

**DIRECTORATE FOR EDUCATION
INSTITUTIONAL MANAGEMENT IN HIGHER EDUCATION GOVERNING BOARD**

Group of National Experts on the AHELO Feasibility Study

PROGRESS REPORT ON GENERIC SKILLS STRAND

Paris, 15-16 March 2010

This document presents a brief status update of the progress of work for the Generic Skills strand and provides responses to questions and clarifications asked by the GNE members.

The AHELO GNE is invited to:

- *TAKE NOTE of progress report, and COMMENT on progress as necessary.*

Contact:
CAE: cae.ahelo@gmail.com
OECD Directorate for Education: Diane.Lalancette@oecd.org

JT03279340



PROGRESS REPORT ON GENERIC SKILLS STRAND

Introduction

1. During its third meeting on 18-19 November 2009 in Paris, the OECD Secretariat reported on progress with contracting of the Generic Skills Strand [EDU/IMHE/AHELO/GNE(2009)16].

2. In the following discussion, the GNE:

- **ASKED** for statistical evidence that reducing the number of performance tasks from three to two will not affect the results for the Generic Skills strand;
- **ASKED** to retain at least nine performance tasks for participating countries to select from as proposed in the original CAE proposal plan;
- **SOUGHT** clarification on whether the “making an argument” task would still be measured;
- **NOTED** that the translation and review teams require many participants and that countries may not be in a position to supply that many people, and **ASKED** for flexibility in reducing the number of participants for the translation and review activities; and
- **ASKED** for clarification on the optional participation of countries in the CLA faculty academies and the number of participants per country, **TOOK NOTE** that a half-day workshop to present the faculty academies is offered to interested countries in January in Paris and **NOTED** that an official invitation would be sent to all countries.

3. This document presents a brief status update of the progress of work for the Generic Skills strand and provides responses to questions and clarifications asked by the GNE members.

4. The AHELO GNE is invited to:

- **TAKE NOTE** of progress report, and **COMMENT** on progress as necessary.

Summary of Contents

5. Included in this brief update is a status checklist based on the CAE Work Plan, and an overview of accomplishments from the Generic Strand meeting that took place on 15-18 February, 2010. Also included is a summary of collaboration between CAE and ACER, and general questions/issues raised that are of potential importance to the GNE. Note that: 1) the official start date of Module A was February 1, 2010, and 2) a report will be provided at Milestone 1 as agreed upon between OECD and CAE (May 31, 2010): “Deliver to OECD descriptions of the two PTs that have been adapted by Country Teams for the test translation process to begin.”

AHELO Module A Work Plan

Project Phase 1: Pre-Contract Planning

A. Activity:

The Participating Countries in collaboration with CAE and following established criteria will select Country Team Members for the project including AHELO Group of National Experts (GNE) and Assessment Expert(s).

Status: *Completed*

In anticipation of phone calls between CAE staff and Country Teams (see below), countries nominated and sent vita for proposed assessment experts. Because of the wide range of skills required in the person that serves as assessment representative for each country, we encouraged countries to nominate several candidates from which CAE collaboratively selected one. Country Team nominations were guided by criteria established by CAE (see Annex 1). Of course not all criteria could typically be fulfilled by one person; the criteria served as a guide for selection. One (and in some cases several) assessment representative was selected for each country based on her or his formal training, professional experience, and set of technical qualifications that are relevant to the characteristics of the project. In each instance, we concurred with the recommendations of the country.

B. Activity:

All parties, including the Country Team Members, OECD, and CAE personnel, must sign appropriate confidentiality agreements that protect CAE's trade secrets, know-how, and proprietary methods and to comply with OECD's confidentiality requirements as stipulated in the contract.

Status: *Completed*

The Collegiate Learning Assessment contains proprietary performance tasks. If disclosed to the public, the assessment would be severely compromised. Consequently, and in cooperation with OECD/AHELO, all participating country team members were asked to sign a confidentiality agreement that defined the terms of confidentiality and consequences if that agreement were broken.

CAE staff and country team member jointly identified personnel who had a "need to know" in order to advise on performance task selection. From a total of nine performance assessments reviewed two would be selected for the AHELO feasibility study. All such team members signed confidentiality agreements, with one exception. One United States team member, a state employee, was forbidden to sign the agreement by that state's legal office (one of four participating states).

C. Activity:

CAE will discuss with designated Participating Country representatives (by phone) any special needs or concerns that need to be given consideration during the first meeting of the Country Representatives.

Status: *Completed*

Immediately after the December 2009 holiday break, CAE staff contacted Country Representatives with a package of material including an introductory letter (see Annex 2).

Following the letter, we contacted country teams and arranged a phone call with each. Phone calls took place between January 13 and January 21, 2010. The call lasted from 45 minutes to almost two hours, depending on the experience of the country representatives with AHELO. Typically those teams that were newest to the project needed more time to learn about it. The questions and discussions were collegial and covered a wide range of topics including: (a) how might country teams organize themselves to carry out the study, (b) what should country teams tell colleges and universities (institutions) regarding what is required of them if they were to participate in the project, (c) selection of an assessment expert for participation in the project, (d) staffing requirements, (e) progress in recruiting institutions, (f) what would be involved task selection, (g) what would be involved in the task adaptation and translation process, (h) the nature of the confidentiality agreement.

D. Activity:

CAE will develop a conceptual framework and procedures for training Country Representatives to adapt each CLA Performance Task's (PT) contents to reflect their contexts (see GS.4¹ Conceptual Framework for Performance Task Adaptation and Translation; GS.13 Translation and Adaptation Guidelines PowerPoint; GS.14 Translation Flowchart).

Status: *Completed*

The Conceptual Framework serves as a guide for task adaptation and translation, and also as a basis for training teams at the meeting held in New York City 15-18 February 2009. The Framework notes that the "Generic Strand" can be divided into two broad parts. The first part involves the selection, adaptation and translation of two performance tasks from the Collegiate Learning Assessment for their use in the project. The second part involves implementation including data collection, analysis, and reporting. The conceptual framework presented focuses on the first part of the Generic Strand.

The Conceptual Framework covered the following topics: (a) architecture of the CLA (task description, development, scoring and example), (b) task adaptation, (c) performance task translation, and (d) performance task translation and review. It provided a set of guidelines for each topic in a series of tables that could be used in practice.

E. Milestone:

CAE will plan and organize the Country Team Meeting in New York and assemble materials for circulation to the Country Team Members (see GS.15 Communication Log for AHELO Study).

Status: *Completed*

The meeting was held as planned, in New York City 15-18 February, 2009. By all accounts, it achieved its objectives of informing and training country teams about: (a) the nature of the Collegiate Learning Assessment (its tasks, response format, and scoring rubric), (b) the process of adapting and translating CLA Performance Tasks, (c) the selection of two Performance Tasks

¹ All Generic Strand documents are numbered sequentially: GS.#. A complete listing of GS documents, to date, may be found in GS.15. Documents are available on the AHELO Exchange website.

from nine candidates—Catfish and Lake-to-River (described briefly below in this report), and (d) practice in adapting the Catfish Performance Task followed by detailed discussion of how variation among countries might affect translation (e.g., name of river and lake; metric system).

The CAE team will issue detailed notes from the meeting in the next several weeks. They are too extensive and undeveloped at this point in time (February 26, 2010) to put into this progress report.

Project Phase 2: Country Team Meeting, NY

A. Activity:

Familiarize teams with the CLA (see GS.12 AHELO GS Meeting Overview)

Status: *Completed*

Marc Chun led a full day workshop on the CLA for participants that aimed to get them deeply involved in the architecture of CLA performance tasks, response formats, and scoring. Participants worked in teams throughout the day: (a) taking one performance task—“Crime”—and discussing among teams just what kinds of thinking and performance were being measured. (b) This was followed by reading typical student responses to the task to develop an understanding of what kinds of performance the task evoked. (c) Participants then turned to scoring and initially intuitively scored responses learning that even without a rubric or training, they pretty much agreed on performance levels; they then applied the CLA scoring rubric for the performance task. The day was concluded with a presentation and video of the CLA internet platform and how it worked.

B. Activity:

Review a subset of at least nine CLA PTs and select four PTs considered suitable and valid in an international context, two of which will be selected for modification and use in the AHELO feasibility study based on recommendation of CAE. The subset will be selected using certain criteria (see GS.5 Criteria for Choosing a Subset of CLA PTs).

Status: *Completed*

At the beginning of the NYC meeting, teams discussed all of the performance tasks: 1. *Airplane* 2. *BrainBoost* 3. *Catfish* 4. *Crime Reduction* 5. *Dogtooth* 6. *Lake to River* 7. *Museum* 8. *Parks* 9. *Skating Accident*. They agreed to take a “straw vote” to see how much they agreed or disagreed about the top candidates for inclusion in the study. The results showed “Catfish” and “Lake-to-River” to be widely popular. Nevertheless, the teams spent much of two days off and on discussing each performance task. “Catfish” raised only minor issues of translation and all agreed there were no obstacles. “Lake-to-River” raised the issue of student familiarity with lakes and rivers and led to a discussion of universality of performance task. The issue was resolved by noting that while higher education students may not directly experience lakes and rivers every day, a higher education student should be able to reason about environment issues surrounding such a situation. Hence, Lake was unanimously selected as well. (Both tasks are briefly described below.)

C. Activity:

The Assessment Experts from the Country Teams will form an Assessment Adaptation Group (AAG), which CAE will train in task adaptation, the translation and review process, procedures, and the recruiting of test translation teams.

Status: *Pending* CAE visits to respective countries in May/June/July 2010.

Initial training, as described in various activities described above, was carried out in NYC. Follow up training will be carried out on March 17, 2010 at OECD in conjunction with the GNE meeting. Representatives from the Country Teams will meet and work through the two performance tasks, raising issues and reaching agreement as to how the English version of the tasks will be modified. CAE staff will take recommendations under advisement and create a revised version of each task for further Country Team review. Finally, a member of the CAE staff will visit each country and work with assessment experts and translators to translate, field test, and revise the tasks.

AHELO Module A Meeting: February 15-18, 2010

Major Accomplishments:

- Two CLA Performance Tasks (PTs) were agreed upon for the feasibility study: “Catfish” and “Lake-to-River.” The “Catfish” PT sets up the following situation: A deformed catfish was recently discovered in a local lake. The mayor of the city asks you to review evidence (documents) that support or refute three alternative explanations for the appearance of a 3-eyed catfish. The “Lake-to-River” PT sets up the following situation: A county Board of Supervisors is considering the pros and cons of removing the county’s dam. You have been asked to review a set of documents and recommend a course of action.
 - Adaptation was begun on one of these PTs
 - Meeting scheduled for 17 March (Paris) as many of the country representatives and Rich Shavelson will be together for the GNE meeting – main agenda item will be to discuss adaptation of the two selected performance tasks
- *CLA in the Classroom* workshop was conducted. The workshop provided participants an in-depth understanding of:
 - Elements of a CLA performance task
 - What to look for in scoring a performance task
 - How to use the scoring rubric
- Framework for adapting and translating the performance tasks was presented and discussed in detail.
- Implementation questions/concerns were addressed²:

² Detailed meeting notes will be distributed later in March.

- Agreement that, to the extent possible, dissemination of information about the feasibility study to institutions, faculty, staff, and students will remain constant across all countries so as to mitigate potential advantages.
- Universality of performance tasks—there is a duality to this construct. It can refer to commonality of local situations across countries or to knowledge and experiences expected of most all college students across countries. We adopted the latter definition for this project.
- How to plan for comparability among diversely organized higher education institutions within and across the participating countries (this will be an ongoing effort).
- Each of the participating country teams provided brief overviews of their respective higher education systems so as to familiarize each other on similarities and differences.
- How to maintain inter-rater reliability not just among country scorers, but also across country scorers (answers: translate a small number of student responses from each country into English for inter-scorer training; place common “anchor” tests throughout each country; have scorers collaborate with translators so that there is a more comprehensive understanding of the PTs and scoring rubric).
- Recommended 100% double scoring, although sampling and scoring frameworks have yet to be agreed upon; working with ACER, cApStAn, and other related groups on this.
- AHELO Feasibility study organization – what are the various modules, how do they interact with Module A, ACER and implementation phase, etc. were addressed (see General Questions/Issues below for more information).
- Administrative concerns – Examples of general administrative/logistical questions/answers included:
 - number of computers needed
 - no minimum requirement of number of computers as there will be multiple testing sessions
 - what the CLA internet interface looks like
 - a video was shown during the meeting that demonstrated the interface
 - test proctor responsibilities
 - the proctor interface was also shown during the video; additionally, CAE will provide countries with a proctor manual that will need to be translated
 - number of scorers needed
 - suggested 10 scorers total: with the two tasks administered in each institution, 5 scorers per task
 - training of scorers
 - CAE will visit each country during testing period to conduct a rigorous scorer training program
 - what happens in a power outage
 - the students’ responses are saved every thirty seconds so only a small portion of the test may be lost due to power outage

Collaboration and coordination with ACER

6. The following activities were undertaken with the ACER Consortium:

- Participating in all ACER conference calls and meetings.
- Coordinating Module A activities with Module E efforts.
- Reviewing, editing, and making recommendations on all documentation, including those listed below.
 - <https://ahelo.acer.edu.au/display/ModuleE/Reporting+Guidelines> New: first draft for review
 - <https://ahelo.acer.edu.au/display/ModuleE/Analysis+Plan> First draft for review
 - <https://ahelo.acer.edu.au/display/ModuleE/Assessment+Design> For final review
 - <https://ahelo.acer.edu.au/display/ModuleE/Sampling> Draft with queries for review
 - <https://ahelo.acer.edu.au/display/AC/10-02-17> Notes from last Consortium meeting for review and revision
- Posting all Module A documentation and general emails to the AHELO Exchange website.
- Collaborating on technical issues including: sampling, modeling, language translation (Statistics Canada, cApStAn, and Data Procession Center).

General questions/issues

7. There is still a lot of confusion about the implementation phase of the study and how all the other modules interrelate. Diane Lalancette (OECD) provided a brief overview (see GS.22 OECD AHELO Overview for CAE GS Meeting PowerPoint); however, participants suggested that an overall timeline/flowchart for the entire Module A (development and implementation) would be very helpful in understanding project implementation.

8. Some countries are still unsure of how they need to be organized for the study (including infrastructure and resource costs for the national teams), but particularly expected time and monetary costs for the participating colleges and universities (“institutions”), potential incentives costs (see #5 below), and training and scoring time/monetary costs.

9. There was a general discussion on student incentives, but it seemed the countries would like to hear about alternative ways to motivate students to participate in the study (e.g., more than just a stipend).

10. There was continual discussion about how to maintain consistency across countries for potential comparison purposes in: a) type of institutions, b) majors or programmes of study, c) age of students, d) scoring, and e) representativeness of student/institution (including how to best select institutions).

11. The following items provide responses to the questions and clarifications asked by the GNE members at the November meeting.

- The GNE members ASKED for statistical evidence that reducing the number of performance tasks from three to two will not affect the results for the Generic Skills strand.

ANSWER: For cost reasons, the decision was made to reduce the number of performance task from three to two. Two tasks is the minimum needed to estimate the variance of performance on the tasks, and to test out the feasibility of task adaptation and translation guidelines and procedures. (One criterion for task selection was that the tasks be demonstrably different.)

Of course, ideally, we would use six to nine performance tasks for this purpose. The larger the sample the better, ignoring costs. But due to cost constraints, we are using the absolute minimum size task sample possible (two tasks).

- The GNE members SOUGHT clarification on whether the “making an argument” task would still be measured.

ANSWER: No, this will not be measured. Note this was also clarified in document *GS.4: Conceptual Framework for Adaptation and Translation* sent to country participants and provided to participants at the Generic Strand meeting in NY.

- The GNE members NOTED that the translation and review teams require many participants and that countries may not be in a position to supply that many people, and ASKED for flexibility in reducing the number of participants for the translation and review activities.

ANSWER: Yes, there is flexibility within countries not to use all personnel as originally envisioned in our proposal. At a minimum a country will need three translators: two translators for dual translation at one point in time and one (different) translator later on during translation review. In addition, a minimum of one assessment expert/translation & adaptation advisor would be involved in the process as well (see GS. 14).

**ANNEX 1 - REQUIRED QUALIFICATIONS FOR THE COUNTRY TEAM REPRESENTATIVE
– GENERIC SKILLS STRAND**

Education and formal training

Ph.D. in psychometrics, statistics, social-science measurement or related discipline

English proficiency

Fluent in English (both conversational and reading/writing)

Professional experience

Demonstrated experience providing technical expertise and leadership in at least two projects related to the following areas:

- higher education
- international or multicultural education
- academic achievement assessment
- assessment development, preferably constructed-response tasks
- computer simulations or computer-based testing
- test translation and test translation review

Desired Qualifications

- Experience coordinating the collection, management, and analysis of data for educational projects
- Ability to successfully work with multidisciplinary and/or multicultural teams
- Record of technical publications in the areas of expertise

**ANNEX 2
INTRODUCTORY LETTER TO PARTICIPATING COUNTRIES
IN THE GENERIC SKILLS STRAND**

The introductory letter began as follows:

“January 8, 2010

“Dear AHELO Representative(s),

We enthusiastically welcome you to the Assessment of Higher Education Learning Outcomes (AHELO) feasibility study. This is the beginning of what we hope to be an exciting study of if and how a performance assessment of college students’ critical thinking, analytic reasoning, problem solving and communication skills can be used across cultures and languages.

We look forward to meeting you in person at our scheduled meeting from February 15-18, 2010 in New York City at the Yale Club (<http://www.yaleclubnyc.org/>).

We have provided some initial details of this meeting in Annex F, attached. Our next correspondence will also provide more extensive information (e.g., conceptual framework for modifying performance tasks to fit country contexts and for translating performance tasks from English into home language, general instructions for taking the CLA on an internet platform, summary descriptions of 9 CLA performance tasks, our recommendations and rationale for focusing on a subset [4-5] of the performance tasks)....

In the remainder of this letter, we aim to introduce ourselves and the project in a bit more detail. To do so, we have provided biographies and a contact list for the primary CAE actors in Annexes A and B, respectively. A second purpose is to provide some specific project background documents in Annex C. These provide a rationale for and a description of the Collegiate Learning Assessment (CLA). Third is to provide the currently envisioned work plan for the first year of the feasibility study in Annex D. The fourth is to provide an initial listing of contacts in the participating countries, also in Annex B. Fifth is to outline, in Annex E, criteria for selecting country assessment representatives. And lastly, we have included some information about our first meeting in New York City from February 15-18, 2010.”