for an OECD Assessment of Higher Education Learning Outcomes (AHELO). Detailed planning for the AHELO feasibility study started in May 2008 and a first version of the roadmap was published in July 2008. This revised version of the roadmap reflects current thinking and will be expanded and updated over time as the feasibility study unfolds.

Background context

High quality higher education is essential for success in the knowledge economy. Investment in higher education is significant and growing, both from public and private sources. Following decades of rapid expansion of higher education and its growing internationalisation, there is increasing recognition that greater attention should now be paid to its quality and relevance to ensure quality provision for all. Policymakers as well as the public devote considerable attention to the outcomes of higher education given its importance for human capital development, its cost to public finances as well as to students and their families, and the needs of business and industry.

At the same time, efforts to improve the quality of teaching and enhance the learning outcomes of students enrolled in higher education suffer from a considerable information gap. There is no reliable information which enables comparative judgments to be made about the competences of students in different countries and different institutions, or about the quality of teaching. The reputations of higher education institutions (HEIs) are based largely on historical research performance. It seems that international rankings derived from inputs or research-driven outputs are distorting decision-making by individuals, institutions and governments. Developing measures that give due weight to teaching practices and learning outcomes has thus become essential.

In response, the OECD has embarked on a feasibility study to explore the viability of developing an international Assessment of Higher Education Learning Outcomes (AHELO) that would measure learning outcomes in ways that are valid across cultures and languages – but also across the diversity of institutional settings and missions.

AHELO goals and challenges

The purpose of the AHELO feasibility study is to assess whether it is possible to measure what undergraduate degree students know and can do at the international level, in order to provide relevant information to HEIs, governments, and other stakeholders including students and employers.

Strategic goals

The underlying motivation for an AHELO is that this information could contribute to HEIs' knowledge of their teaching performance, and thereby provide a tool for development and improvement. As such, the AHELO central emphasis is on the improvement of teaching and learning and in providing higher education leaders with tools to empower them and foster positive change and enhanced learning.

The feasibility study has two key aims. First, to test the science of the assessment – whether it is possible to devise an assessment of the outcomes of higher education which enables to make valid and reliable statements about the performance/effectiveness of learning in HEIs of very different types, and in countries with different cultures and languages. Second, to test the practicality of implementation – whether it is possible to motivate HEIs and students to participate. In addition, the feasibility study will involve related work exploring other options for capturing indicators of higher education quality indirectly.

Challenges to address in AHELO

The development of an AHELO presents a number of scientific and practical challenges. It is of crucial importance that an assessment has both reliability and validity. Constructing an assessment that is valid across HEIs, cultures and disciplines implies taking into account:

- The diversity of HEIs;
- Differences between national systems of higher education;
- The absence or presence of selection in the system or certain HEIs;
- Variations in the duration and content of programmes; and
- Cultural and linguistic diversity.

Practical and operational challenges that also have to be addressed are how to motivate students and HEIs to participate and ensuring a fair assessment of HEIs and programmes.

Why the OECD?

The OECD is well placed to lead and implement the feasibility study given its institutional framework for cooperative and international comparative work, as well as its credibility and demonstrated expertise in developing large-scale international assessments such as the Programme for International Student Assessment (PISA) at school level and the Programme for International Assessment of Adult Competencies (PIAAC) which focuses on the skills of the adult workforce.

Within the OECD, the Programme on Institutional Management in Higher Education (IMHE) provides a platform for the engagement of HEIs with Governments in AHELO in a way which will ensure that the approaches adopted take account of institutional needs and concerns. IMHE is open to all recognised HEIs in OECD countries, as well as associations of HEIs and representatives of Governments. It can therefore provide a locus for consultation and steering of the feasibility study which also allows countries which are not directly involved in it to have their say and monitor progress.

The feasibility study is managed by a relatively small team of at the OECD. The bulk of the technical work involved in the different strands of the AHELO feasibility study work will be carried out through collaboration with a number of contracted outside experts, consultants and organisations given the scope of the task and the high level of specialised expertise required. A number of countries and organisations have already expressed interest in participating and/or supporting this feasibility study work in-kind or through voluntary contributions. These include HEIs' associations, quality assurance agencies, foundations as well as countries interested in taking part in the feasibility study (see Table 1 for a list of participating countries, as of 20 May, 2009). These organisations will be invited to support the various strands of the feasibility study while countries taking part in the different strands will contribute through voluntary contributions as well as national implementation costs.

The AHELO approach and its rationale

A multi-dimensional definition of quality

Measures of higher education learning outcomes hold out important promises for students, HEIs and public policy more generally. The extent to which the information needs of the various stakeholders can be met will depend on the units from which data are collected and for which outcomes are reported. The approach underlying the AHELO feasibility study considers these units in the form of a three-dimensional matrix.

The first dimension of the matrix comprises users, *i.e.* the stakeholders who would be served and their information needs. These include:

- *Individuals*, such as students wishing to make better-informed choices or employers seeking to benchmark qualifications against actual skill measures;
- *HEIs, departments or faculties* seeking a better understanding of their comparative strengths and weaknesses; and
- *Public policymakers* seeking to quantify stocks and flows in high-level skills to obtain better insights into the quality, equity and efficiency of higher education services and to assess the impact of policy decisions.

The second dimension of the matrix relates to the use of the measures on learning outcomes, ranging from summative comparisons of institutional performance to informative and diagnostic tools at programme and faculty level.

The third dimension of the matrix considers the kind of instruments needed to serve the respective users and uses.

Recognising the multiplicity of potential uses and users, the OECD does not aim for the establishment of a single performance measure that would then be used for a unidimensional ranking of HEIs or countries. The OECD also acknowledges that any effort to bring together all HEIs on one standard, would risk driving the assessment down to the lowest common denominator. The AHELO aim is rather to establish a “multi-dimensional quality space”, in which quantifiable criteria for quality establish the dimensions of the space. Within this concept of the “quality space” higher education systems, HEIs, departments and faculties can then be situated depending on the prevalence of the different quality attributes. Students would then be able to choose programmes and HEIs depending on the configuration of the quality attributes that are most relevant to them, rather than depend on ratings that combine quality attributes in predefined ways, which may not necessarily be the most relevant ones for either students or providers. It would also become possible to portray policy trajectories of HEIs and systems over time, as they change their position on the different dimensions of this “quality space”, which in turn could become a powerful tool for public policy.

For instance, two kinds of outcomes measures can be considered at the level of HEIs or departments, both of which the OECD is committed to explore to meet the information needs of the various groups of stakeholders:

- *Individuals*, whether prospective students or employers, would want to know the “bottom line” of the performance of HEIs, departments or faculties, in terms of the absolute student performance or raw scores, recognising that such an assessment would not only measure the quality of

educational services provided, but also other aspects such as the effects of selection and the socio-economic makeup of students.

- Meanwhile, individuals, HEIs and public policymakers wishing to assess the quality of the teaching services provided by HEIs would primarily be interested in the “value-added” or “gain” provided by the HEIs, *i.e.* the scores an institution would attain after accounting for the quality of prior schooling or the degree of selectivity of the programmes and HEIs.

The OECD also recognises that learning outcomes are only one component of the quality of HEIs, and that the “quality space” concept would need to recognise other dimensions of the quality of outcomes as well, such that HEIs and systems could be appropriately represented, in accordance with their respective missions. Other dimensions for which various instruments have been established include curriculum and consumer ratings. Alumni ratings, such as used by the German Centre for Higher Education (CHE), are considered one way to complement the assessment of competencies linked to “employability” with actual labour-market outcomes. Institutional factors, non-cognitive characteristics that are known to be tied to successful study and achievement, and measures of institutional efficacy are also seen as relevant. Last but not least, there is a wide area of direct and indirect measures of research outcomes that can be utilised to complement the work undertaken by the OECD.

HEIs as units of analysis

For several reasons, the OECD does not consider it feasible to develop internationally comparative information on higher education learning outcomes at the country/system level:

- First of all, any system-level assessment would have to confront issues of differentiation within the system, as variation in institutional structures challenges the establishment of cross-nationally comparable classes of HEIs, and rigid institutional boundaries would risk undermining the richness of horizontal differentiation.
- Second, in many countries, governments have limited options to incentivise the participation of HEIs in assessments, particularly those carried out at an international level.
- Third, even if it were possible, large differences in enrolment rates would raise questions about how to interpret results since, other things being equal, it would only be natural that those countries with more selective systems and lower enrolment rates would likely perform better than countries where higher education has become largely universal.
- Last, but not least, mandated assessment in the form that would be required to obtain a system-wide representative sample of HEIs, may lead to administrative response, not faculty response and would thus not be effective as a tool for improvement at the level of service provision.

The AHELO approach therefore centres on the establishment of measures of learning outcomes at the level of HEIs, departments or faculties, the idea being to combine the definition of OECD measures of quality with valid and reliable assessment methods to which HEIs could, with an appropriate set of incentives, voluntarily subscribe and which could progressively find acceptance in a widening range of HEIs. Eventually, if an OECD assessment gathered pace and found wide acceptance, issues of assessing system-level performance might be addressed, and this would make the work relevant to a much wider range of stakeholders in the longer term.

A target population of students at the end of their undergraduate degree

The OECD has devoted consideration to what would constitute meaningful target populations for an AHELO. One possibility, that would facilitate comparisons across HEIs and countries, would be to focus the assessment on comparable age bands. However, this would make it very difficult to link results to national degree and qualification structures and thus make it difficult to interpret them in the national institutional context. Such an approach would also be very difficult to implement, as it would require the selection of age-based samples that may spread widely across years of study. The approach therefore considers an assessment towards the end of the first undergraduate degree (of a three- or four-year duration depending on the programmes offered in participating countries) as a more practical solution for the feasibility study.

Reporting at institutional level to permit a wide substantive coverage with limited response burden

The OECD recognises diverging views on how the knowledge about learning outcomes in HEIs can and should be used. Some see such information primarily as a tool to reveal best practices, to identify shared problems among HEIs and to encourage collaboration and lateral capacity building among research and teaching personnel. With this approach, emphasis would be placed on the relevance of performance information for the institutions themselves and on contextualising performance data with other information on the learning environment in HEIs. Other views extend the purpose of learning outcome measures to support contestability of public services or market-mechanisms in the allocation of resources, *e.g.* by making comparative results of HEIs publicly available to facilitate choice.

Irrespective of where the focus would eventually be placed in different national contexts, an AHELO would not seek to report data at the level of individual students. This approach has important technical implications since it will allow the use of matrix sampling techniques¹ through which the coverage of the assessment areas could be widened for an eventual main study, without unduly extending the response time demands on individual students.

The AHELO feasibility study

Expected outcomes

There is currently considerable interest within institutional, political and scientific circles for measures of higher education learning outcomes, but uncertainties and doubts of some actors as to whether it is scientifically and operationally feasible to measure learning outcomes across HEIs of very different types, and in countries with different cultures and languages. In order to answer this question, a number of international experts have been consulted over the past two years. Three meetings were held in 2007, bringing together international specialists in the field². The main conclusion of the experts was that while it

¹ This approach allows significant gains in testing time and is intended to overcome the conflicting demands of limited student testing time and broad coverage of assessment domains. The difficulty of test items is linked to the abilities of respondents (via a Rasch model) by a logistic function which makes it possible to compute the probability that a student with a given level of ability succeeds on an item. The Rasch model thus creates a scale on which every item and every respondent can be located. Due to this probabilistic link, it is no longer necessary to administer the whole battery of assessment items to every respondent, and it is possible to assign a subset of the items pool to certain sub-samples of students provided that some link items are administered to several sub-samples of students. In this way, the performance levels attained by respondents on their subset of items provide a reliable estimate of their performance had they responded to the whole battery of items.

² The expert meetings took place in Washington (28 April), Paris (5-6 July) and Seoul (26-27 October). The Washington meeting was primarily about the usefulness and desirability of an OECD international

might be both desirable in terms of public policy and theoretically possible to assess and compare central components of education outcomes, it would be necessary to conduct a feasibility study to test this proposition before undertaking any more systematic assessment. The feasibility study would have to do two things:

- Test the science of the assessment – whether it is possible to devise an assessment as well as associated contextual data which enables reliable statements to be made about the performance of students in institutions of very different types, and in countries with different cultures and languages; and
- Test the practicality of implementation – whether it is possible to motivate institutions and students to take part in such an assessment and find solutions to implement such an assessment.

The feasibility study is expected to demonstrate the feasibility – or otherwise – of comparing HEIs' performance from the perspective of student learning rather than relying upon research-based measures which are currently being used across the globe as overall proxies of institutional quality. To this aim, it will implement several instruments in different countries. It is anticipated that the assessment instruments and contextual surveys will be administered to students who are almost at the end of a bachelor-type degree, with the option of paper-and-pencil or electronic delivery, depending on operational considerations (logistics implications, security of testing, access to computers in some sub-systems for some participating countries). Because of the national differences in academic year, a window of testing time needs to be allowed, giving flexibility to the HEIs to ensure that the timing does not coincide with major exams.

The main criteria to assess the success of the feasibility study is to provide a proof of concept that the various instruments considered can be applied in diverse institutional, cultural and linguistic settings with appropriate adaptations and yet provide valid, reliable and free-of-bias measures of student learning outcomes as well as indirect measures of higher education quality. With respect to implementation, success would be defined in terms of meeting minimum student participation and response rates thresholds.

The design of the feasibility study will involve careful psychometric analysis, technical reviews by international experts as well as an international conference gathering different stakeholder groups to discuss whether, and how, to take the results from the feasibility study forward. In case of positive results, this proof of concept would constitute a key pillar for longer-term work as it would assist OECD countries in deciding whether to launch a fully-fledged AHELO study.

Several strands of work reflecting several aspects of quality

Any assessment of higher education learning outcomes will need to define and operationalise criteria for what constitutes quality. The OECD acknowledges that there is no generally accepted definition of what higher education outcomes ought to be, but considers that there are promising ways to examine various facets of learning outcomes underway. The feasibility study will thus involve different kinds of activities to explore the feasibility of directly measuring or indirectly capturing various aspects of quality.

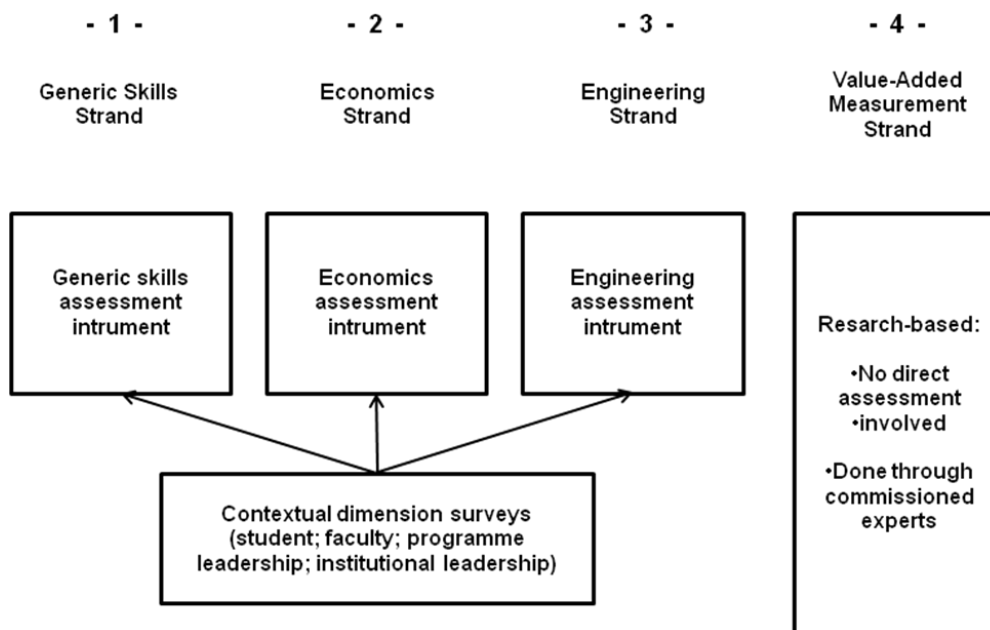
For the sake of a feasibility study, it is not necessary to develop comprehensive instruments to assess student performance. Instead, the focus is on providing proof of concept, and it should be possible to take advantage of this feasibility study to explore different approaches, methodologies and instruments that

assessment of higher education learning outcomes, the Paris meeting focused on the conceptual possibility, and the Seoul meeting concerned how to move from possibility to feasibility. See www.oecd.org/edu/ahelo for the summary records and lists of participants from these meetings.

might eventually be envisaged as parts of a fully-fledged assessment. As a result, the work is divided in four distinct strands of work (see Figure 1 below) to be undertaken separately but coherently:

1. A Generic Skills Strand to assess the feasibility of measuring generic skills across diverse institutional, national, linguistic and cultural contexts;
2. An Economics Strand to assess the feasibility of measuring discipline-specific learning outcomes in this disciplinary field across diverse institutional, national, linguistic and cultural contexts;
3. An Engineering Strand to assess the feasibility of measuring discipline-specific learning outcomes in this disciplinary field across diverse institutional, national, linguistic and cultural contexts;
4. A Value-Added Measurement Strand to reflect – from a research perspective and with no test of instruments – on possible methodologies to capture learning gain during a student higher education experience, net of incoming abilities.

Figure 1 – AHELO feasibility study strands of work



Three small-scale tests of instruments will be undertaken, focusing on generic skills and discipline-specific assessments in two contrasted fields, with the understanding that an AHELO would aim at expanding the number of disciplines covered over time. Economics and engineering have been chosen for the feasibility study to get insight into the feasibility of measuring discipline-specific skills in both scientific and social sciences domains. The focus will be on competencies that are fundamental to the discipline and “above content”. Indeed, the AHELO approach follows the PISA dynamic model of lifelong learning in which new knowledge and skills necessary for successful adaptation to a changing world are continuously acquired throughout life. AHELO therefore focuses on aspects that higher education students will need in the future and seeks to assess what they can do with what they have learned. The assessments are not constrained by the common denominator of programme curricula which are very diverse in higher

education, but instead examine students' ability to reflect, and to apply their knowledge and experience to novel and real world tasks and challenges.

The purpose of these small-scale tests of instruments will be to assess the international validity and reliability of the instruments used, taking account of institutional and individual characteristics. To accomplish this, each instrument will involve a small number of countries sufficiently different in terms of their higher education system organisation, geographic setting, and linguistic and cultural background to provide proof of concept of the cross-national validity of the instrument, and about 10 different types of HEIs within each country to provide proof of concept of its cross-institutional validity.

A Contextual Dimension will be embedded in each of the three small-scale tests of instruments to assess the feasibility of capturing contextual information on institutional settings, teaching practices and environment characteristics that may affect learning, as well as indirect proxies of quality. This Contextual Dimension will be collected to add essential analytic dimensions to AHELO.

But summative measures of learning outcomes are only one facet of quality, and the OECD is also committed to explore outcomes measures in terms of value-added to meet the information needs of the various groups of stakeholders. The fourth strand of work will thus explore the issue of value-added measurement in higher education – *i.e.* the learning gain that takes place during the higher education experience. However, it has to be recognised that assessing the value-added or marginal gain by HEIs imposes layers of complexity that, though theoretically well understood, are difficult to resolve in large scale assessments. Given the complexity of measuring marginal gain, the proposed approach will first scrutinize possible methods for capturing marginal learning outcomes that can be attributed to attendance at different HEI's, both from a conceptual/theoretical perspective and in terms of psychometric approaches. It will build upon similar work carried out at school level by the OECD and review options for value-added measurement in higher education. Researchers will be invited to study potential data sources, methodologies and psychometric evidence on the basis of datasets existing at the national level, with a view to providing guidance towards the development of a value-added measurement approach for a potential fully-fledged AHELO study.

Maximising insight from a broad geographic and linguistic coverage in a manageable way

In order to maximise information gains while recognising the need to ensure a broad geographic and linguistic coverage as well as various practical and resource constraints the different strands of work will be carried out separately but coherently. Each strand will involve 4 or 5 countries as diverse as possible, with 10 HEIs from each country reflecting the diversity of higher education at national level. This number is large enough to assess the measurement properties of the various instruments, and small enough to keep the process manageable and avoid that validity gains are sacrificed over efficiency gains.

Countries involved

Countries supporting or involved in the AHELO feasibility study include – as of 11 June 2009 – Australia, Belgium (Flemish Community), Finland, Italy, Japan, Korea, Mexico, the Netherlands, Norway, Sweden and the United Kingdom. A number of other countries are still considering participation.

Countries directly involved in the implementation of the AHELO assessments in the different strands of work are the following:

- *Generic Skills Strand*: Finland, Korea, Mexico and Norway;
- *Economics Strand*: Belgium (Flemish Community), Italy, Mexico and the Netherlands; and

- *Engineering Strand*: Australia, Japan and Sweden.

A transparent and collaborative process

The support of both governments and HEIs is necessary for the success of this initiative. It is thus essential for the AHELO instruments to be perceived and used as a tool for improvement and that the work be led and steered jointly by the higher education sector and governments, making the process transparent to all stakeholders – irrespective of whether they participate in the feasibility study or not.

A unique governance structure within the OECD

The AHELO feasibility study benefits from a unique governance structure within the OECD, whereby the project is jointly steered by governments, HEIs and agencies through the Programme for Institutional Management in Higher Education (IMHE) Governing Board – which brings together these different groups with a common interest in improving institutional management and effectiveness. This peculiar governance arrangement reflects the stakes and ambition of the study, and the deliberate inclusiveness of the process for its development. Through the IMHE Governing Board, governments and institutions determine, in the context of OECD objectives, AHELO's policy priorities and oversee adherence to these priorities during implementation. The Education Policy Committee – which steers and provides strategic direction for all OECD work on education – bears responsibility to make decisions about whether and how to proceed with the work once the feasibility study is completed.

However, the technical nature of the project has led the IMHE Governing Board and the Education Policy Committee to delegate decisions on the methods, timing and principles of the AHELO feasibility study to an AHELO Group of National Experts (GNE). The AHELO GNE is expected to meet about twice a year over the life of the project. The OECD provides Secretariat services to the AHELO GNE.

A number of ad-hoc groups play a more indirect role in the steering of the AHELO feasibility study through the sharing of expertise, dialogue or the provision of advice. A range of experts have been or are to be commissioned to provide input during the initial development of the feasibility study features. For instance, a group of experts has been contracted to provide advice about the contents and construction of the contextual dimension surveys and develop conceptual and analytical frameworks for the contextual dimension surveys. Likewise, the Tuning Association has been contracted to convene academics from a range of different countries in order to reflect on definitions of expected/intended learning outcomes in economics and engineering. The outcomes of this Tuning-AHELO project provide an intermediate output of the AHELO feasibility study (Tuning Association, 2009a, 2009b). A Technical Advisory Group will also be established to advise on major technical issues (*e.g.* instrument development procedures, sample and test design, translation procedures, or scoring and verification procedures). Last but not least, a Stakeholders' Consultative Group has been established to discuss and reflect on the unfolding of the AHELO feasibility study and its potential longer-term impact.

The IMHE Governing Board will, in accordance with the principles governing OECD work on education, report on the conduct and outcomes of the feasibility study to the Education Policy Committee which will bear responsibility to take decisions about whether and how to proceed with the work once the feasibility study is completed.

Transparency

Given the high level of interest for higher education quality assurance issues, international rankings and the AHELO initiative in relation to these, a communication plan is essential in ensuring transparency on the goals and progress of the feasibility study. The OECD is already engaging with a wide range of

stakeholders in communicating the purpose and practicalities of the initiative (officials, HEIs, media, students). To this end, an informal Stakeholders' Consultative Group (SCG) gathering the diversity of stakeholders with an interest in higher education quality was created in order to:

- Streamline communication on the goals and progress of the AHELO feasibility study;
- Listen to stakeholders' suggestions, concerns, warnings or advices on the AHELO feasibility study goals and progress;
- Provide a forum for multilateral discussion and cross-fertilization of ideas on this important initiative; and
- Rely on influential networks to disseminate the aims, methods and progress of the AHELO feasibility study.

In addition, the OECD Secretariat is committed to provide regular reports detailing progress on the AHELO feasibility study, through a Website dedicated to the work [www.oecd.org/edu/ahelo], and the participation of the OECD Secretariat in major conferences to reach relevant audiences. Dissemination will also take place as part of the IMHE conferences.

A final conference will also be organised to disseminate the outcomes of the feasibility study and to inform subsequent decision-making. The targeted audience will bring together international experts, quality assurance experts, representatives of HEIs, students, businesses and policymakers. International experts will present the outcomes of the various strands of the feasibility study, and will report on the technical feasibility of developing a fully-fledged AHELO study. Workshops and panels will also be organised to provide participants with a chance to interact and discuss operational issues, political economy constraints and the potential impact of such an assessment.

A cooperative funding mechanism

The cost of the feasibility study will be substantial. OECD staff involved in the management and conduct of the study is funded from available resources, but the bulk of the money for development work and implementation needs to be raised. Indeed, OECD Ministers are keenly interested in measures of learning outcomes, but given uncertainties as to the scientific and operational feasibility of such an assessment, the OECD needs to demonstrate that it is feasible to measure learning outcomes in valid and reliable ways across cultures, countries, languages and institutions before significant resources can be invested in the development of an AHELO as part of the regular OECD programme of work and budget.

Therefore, the initial steps of the AHELO feasibility study are to be funded from a range of grants and voluntary contributions, as is commonly the case for new cutting-edge OECD projects such as PISA or PIAAC. The OECD is now seeking funds for the development phase of this project. It is envisaged to seek support from foundations for development work and voluntary contributions from countries involved in the feasibility study for national implementation costs.

Proposed activities in the different strands of work

Four strands of work

Generic skills strand: international pilot test of the Collegiate Learning Assessment

The generic skills strand is an essential component of the feasibility study. Indeed, competencies in critical thinking, analytic reasoning, problem-solving, or the generation of knowledge and the interaction between substantive and methodological expertise are widely viewed as critical for the success of individuals and of rising relevance in the information age. It is therefore important for an AHELO to measure these transversal higher-order competencies that are necessary for success in both academic and business contexts – not only cognitive knowledge. A key advantage is that such competencies are largely invariant across occupational and cultural contexts and could be applied across HEIs, departments and faculties. Moreover, a focus on higher-order skills allows the coverage of a more diverse population representing the whole undergraduate student body whereas the discipline strands will only cover a subset of students enrolled in given disciplines.

International experts gathered in 2007 and reviewed the various initiatives taken in countries to assess higher education learning outcomes (Nusche, 2007). They were impressed with the Collegiate Learning Assessment (CLA) approach taken by the Council for Aid to Education (CAE) in the United States. The CLA is a national effort to assess the quality of undergraduate education by directly measuring student learning outcomes. It measures student ability and learning in critical thinking, writing, and synthesising quantitative and qualitative data. The measures are focused on skill sets that students will need as they graduate and enter the work force, namely critical thinking, analytical reasoning, problem-solving and written communication. These skills are intertwined. Thus the CLA measures are holistic: they require students to use these skills together to respond to tasks. All CLA measures are administered online, using open-ended prompts that require constructed responses. Each task also has an accompanying library of information which students are instructed to use in preparing their answers. Tasks often require students to marshal evidence from these diverse quantitative and qualitative sources and exercise judgment on their relevance. Tasks are appropriate for students across a wide range of undergraduate academic majors and general education programmes.

As a result, the experts recommended, as part of an AHELO feasibility study, to implement field-testing of this instrument to assess the extent to which higher order skills of the type measured by the CLA can be validly measured across different cultural, linguistic and institutional contexts. In addition, the CAE will select and use a multiple-choice critical thinking test that will be administered in parallel to the CLA performance tasks. Data collected will be used to provide evidence of the construct validity of the CLA performance tasks. As with PISA, these direct assessments of student knowledge and ability will be complemented by contextual information which will then enable policy- and practice-related conclusions to be drawn at institutional level.

Discipline strands: Economics and Engineering

The assessment of student performance in the discipline strands is intended to complement the evidence of the information gathered by the generic skills strand. Indeed, the main limitation of an approach entirely limited to generic competencies, is that it would not assess the kind of subject-matter competencies that many higher education departments or faculties would consider their primary objective. There would thus be a risk that what is measured becomes too far removed from what goes on in faculties and departments and does not capture the competencies that are uniquely the province of HEIs.

The discipline strands will seek to assess discipline-related competencies. The challenge of this approach is that it requires highly differentiated assessment instruments, which would make comparisons across HEIs and countries difficult. It also excludes competency areas that are not easily amenable to large-scale assessment or that are not sufficiently invariant across cultures. However, the experts who convened in 2007 considered that interesting examples do already exist in selected disciplines, such as science, technology, engineering or health care/professions allied to medicine (STEM), on which a feasibility study could be based. The feasibility study will initially focus on one of these areas – engineering has been chosen for this purpose – and then progressively expand the range of subject areas covered over time should an AHELO main study be launched afterwards. The advantage is that this approach covers disciplines that are common among HEIs in OECD countries, less likely to be influenced by unique cultural features, and reflect the dynamics of disciplinary change. One caveat to this approach that the experts noted is that many countries have already established assessments in the STEM disciplines and that there would be greater benefits of international collaboration in areas that are yet unexplored and where national work is less likely, such as the humanities and the social sciences. Therefore, an assessment in the economics discipline is also included in the AHELO feasibility study.

These economics and engineering assessments will permit gauging the viability of measuring discipline-specific skills, representing both scientific and social sciences domains, with the understanding that a fully-fledged AHELO main study would aim at expanding the number of disciplines covered over time.

The aim shall be to assess competencies that are fundamental and “above content”, *i.e.* with the focus on the capacity of students to extrapolate from what they have learned and apply their competencies in novel contexts unfamiliar to them, an approach that is similar to PISA.

As for the generic skills strand, the discipline strands will explore the feasibility of directly measuring learning outcomes in the selected disciplines and across different cultural and linguistic contexts. A prerequisite for this is to reach international agreement on expected learning outcomes in these contrasted disciplines to provide proof of concept that it is possible to develop domain assessment frameworks in the disciplines in the context of great curriculum diversity in higher education programmes. This will be done following the Tuning approach that has been successfully applied in Europe in many disciplinary fields, and is now being piloted in other parts of the world. The outcomes of this Tuning-AHELO project (Tuning Association, 2009a, 2009b) provide an intermediate output of the AHELO feasibility study, to demonstrate that agreements on expected learning outcomes can be achieved in contrasted disciplines.

However, to the extent that the chief outcomes of the feasibility study tests are to provide a proof of concept that an assessment of learning outcomes can be done across countries, the full coverage of internationally-agreed domain assessment frameworks is not an absolutely necessary condition at the feasibility study stage (although it would be for a main study). For the sake of practicality, the feasibility study may therefore use assessment instruments developed in a specific national context, provided that the items being tested across countries gather international agreement as being relevant in each of the national contexts envisaged. In the event that an AHELO main study would be launched in the future, further work on defining assessment domains and developing instruments in a cross-national collaborative fashion would be needed.

Contextual dimension

While the main focus of the AHELO feasibility study is to gauge the feasibility of assessing learning outcomes, it is also necessary to assess the feasibility of gathering the contextual variables that will eventually be needed to interpret performance measures and help institutions understand the performance of their students and improve their teaching accordingly. Furthermore, the contextual data collected

through student, faculty and programme and institutional leadership teams will be used to rehearse some psychometric analyses to identify relevant contextual variables for longer-term development and demonstrate the analytical potential of AHELO for institutional improvement.

This aspect of the feasibility study will also require ensuring that the contextual dimension survey instruments are internationally valid and reflect the cultural context of the countries in which the AHELO feasibility study will be implemented. Some initial development work has already been undertaken. A group of commissioned contextual dimension experts (Ewell *et al.*, 2008, 2009), has developed a conceptual and analytical framework which served as a basis to prioritize desired contextual variables for the feasibility study.

The survey instruments that will capture contextual variables will consist of four instruments: two brief student and faculty questionnaires of a maximum duration of 15 minutes each, a department leadership questionnaire and an institutional leadership questionnaire.

Value-added measurement strand

As indicated above, two types of learning outcomes measures can be considered at the level of HEIs or departments:

- Individuals, whether prospective students or employers, would want to know the “bottom line” of the performance of HEIs, departments or faculties, in terms of the absolute performance or raw scores students enrolled in these would attain on an assessment.
- Individuals, HEIs and public policymakers wishing to assess the quality of the teaching services provided by HEIs would primarily be interested in the incremental learning deriving from higher education attendance, *i.e.* the “value-added” or “gain” provided by the HEIs measured by the scores an institution would attain after accounting for the quality of prior schooling or the degree of selectivity of the programmes and HEIs.

The OECD is committed to exploring both kinds of outcomes measures in an AHELO to meet the information needs of the various groups of stakeholders. However, it has to be recognised that assessing the value-added or marginal gain by HEIs raises a number of scientific and practical issues and the measurement of value-added would impose layers of complexity that, though theoretically well understood, are difficult to resolve in large scale assessments. Two sets of strategies can be considered in this area:

- One strategy is to focus the assessment on skill environments that students would typically not encounter prior to university entry, such as science, technology, engineering or health care/professions allied to medicine (STEM).
- The other strategy involves the *post-hoc* collection of data on prior learning as, for example, done in the Collegiate Learning Assessment (CLA), the synthetic linkage of outcome information of incoming students to outgoing graduates, or the longitudinal follow-up of cohorts of students. Each of these approaches has its merits and drawbacks (Nusche, 2007).

Given the complexity of measuring marginal gain, the proposed approach for the AHELO feasibility study is to explore both types of outcomes measures sequentially, and to examine the value-added perspective only conceptually in a first stage, while the pilot measurement of value-added in an international context would be incorporated into the design at a second stage.

For the purpose of the feasibility study, the value-added measurement strand will therefore consist in a review and analysis of possible methods for capturing marginal learning outcomes that can be attributed to HEI's attendance, both at a theoretical level and in terms of psychometric approaches. It will build upon similar work carried out at school level by the OECD and review options for value-added measurement in higher education. Researchers will be invited to study potential data sources, methodologies and psychometric evidence on the basis of datasets existing at the national level, with a view to providing guidance towards the development of a value-added measurement approach for a fully-fledged AHELO main study.

Main activities for field-testing of assessment and survey instruments

In terms of practical implementation, the Generic Skills strand, the two discipline-related Economics and Engineering strands and the contextual dimension surveys will follow a similar model of implementation, whereby participating countries will be involved in the development and translation/adaptation of the instruments to be tested. Table 2 provides a summary of the proposed activities and timelines.

Planning (June 2008 – October 2009)

- Design the sampling plan

The sampling plan will set out the minimum acceptable sample sizes (given the final design of the study), processes for selecting students, the quality control procedures that must be observed by countries and/or institutional coordinators, and the definition of exclusion criteria (*e.g.* part-time students).

- Recruit pilot HEIs in each participating country

Participating HEIs will need to be sufficiently different to provide a cross section of the sector for each participating country even though they will not be statistically representative of the diversity of the sector. For each small-scale test of instruments, the recruitment of pilot HEIs will therefore seek to ensure that prominent HEIs are included along less well-known HEIs, and where appropriate, that the pilot includes a mix of public and private or urban and rural HEIs. In countries with a binary higher education system, a couple of vocationally-oriented HEIs will be recruited as well.

- Recruit content and measurement experts in each participating country

Ultimately, the success of implementation of the small-scale tests of instruments, just as for any international assessment, will critically depend on having a well-trained team in each participating country. The recruitment and training of the content and measurement experts for each assessment and survey instrument will therefore be an integral part of the planning phase.

- Develop criteria to assess success of the international field testing

Decisions regarding the development of a large-scale implementation of AHELO across OECD countries will depend on the outcomes of the feasibility study. In this respect, it is important that the criteria for assessing the success of field-tests be developed upfront, when planning the initiative. It has to be borne in mind that the ambition at this stage is not to compare HEIs' performance. Rather, success can be defined in terms of providing a proof that the instruments

can be applied in diverse institutional, cultural and linguistic settings and yet provide valid and reliable and free-of-bias measures of student learning outcomes.

With respect to implementation, success will be defined in terms of student participation and response rates. A list of technical criteria and threshold measures to assess the degree of success of field-tests will be developed with the OECD and international experts.

- Establish a technical advisory group

A technical advisory group comprised of experts and individuals who have a leading operational role in the AHELO feasibility study will be established. This technical advisory group will be consulted on major technical issues (*e.g.*, instrument development procedures, sample and test design, translation procedures, or scoring and verification procedures).

Developing assessment frameworks and instruments (August 2009 – October 2010)

- Development of the assessment frameworks

Substantive input from participating countries and subject-matter experts will be imperative to ensure that the resulting instruments will be valid and reflect the cultural context of the countries in which the AHELO feasibility study will be implemented.

For the generic skills strand, the development of the assessment framework will build on the existing work done for the CLA. For the discipline strands, some initial work has been undertaken through the Tuning-AHELO project, the results of which test developers will be free to use as they see fit. For the contextual dimension, survey developers will be guided by the conceptual and analytical frameworks developed by a group of commissioned contextual dimension experts (Ewell *et al.*, 2008, 2009) and the results of a prioritization process of the underlying contextual variables by participating countries.

- Development of the assessment instruments

The development of instruments for implementation in different national, cultural, linguistic and institutional contexts is not straightforward. It is important to note that the goal at this stage is not to compare HEIs' performance, but rather to provide a proof of concept that instruments can credibly be applied in different national and institutional settings. This can be achieved by field-testing the instruments in a relatively small number of voluntary HEIs in each of a small number of countries, and collecting both cognitive and psychometric information on how well task items perform in cross-national and cross-institutional contexts. Test validity involves the accumulation of evidence that supports inferences about the characteristics of individuals/groups.

Student performance instruments will require a maximum duration of 90 minutes testing time, whereas the survey instruments for contextual variables will require a maximum duration of 15 minutes for the student and faculty questionnaires.

Translating, adapting, and pre-testing instruments (March – October 2010)

- Translation and adaptation of instruments

The instruments will undergo translation into the languages of the participating countries. Quality control during the translation process is an integral part of instrument development and is critical

to help ensure that the small-scale tests of the instruments provide data that are comparable across countries. To this aim, the translation process will be carried out drawing upon internationally-recognised techniques and procedures such as those used in PISA or PIAAC. They consist of a double translation (from an English source in the case of the AHELO feasibility study) to the destination language, followed by a process of reconciliation, verification and a final layout check by the national team.

- Pre-testing instruments

Once the instruments are translated and adapted for each country, they will be pre-tested in their home countries and refined based on findings. The pre-testing of each instrument will involve a small number of students per participating country and may comprise cognitive laboratory procedures, with an initial cognitive walk-through by analysts, and a set of cognitive interviews with respondents to review their thought processes retrospectively. Respondents will be probed to identify particular steps that might have been problematic.

Field-testing of instruments (November 2010 – April 2011)

- Implementation of field-tests

Once the instruments have been developed, translated and adapted, pre-tested and revised, they will be field-tested in each participating country (See Table 1 for country-specific testing windows).

Survey operations will be developed and implemented with related aspects of quality control, including the development of test administration procedures, the development of scoring procedures, and the training of country representatives in these procedures. These materials and procedures will be reviewed by the technical advisory group and the OECD Secretariat, and will be adapted to each participating country in light of national input.

Standards must be met with regard to validity, reliability, and comparability across countries. Indeed, perhaps the greatest risk in the AHELO feasibility study is the possibility of seemingly unimportant deviations introduced at the national level during the course of the implementation. Such decisions can undermine the integrity of the entire assessment in a particular country, and hence jeopardize conclusions to be made on the feasibility of an AHELO. The key design parameters shall play an important role in assuring final quality. Strict adherence to the technical standards will be needed to assure a high degree of uniformity in all participating countries, achieving standardized administration of the instruments and maximizing response rates.

- Delivery mechanism

The feasibility study will implement several instruments in different countries. It is anticipated that the assessment instruments and contextual surveys will be administered with the option of paper-and-pencil or electronic delivery, depending on operational considerations (logistics implications, security of testing, access to computers in some sub-systems for some participating countries), and the feasibility to implement electronic delivery within the allocated timeframe.

Analysis of field-test results and reporting (April - July 2011)

- Analysis of field-test results

The field-test results will be analysed by psychometricians to test the validity, cross-cultural appropriateness and linguistic transferability of the instruments. Outcomes of field-tests will be benchmarked against the technical criteria and threshold measures defined during the planning phase to assess the degree of success of the international field-tests. This analysis will also be reviewed by the technical advisory group to ensure that technical qualities have been met.

- Reporting

The field-test results will be reported to IMHE Governing Board and Education Policy Committee on the feasibility of adoption across OECD countries by the end of 2011 to inform decisions on next steps. The views of international experts and participants from the final conference on the feasibility of adoption across OECD countries will then be conveyed to the IMHE Governing Board and Education Policy Committee as background information to inform decisions on next steps from the OECD perspective.

Dissemination of the feasibility study results

- Conference

A final conference to discuss the outcomes of the feasibility study as well as the potential impact of an AHELO will be organised at the end of 2011. The OECD will then publish a comprehensive technical report covering all aspects of the AHELO feasibility study, summarizing and describing all data and statistical conventions or approaches followed in the AHELO feasibility study. A final report will summarise the results of the AHELO feasibility study, based on the analysis plan and analyses conducted for the feasibility study.

The targeted audience for the conference will bring together international experts, quality assurance experts, representatives of HEIs, students, businesses and policymakers. International experts will present the outcomes of the various strands of the feasibility study, and will report on the technical feasibility of developing a large-scale, fully-fledged, AHELO main study. Workshops and panels will also be organised to provide participants with a chance to interact and discuss operational issues, political economy constraints and the potential impact of such an assessment.

REFERENCES

- Collegiate Learning Assessment (United States): www.cae.org
- Dumais, J. (2009), *Higher education institutions, sampling issues preparing the AHELO feasibility study*, OECD, available from www.oecd.org/edu/ahelo/callfortenders.
- Ewell *et al.* (2008), *Report on the 1st meeting of experts on the AHELO contextual dimension*, OECD, available from www.oecd.org/edu/ahelo/callfortenders
- Ewell *et al.* (2009), *Analytical framework for the contextual dimension of the AHELO feasibility study*, OECD, available from www.oecd.org/edu/ahelo/callfortenders
- German Centre for Higher Education (CHE): www.che-concept.de/cms/?getObject=302&getLang=en
- Nusche, D. (2007), *Assessment of Learning Outcomes in Higher Education: A Comparative Review of Selected Practices*, OECD Education Working Paper No. 15, accessed 3 June 2008 from www.oecd.org/dataoecd/14/8/40257354.pdf
- Tuning Association (2009a), *A Tuning-AHELO conceptual framework of expected/desired learning outcomes in the Science of Economics*, OECD, available from www.oecd.org/edu/ahelo/callfortenders
- Tuning Association (2009b), *A Tuning-AHELO conceptual framework of expected/desired learning outcomes in Engineering*, OECD, available from www.oecd.org/edu/ahelo/callfortenders

Table 1 – Country-specific testing windows – AHELO feasibility study

Dates	
<i>Generic Skills Strand</i> *	
Korea	01 Nov - 26 Nov 2010
Finland	07 Feb - 04 Mar 2011
Norway	07 Feb - 04 Mar 2011
Mexico	07 Mar - 01 Apr 2011
<i>Economics Strand</i> *	
Netherlands	28 Feb - 25 Mar 2011
Belgium (Fl.)	07 Mar - 01 Apr 2011
Italy	07 Mar - 01 Apr 2011
Mexico	07 Mar - 01 Apr 2011
<i>Engineering Strand</i> *	
Australia	02 Aug - 27 Aug 2010
Japan	29 Nov - 24 Dec 2010
Sweden	14 Feb - 11 Mar 2011

* Participation in the Contextual Dimension and the extent of contextual data collection is optional at national level.

Table 2 –Proposed activities and timelines

• 2009	<ul style="list-style-type: none">• Review, develop and finalise frameworks• Start item development• Develop survey questionnaires and documentation collection process• Develop the overall assessment design and analysis plan• Develop translation/adaptation guidelines
• 2010	<ul style="list-style-type: none">• Item development• Finalise instruments• Translation/adaptation of instruments and surveys• Develop sampling plan• Finalise quality assurance plan and survey procedures• Provide training to national staff prior to administration• Administer instruments and survey procedures in Southern Hemisphere
• 2011	<ul style="list-style-type: none">• Administer instruments and survey procedures in Northern Hemisphere• Provide clean data set for analysis• Evaluate implementation and outcomes of the feasibility study• Data analysis• Prepare final and technical reports• Release of final and technical reports
