OECD-AHELO State of Progress

- OECD-AHELO: An international test that aims to measure what students in higher education know and can do upon graduation.

- OECD-AHELO Feasibility Study
  - Is the assessment scientifically possible?
    - Measuring what? Defining competence frameworks
    - How? Defining measurable learning outcomes that demonstrate mastery of a given competence, and developing instruments that measure attainment of those learning outcomes.
    - Scoring: Defining through scoring rubrics the scope and level of learning outcomes to be attained, and sharing an understanding of the framework.
  - Is the assessment practically possible?
    - Getting universities and students to participate.
  - Can we verify the reliability and validity of the instruments?
    - The conclusion of the Technical Advisory Group.

OECD-AHELO進捗状況

- OECD-AHELO: 大学教育の成果を世界共通のテストで測定する国際事業。

- OECD-AHELOフィージビリティ・スタディ（実施可能性の調査）
  - テスト問題を開発することはできるのか。
    - 何を測定するのか･･･コンピテンス枠組を定義する。
    - どう測定するのか･･･コンピテンスを学習成果に具体化して設問に落とし込む。
    - どのように解答するのか･･･解答ルーブリックをとおして、学習成果の範囲（観点）と準拠を規定し、認識を共有する。

- テストを実施することはできるのか。
  - 大学・学生の協力を獲得する。

- テストの妥当性と信頼性を検証することはできるのか。
  - TAG（技術諮問グループ）の結論。

Informing Universities for Educational Improvement

The AHELO Feasibility Study Experience in Japan, Canada, and Australia

Satoko Fukahori
National Institute for Educational Policy Research
OECD-AHELO本調査にむけて
-だれのため・何のために実施するのか-

- 調査を実施することは、だれに対して、どのようなメリットをもたらすのか。
- 調査結果をどのように分析するのか。
- だれに対して、どのように報告するのか。

※ 明確なビジョンを共有することなく、取組を発展させることはできない。

AHELO-FSの調査設計
- 組み込まれた方向性 -

- 任意の大学（便宜サンプル）
  - 各国の大学システムを代表するサンプルではない。
    - 「大学（高等教育機関）」母集団の定義は一律ではない。
    - 大学を母集団から無作為に抽出し、調査への参加を取りつけるのは容易ではない。
  - 「国」の比較は適切ではない。

- 各大学で学生を無作為に抽出
  - 学生在籍する大学（プログラム）を代表するサンプル。
  - 項目反応理論にもとづいて、学生は異なる問題セットに解答。
    - 「学生」の定義を標準化して比較可能にすれば、問題の難易度や分析に関する情報が必要（現段階では情報の蓄積がない）。学生を分析の単位とした調査結果の活用は、直ちには難しい。

- 分析の単位は「大学」
  - AHELOは、大学に対して、その教育改善に資する情報を提供する
    ために実施する。

If there is a Future AHELO
-For whom and for what?-

- Who will benefit from a future AHELO and how?
  - How will the results be analyzed, and how will they be reported?

※ In order for a future AHELO to be successful, we will need to share a clear vision on “for whom and for what”

The AHELO-FS Research Design

- A convenience sample of universities
  - The university sample does not represent the country’s HE system.
    - The “university population” is not self evident – what is a university?
    - Randomly sampling a university from the university population, and gaining their participation at a high rate will not be easy.
  - AHELO-FS was not designed to compare countries.

- A random sample of students
  - The student sample represents the university’s/program’s student population.
  - Students responded to different sets of items based on the item response theory model.
    - Information such as the difficulty level and student distribution of each item are necessary in order to standardize and make comparable student test scores. Since such information is not yet available, it is not at this point possible to conduct a robust analysis with students as the unit of analysis.

- The unit of analysis is the “university.”
  - The purpose of AHELO is to provide institutions with information that will help them improve their education.
What are the kinds of information that universities seek from AHELO?

- **Consultation:**
  - NIER Project Study on the Analysis of the AHELO-FS Results (2013).

- **1. International benchmarking**
  - How are their students performing in relation to their international peers?
  - Not ranking.

- **2. Diagnosis of student competencies.**
  - What are the strengths and weaknesses of their students’ competencies?

- **3. The relationship between teaching and learning environments and student performance.**
  - How can programs and student support systems be improved?
The information that universities seek.

- Hypothetical Data -

Institutions must agree in advance to share information on the names of institutions constituting the benchmark group. Ideally, institutions should be able to choose their peers.

Instruments need to be designed in a way that will allow for the analysis of competence clusters.

The Relationship Between Teaching and Learning Environments and Student Test Scores - an Exploratory Analysis -


Recommendation 1: Induce qualitative change in Bachelor level education through encouraging autonomous study.

- Increasing time spent on active learning such as discussion and debate.
- Encouraging learning outside the classroom, such as internship.
- Hypothesis 1: The lower the proportion of "lectures" and higher the proportion of "seminars" and "group-work," the stronger the capacity for "autonomous study."

Recommendation 2: Increase quality learning time.

- Securing sufficient total time for autonomous study — preparing for class, taking classes, and later developments.
- Hypothesis 2: The longer total time spent on study, the stronger the capacity for "autonomous study."
**Test Scores**

Do they measure “ability to think autonomously?”

- **SWLE (Scaled Weighted Likelihood Estimate)**
  - Combined total score of Multiple Choice Questions and Constructed Response Tasks.
  - Estimates of the student ability calculated based on the Item Response Theory model.
  - Average 500, Standard Deviation 100.

- **Student Comments to the Test.**
  - **Multiple Choice Questions**
    - Most of the contents were covered in class, so everyone should be able to solve them.
    - The questions were sporadic, superficial, and unconnected. It would be better to delve more deeply into the subject matter.

  - **Constructed Response Tasks.**
    - Interesting questions, asking what caused the problem and how to solve them. Useful in practice.
    - In university, we deal mostly with theoretical and abstract problems. The test was interesting because it dealt with real world problems. I liked the ethical questions.

    - If constructed response tasks are important, we should have more classes with group work and case study. We have not had such experience, nor chances to learn approaches to solving these types of questions.

---

**Sharing National Data**

An MOU between Australia, Canada, and Japan

- For the purposes of educational research each Party desires to disclose and share its respective national Data with each other Party (“the Purpose”).

- The Parties agree that no Party shall disclose to any third Party, publish or otherwise disseminate any other Party’s national Data disclosed to it under this Memorandum of Understanding or findings or research based on such Data without the prior written permission of the other relevant Party or Parties.

- Any major public announcement in connection with this Memorandum of Understanding must be agreed by the Parties before it is made, except if required by law or a regulatory body, in which case the party required to make an announcement must, to the extent practicable, first consult with and take into account the reasonable requirements of each other party.
Hypothesis 1: The lower the proportion of "lectures" and higher the proportion of "seminars" and "group-work," the stronger the capacity for "autonomous study."

- The proportion of "lectures"
  - 81% of the students in the Japanese sample responded that "61-75%" or "75%+" of their classes were of "lecture" type. The test scores for these groups were relatively high.
  - A similar pattern was found in the Country A sample.
  - In the Country B sample, students responding "41-60%" and "61-75%" scored relatively high on the test.

- For the three country-samples, around 60% "lecture" type seems to be the preferable proportion.

For the Japanese sample, students spend relatively little time on active learning activities. However, those taking more "seminar" or "group-work" type classes do not necessarily score better. The preferable proportion seems to be about 25%?

- "Seminars"
  - 41% (4%)
  - 25-40% (36%)
  - 25-40% (28%)
  - 0-25% (16%)

- "Group work"
  - 41% (16%)
  - 25-40% (28%)
  - 25-40% (15%)
  - 0-25% (11%)

- "None"
  - 41% (16%)
  - 25-40% (28%)
  - 25-40% (15%)
  - 0-25% (11%)

- "Less than 25%"
  - 41% (4%)
  - 25-40% (36%)
  - 25-40% (28%)
  - 0-25% (16%)

- "More than 50%"
  - 41% (16%)
  - 25-40% (28%)
  - 25-40% (15%)
  - 0-25% (11%)

Satoko Fukahori 2013.12.10 NIER Symposium
The ideal pattern and implications for improvement may vary among institutions.

### Hypothesis 2: The longer total time spent on study, the stronger the capacity for “autonomous study.”

<table>
<thead>
<tr>
<th>Number of Hours Spent on Activities</th>
<th>First Semester of Third Year (per 7-day week) [average]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attending formal classes</td>
<td>Japan: 4.5 hours, A: 6.2 hours, B: 4.3 hours</td>
</tr>
<tr>
<td>Preparing for class</td>
<td></td>
</tr>
<tr>
<td>Practical training</td>
<td></td>
</tr>
<tr>
<td>Paid work related to study</td>
<td></td>
</tr>
<tr>
<td>Paid work unrelated to study</td>
<td></td>
</tr>
<tr>
<td>Co-curricular activities</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Total time spent on study per day

- **Japan**: 4.5 hours
- **A**: 6.2 hours
- **B**: 4.3 hours

### HE Sample
- **HE Sample in Japan**
- **HE Sample in Country A**
- **HE Sample in Country B**

---

**Satoko Fukahori**

2013.12.10 NIER Symposium
国際的な学習成果アセスメントによる
大学の教育改善に資する情報提供にむけて

・大学が求めている教育情報
  - 自校の生徒の学習成果の特徴を、国際的な水準と照らし合わせて客観的に捉えるための情報。
  - 国際的なベンチマーク
  - コンピテンシー・プロフィール
  - 教育改善の方向性に係る示唆を得るための情報。
  - 大学の教育環境や生徒の学習実態を、生徒の学習成果との関連のなかで捉える。

・大学教育の改善に資する学習成果アセスメントの条件
  - テストで測定するコンピテンシーが、大学教育をとおして育成しようとしているコンピテンシーと一致していること（整合性）。
  - コンピテンシーにもとづく学位プログラムの体系化

Informing Universities for Educational Improvement based on Higher Education Learning Outcomes

- The information on education that universities seek
  - Information that will allow universities to better understand the state of learning outcomes of their students in relation to international standards.
    - International benchmarks
    - Competence frameworks
  - Information that will provide implications on ways to improve education.
    - The relationship between teaching and learning environment and student learning outcomes.

- When does learning outcomes assessment most effectively inform universities?
  - When the competences being measured are aligned with the competencies being pursued through university education.
  - The importance of developing competence based degree programs.


Tuning Association. Tuning Educational Structures in Europe. (http://www.unideusto.org/tuningeu/)


Satoko Fukahori

References

• OECD. Testing student and university performance globally: OECD’s AHELO. (http://www.oecd.org/edu/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdsahelo.htm)

Tuning Association. Tuning Educational Structures in Europe. (http://www.unideusto.org/tuningeu/)


Satoko Fukahori

Thank you for your attention!

fukahori@nier.go.jp
tuning-ahelo@nier.go.jp