

TUNING-AHELO コンピテンス枠組の 共有と水準規定による グローバル質保証

TUNING-AHELO Global Quality Assurance through Sharing Competence Frameworks and Degree Level Specifications

平成25年12月10日(火)

December 10, 2013

国立教育政策研究所

National Institute for Educational Policy Research

は し が き

「大学を修了する時点で、学生はどのような『力（コンピテンス）』を身に付けていることが期待されるのか」「大学はどうすれば、国際的に通用する水準の教育を提供しているといえるのか」大学進学率が上昇し、グローバル化が進展する中で、これらの視点は、大学教育の質保証を検討する上で欠くことができなくなってきました。

こうした状況に鑑み、コンピテンス重視の大学教育を推進する先駆的な取組である「チューニング（Tuning）」と経済協力開発機構（OECD）による「高等教育における学習成果調査（AHELO）」に注目し、グローバル化時代における大学教育の質保証に関する理解を深めることを目指すシンポジウムを平成25年12月に開催しました。

今回のテーマを企画した背景には、コンピテンス重視に転換してきた我が国の高等教育政策の動向があります。1980年代末より多様化・個性化の路線をたどってきましたが、近年では質保証の基盤となる共通の枠組みづくりに重点がシフトしてきています。中教審による平成20年のいわゆる「学士力」答申がそれを象徴していますし、平成24年のいわゆる「質的転換」答申でも、学位プログラムを確立する必要性や教学マネジメントの在り方に焦点が当てられています。こうした環境の中で、「コンピテンス枠組の共有と水準規定」のための具体的な方法としての「Tuning」や「AHELO」は、極めて時機を得た重要なテーマだったと考えています。

我が国では、グローバル化する知識基盤社会をけん引する人材を養成するために、大学教育にますます大きな期待が寄せられています。その一方で、大学進学人口の拡大と18歳人口の減少によって、大学教育の質保証の問題が顕在化してきています。

今回のシンポジウムは、大学の質保証をめぐる国際的な動向に注目することで、グローバル化時代における大学教育の質保証の在り方を考察する重要な場とすることができたと考えています。

今回のシンポジウムの内容が、この報告書によって更に多くの方々に広がり、我が国の大学教育のリ・デザインに役立てば幸いです。

平成26年3月

国立教育政策研究所長

尾 崎 春 樹

平成25年度教育改革国際シンポジウム
TUNING-AHELO
コンピテンス枠組の共有と水準規定によるグローバル質保証

平成25年12月10日(火)

目次

開会挨拶	尾崎 春樹 (国立教育政策研究所長)	7
第一部 講演		
基調講演	現行の大学評価制度はどのように導入され、大学によってどのように受け止められているか (木村 孟)	11
講演 1	TUNINGコンピテンス枠組：教育改革と学習のグローバルな測定・比較に向けて (ローベルト・ワーヘナール)	19
講演 2	AHELO フィージビリティ・スタディ：調査結果と技術諮問グループ(TAG)の結論 (ピーター・ユーウェル)	37
講演 3	日本におけるAHELOの取り組みと今後の展望 (岸本 喜久雄)	47
第二部 パネルディスカッション		
報告	大学の教育改善に資する情報提供にむけて ー日本・豪州・カナダにおけるAHELOフィージビリティ・スタディの取組 (深堀 聡子)	59
パネルディスカッション		76
閉会挨拶	杉野 剛 (国立教育政策研究所次長)	89
資料		
	AHELOの経験：オーストラリアにおける実施、成果及び教訓 (ダニエル・エドワーズ)	93
	AHELO：オンタリオ州教育機関の経験 (メアリーキャサリン・レノン)	110
	登壇者プロフィール	120

International Symposium on Educational Reform 2013
TUNING-AHELO
Global Quality Assurance through
Sharing Competence Frameworks and Degree Level Specifications

December 10, 2013

Contents

Opening Remarks	(Haruki Ozaki Director General, NIER)	129
Section 1: Presentations		
Keynote Speech	How has the current assessment scheme been introduced and how has it been received by HEIs? (Tsutomu Kimura)	133
Presentation 1	TUNING Competences Frameworks : Key Tools for Educational Reform and the Assessment and Comparison of Learning in Global Perspective (Robert Wagenaar)	141
Presentation 2	The AHELO Feasibility Study : Study Results and the Conclusions of the Technical Advisory Group (TAG) (Peter Ewell)	160
Presentation 3	AHELO in Japan and Its Prospects (Kikuo Kishimoto)	170
Section 2: Panel Discussion		
Report	Informing Universities for Educational Improvement : The AHELO Feasibility Study Experience in Japan, Canada, and Australia (Satoko Fukahori)	181
Panel Discussion		199
Closing Remarks	(Tsuyoshi Sugino Deputy Director-General, NIER)	213
Appendix		215

開会挨拶

尾崎 春樹
国立教育政策研究所長

- 報告書作成に当たり、当日の発言内容に修正を加えていることがあります。
- 所属団体、職名は2013年12月10日現在のものであります。

開会挨拶

国立教育政策研究所長
尾崎 春樹



こんにちは。本年度の教育改革国際シンポジウムの開催に当たりまして、主催者を代表して一言御挨拶を申し上げます。

この教育改革国際シンポジウムは、諸外国の教育改革の最前線で活躍されている専門家の方々をお招きいたしまして、各国の経験から学ぶ、また、我が国の教育改革の実践にそれを生かしていくということを目的として、平成13年度から開催をして参りました。

本年度のテーマは、今、深掘先生から御紹介がありましたとおり、上に書いてありますけれども、「TUNING-AHELO コンピテンス枠組の共有と水準規定によるグローバル質保証」をテーマといたしました。今回のテーマに興味、関心をいただき、在日の約20か国近い大使館の皆様をはじめといたしまして、400名近くの方々に御出席をいただいております。誠にありがとうございます。

AHELOはOECDが進める卒業直前の大学生を対象とする調査で、大学生が大学教育を通して、どのような知識、技能を習得しているのか、どのような学習成果が達成されたのかということ、世界共通のテストを用いて測定することを目的としております。

平成24年度まで、そうした国際的な学習成果調査が可能かどうかという検証のためのfeasibility studyが行われて参りまして、17か国が参加をして、一般的技能、経済学、工学の3分野で展開をされてきました。我が国は平成20年1月に東京で開催されました「OECD非公式教育大臣会合」で、当時の渡海文部科学大臣からfeasibility studyへの参加の意思を表明されまして、中教審大学分科会の、OECD高等教育における学習成果の評価に関するワーキンググループの議論をいただいた後、工学分野で参加することを決定いたしました。我々国立教育政策研究所といたしましては、OECDの委託を受けた国際コンソーシアムのメンバーといたしまして、テスト問題の開発に取り組みますとともに、文部科学省の委託を受けて、ナショナル・センターとして、テスト実施のための事務局の役割を果たして参りました。

また、チューニングとは、学生が身につけるべき知識や技能、すなわちコンピテンスを専門分野の文脈の中で定義する方法、そしてそのコンピテンス枠組みを大学間で共有しながら、各大学がそれぞれの学位プログラムを構築する方法を指します。チューニングは2000年に欧州で発案されて以来、趣旨に賛同する大学が欧州にとどまらず、南米、ロシア、北米、アフリカ、オーストラリア、カナダ、中央アジア、タイ、中国など、各地域に拡大をして、世界的な展開を見せてきました。

AHELOの経済学分野と工学分野では、このチューニングの方法を用いて、コンピテンス枠組みを定義されまして、問題が作成されております。本シンポジウムのタイトルをTUNING-AHELOとしたゆえんでございます。

さて、このようなチューニングとAHELOの取組は、我が国の高等教育政策にとっても極めて重要な示唆を含んでいるというふうに考えます。日本では御案内のとおりグローバル化する知識基盤社会を牽引する人材を養成するために、大学教育にますます大きな期待が寄せられております。その一方で18歳人口が減少するその中で、大学進学人口が拡大するといったようなことの結果として、大学教育の質保証の問題が顕在化してきております。

日本の高等教育政策は1980年代の末から、多様化、個性化の路線をたどって参りましたが、近年では質保

証の基盤となる共通の枠組みづくりに重点がシフトしてきております。中教審による平成20年のいわゆる「学士力」答申がそれを象徴しておりますし、平成24年のいわゆる「質的転換」答申でも学位プログラムを確立する必要性とか、教学マネジメントの在り方に焦点が当てられております。このような環境の中でコンピテンス枠組みの共有と水準規定のための、具体的な方法としてのチューニング、あるいはAHELOといったようなものは極めて時機を得た重要なテーマというふうに考えております。

さて、今回のシンポジウムでは、国内外から7名の先生方をお招きいたしました。各先生には御多忙の中、本シンポジウムへの御出席を遠路はるばるいただきまして、心から御礼を申し上げます。

まず、基調講演は、長らく大学評価・学位授与機構の機構長として、我が国の高等教育の質保証を文字どおりリードしてこられました木村孟文部科学省顧問から、現行の大学評価制度の課題について、御講演をいただきます。第1部では、始めにオランダのフローニンゲン大学教授であるとともに、欧州単位互換・累積制度、ECTSですとか、欧州高等教育圏資格枠組み、欧州生涯学習資格枠組みといった欧州の高等教育を調和させるための主要なイニシアチブに携わってこられましたローベルト・ワーヘナール教授に、教育改革と学習のグローバルな測定や比較にチューニングコンピテンス枠組みがどのような役割を果たすのかという点について御紹介いただくことになると思います。

続きまして、全米高等教育経営システム研究所の副所長であるとともに、AHELOの技術諮問グループの座長でもありますピーター・ユーウェル博士に、AHELO feasibility studyの調査結果と、技術諮問グループの結論についてお話をいただきます。

さらに、AHELOの工学分野の専門家として、世界共通のテスト問題の開発で中心的役割を果たすとともに、国内12の参加大学の調査実施をリードしていただきました東京工業大学の岸本喜久雄教授から、日本におけるAHELOの取組と今後の展望についてお話をいただきます。

第2部のパネルディスカッションでは、筑波大学の金子元久教授にコーディネーターをお願いいたしました。金子先生はユーウェル博士と同様に、AHELOの技術諮問グループの委員としてfeasibility studyの在り方を御検討いただいた方でございます。

このパネルディスカッションでは、第1部で御登壇をいただいた方に加えまして、AHELO feasibility studyのカナダでの国内実施プロジェクトマネージャーを務められましたオンタリオ高等教育質保証カOUNシルのメアリーキャサリン・レノン上級研究アナリスト。それから、オーストラリアの国内実施プロジェクトマネージャーをお務めになりましたオーストラリア教育研究所のダニエル・エドワーズ主任研究員をパネリストとしてお迎えをいたしましたので、ディスカッションを行っていただきたいと思っております。このディスカッションを始めるに当たりまして、当研究所の先ほど皮切りをやっていただきました深堀聰子総括研究官から、日本、オーストラリア、カナダにおけるAHELO feasibility studyの取組につきまして、調査結果の活用の観点から報告をしていただいて、その後、深堀先生もディスカッションに加わっていただく予定になっております。

終わりになりますけれども、AHELOのfeasibility studyは、昨年12月に終了いたしまして、現在、OECDの方で本調査の実施の有無が検討されている段階でございます。その行方にかかわらず、学位プログラムの体系化を目指す我が国の高等教育政策や、具体的な取組に着手している各大学にとりまして、本シンポジウムTUNING-AHELOから得られる示唆は、極めて大きいのではないかと考えております。

本シンポジウムが、我が国の大学教育の改善に大いに寄与することを期待いたしまして、主催者からの御挨拶とさせていただきます。以上でございます。

第1部 講演

基調講演

木村 孟*

文部科学省顧問

「現行の大学評価制度はどのように導入され、大学によってどのように受け止められているか」

講演1

ローベルト・ワーヘナール*

オランダ フローニンゲン大学人文学部長・同大学院研究科長

「TUNINGコンピテンス枠組：教育改革と学習のグローバルな測定・比較に向けて」

講演2

ピーター・ユーウェル*

全米高等教育経営システム研究所副所長

「AHELOフィージビリティ・スタディ：調査結果と技術諮問グループ(TAG)の結論」

講演3

岸本 喜久雄

東京工業大学 大学院理工学研究科工学系長

「日本におけるAHELOの取り組みと今後の展望」

- *印が付いている講演者は、当日英語で講演を行っており、本原稿は仮訳です。
- 報告書作成に当たり、当日の発言内容に修正を加えていることがあります。
- 所属団体、職名は2013年12月10日現在のものです。

基調講演

現行の大学評価制度は
どのように導入され、
大学によってどのように
受け止められているか

文部科学省顧問
木村 孟



皆さん、こんにちは。国立教育政策研究所主催のこのシンポジウムにおきまして、基調講演の機会を与えていただきましたこと、大変光栄に存じます。講演に先立ち、まず海外から参加していただきました皆様に、心から御礼を申し上げます。遠路はるばるお越しいただきありがとうございました。実は、私、先週はブリュッセルに出張し、3日前の土曜日に日本に帰ってまいりました。まだ、かなりひどいジェットラグに苦しんでおります。海外からの参加者の皆様も、多分同様の状況ではないかと思いますが、よろしく申し上げます。

本日の私のプレゼンテーションは、我が国において、現行の大学評価制度がどのような経過をたどって実施されるに至ったかについて、お話をさせていただきたいと思っております。TUNING-AHELOとは直接の関係はありませんが、チューニングの前の段階として我が国がどのような状況にあるのか、お話をさせていただきます。世界の他の国々と同じように、日本も高等教育、すなわち大学の改革を、ここ20年ほどの間、推し進めております。現行の大学評価制度はその努力の一つの産物であります。私のプレゼンテーションは2部からなっております。最初のパートにおきましては、この新しいスキームがどのようにして生まれてきたのか、大学改革のプロセスの中でどのように推し進められてきたのかについてお話をいたします。2番目のパートにおきましては、現行の大学評価制度が高等教育機関によってどのように受け止められているのかを、アンケート調査の結果に基づいて御紹介したいと思います。このアンケート調査は私が数年前までヘッドを務めておりました、大学評価・学位授与機構が行ったものであります。

我が国には長年大学審議会というものが存在しておりました。大学審議会は、我が国の歴代首相の中でも特に強いリーダーシップを発揮された中曽根首相（当時）によって創設された臨時教育審議会の答申によって設立されたものです。この臨時教育審議会は大変影響力の強い審議会でありまして、我が国の教育に関する審議会としては、最も強い影響力を持った審議会であったと申し上げても良いかと思っております。臨時教育審議会は、我が国の教育が抱えるあらゆる問題を指摘し、それについての議論を行いました。特に、強く高等教育に着目し、「高等教育が我が国の弱点である」という強いメッセージを発しました。当時、我が国は極めて高い経済成長を遂げており、日本人は、全てのことに強い自信を持っておりました。ジャパン・アズ・ナンバーワンという言葉がよく使われていました。このような時代に臨時教育審議会から出された批判的なメッセージは、様々なセクターから多くの批判を受けましたが、私は、臨時教育審議会のこのような批判は、当を得たものであると思っています。文部省（当時）は1987年それを受け、大学改革を加速するため大学審議会を設置いたしました。この審議会は、省庁統合で、中央教育審議会に統合される2000年まで存続しました。大学審議会では大変活発な議論が行われました。15年弱の間に100以上の答申を世に問い、20以上の報告書を発表いたしました。

1988年に、最初の答申が提出されました。その答申では、我が国の大学院制度を弾力化しなければいけないということがうたわれました。それまでは、大学院は質の高い研究者又は学者を生み出すということを使命としてきたのですが、大学審議会はそのミッションを多様化すべきであると提案しました。具体的には、大学院において、産業界又は研究機関以外の政府機関で働けるような専門性の高い人材を育てるということも目指すべきだとしたのです。この点については、かなりの改善が行われましたが、まだ問題は残っています。依然として、大学院の卒業生の多くは、

研究機関で働いています。

1991年には、大変重要な包括的な答申がなされました。非常に包括的なもので、私はこれを、第1ハイライト答申と呼んでいます。様々な重要な提案が含まれています。例えば最初の提案ですが、我が国の高等教育に関する議論の中で初めて、大学レベルの教育の改善を行うべきであると提言いたしました。ここで初めて、大学教育に焦点が当てられたのです。これまで大学のミッションは、第一線の研究活動の遂行でしたが、大学審議会は、大学に対して、大学教育の重要性も認識しなければいけないと提言いたしました。その結果、大学設置基準の大綱化が実施されました。これは日本の大学に非常に大きなインパクト与えました。我が国では、第二次大戦後直後に、カリキュラムの組み方が変更され、二つの教育プログラムが設定されることになりました。一つが一般教育プログラム、すなわちリベラルアーツあるいは教養教育、そして、もう一つが専門教育プログラムです。一般教育は、大学の1年生又は2年生の時点で行われ、その後に専門教育を行うというシステムになりました。しかし、残念ながら、その結果、一般教育と専門教育の間に不連続が生じてしまいました。そのような状況が一般的になったことから、大学審議会は、一般教育と専門教育の区分を廃止すべきであるとの提言をいたしました。大学審議会のこの答申によって、大学は、カリキュラム編成の自由度を与えられたのです。

この答申の中で最も重要な提言は、大学の教育・研究活動に対して、自己点検、評価のシステムを導入すべきであるという提言です。これは非常に重要な提言でした。この時点で初めて、大学評価の重要性が認識され、評価の芽が出てきたと言っても過言ではありません。この時点までの我が国の大学は、学部教育にのみ重点を置いていました。それに対し、大学審議会は、その優先順位をシフトすべきであると提言しました。すなわち、学部教育から大学院教育に重点をシフトしていくべきだとしたのです。このことは、世界の他の国々では既に行われていることでした。

1993年、また別の答申が発表されました。全ての大学は社会に対してその門を広げていかなければいけないとの提言を行いました。すなわち、社会人学生の受入れなどを行うべきであると提案いたしました。それまで、我が国の大学では、社会人学生の受入れは殆ど行われていませんでした。

次に、1998年、また包括的なレポートが出されました。これを私は第2ハイライトと呼んでいます。これにもたくさん重要な提言が含まれています。例えば、学部教育の再構築という提案を行っています。また、大学に対して、積極的にファカルティ・デベロップメントプログラムを開発すべきであるとの提言もしています。もう一つの重要な提言がなされています。日本の大学は、これまで、象牙の塔と言われる存在で、文部科学省ですら、大学のリーダーシップはどうあるべきか、組織運営はどうあるべきかということについて、口出しすることができませんでした。大学審議会は非常に勇気ある提言を行いました。すなわち、大学の運営においては、強いリーダーシップが必要であり、それを土台にした、効率的な組織運営が必須であるとの提言を行ったのです。これは正に正鵠を得た指摘であります。最も重要な提言は、第三者評価システムの導入です。先ほど御紹介しましたように、評価の芽は1991年に出ていますので、その7年後に本格的な第三者評価システム導入が提案されたということです。この提言を受けて、文部省（当時）は非常に迅速に動き、2000年には大学評価・学位授与機構が設置されました。もともとイングランドのCNAに相当する機関として、学位授与機構が存在していましたが、その組織に大学評価の業務が付加されたのです。大学評価・学位授与機構の設置によって、評価、認証などの新しいスキームが実行されることになりました。ここに紹介した二つの協会ですが、この二つの組織も第三者評価機関として、文部科学省の承認を受けています。その一つは、大学基準協会JUAAです。JUAAには、歴史の長い私立大学の多くが加入しています。そして次が、日本高等教育評価機構JIHEEです。比較的新しく設立された私立大学の多くがメンバーとなっています。4年制の大学の評価については、この三つの機関が国から認証を受け、公式な認証評価機関として活動を行っています。各大学はどの組織の評価を受けるかを選択することができます。例えば、A大学は大学基準協会、B大学は日本高等教育評価機構といった具合です。2004年に、初めて認証評価が実施されました。

ここで、これまでのまとめをしたいと思います。高等教育の改革の方向性におけるキーワードはこの三つです。まず多様化、次に個性化です。大学審議会は我が国の大学はそれぞれ個性を持たなければいけないと言い続けてきました。そして、第三者評価です。この第三者評価は、英語ではcertified evaluation and accreditationと呼んでいます。750ほどの我が国の全ての大学は、7年に1度はこれを受けなければいけないという規則になっています。大学評価・学位授与機構では、2011年に7年周期の1サイクル目を終えました。後ほど御説明しますが、このスキームは全体的にはうまくいっていると申し上げて良いと思いますが、一つ問題があります。研究面の評価が義務でないことです。

目下のところは、教育面だけを評価するということになっています。これで本当に競争力の高い大学を作ることができるのかという疑問の声が各方面から挙がっています。このような状況下では、新しいファンディングスキームが必要であると考えられました。すなわち、世界レベルの研究を推進するための競争的資金の導入です。最近はこの種のスキームが多く打ち出されています。例えばCOE（Center of Excellence）形成プログラムですとか、世界トップレベル研究拠点プログラム（WPI）などがその例です。こういった競争的資金の導入によって、研究大学が多くの研究資金を得ることができるようなシステムが、整備されつつあります。やはり教育と研究の双方について評価を行っていくことが必要だと思えます。教育面をきちんと評価しつつ、競争的な大型の資金を導入していくことによって、教育と研究の双方のクオリティを高めていくというのが、我が国の目下の方針です。

それでは、第2部に進ませていただきます。ここでは認証評価の検証結果についてお話をしたいと思えます。検証に関わる調査は、大学評価・学位授与機構が行ったものです。様々な質問が大学並びに評価者に対してなされました。

まず、大学評価・学位授与機構が設定した評価基準、及び観点についてです。これら評価基準、観点が各大学の教育研究活動の質を保証するために適切であったかどうかと聞いています。Y05、これは2005年度、Y11は2011年度を示しております。1サイクル目における結果です。Tは大学評価・学位授与機構の認証評価を受けた大学の総数です。例えば、09年度に関しましては、38校が評価を受けたということです。5段階評価になっております。5は強くそう思う、4はそう思う、3はどちらとも思わない、2はそう思わない、1は全くそう思わないです。ここで一つ注意していただかなくてはいけないことがあります。日本人は、性格上マイナスの評価をしたがらないという点です。そういう意味では、3は中立的な表現になっていますが、日本人の場合にはややマイナス面の評価としてカウントした方が良いということをお知らせください。これを見ていきましょう。全体として、レスポンスは大変よく、大学が満足しているということが分かります。では、次の質問についてはどうでしょうか。ポジティブな答えが出てきています。しかし、この評価基準が大学の教育研究活動について、社会から理解と支持を得るために適切であったかという質問に対しては、そうは思っていないという様子が浮かび上がっています。つまり、これに対しては少し疑問の目を持っているということです。大学評価・学位授与機構は、今後、この評価基準、観点については修正することを考えなければならないと思えます。次は、自己評価についてです。認証評価を受けるには、各大学は自己評価書を提出しなければなりません。自分たちが提出した自己評価書が、自分たちの大学の総合的な状況が広く社会等の理解を得るために、分かりやすいものになっているかということについて、余り自信がない大学が多かったという結果になっています。次の質問は、大学評価・学位授与機構が作成した最終的な評価報告書についてです。これに対する反応は全体的には良いようです。しかしながら、先ほどと同じ社会から理解と支持を得るのに役立つかとの質問に戻りますと、ここでもマイナスの回答がかなり見られます。次に、最終的な報告書を通じて、教育研究活動等に関して新たな視点が得られたか否かという質問に対しては、非常にネガティブな回答が寄せられています。この評価によって、教育研究活動に関して新たな視点を獲得することは非常に難しいようです。結論として、このタイプの評価は、将来、大学がどのような方向に進めばいいかという点に関しては示唆を与え得ないと言ってよいようです。四つ目の質問が、自己評価を行ったことによる効果、影響についてです。これも全体としてはよい結果になっています。しかしながら、自己評価を行うことの重要性が教職員のほとんどに浸透したかという質問に関してはかなりマイナスの回答が多くなっています。大学の教職員に対して現行制度のメリットを理解させる努力が、まだ不十分だということだと思えます。また、教育研究活動等を組織的に運営することの重要性が教職員に浸透したか、機関全体に浸透したかということの質問に対してですが、回答はこの国が抱える課題を如実に表しています。日本の大学においては、これまで学部の独立性が異常に強かった。全学横断的な効率的な大学運営が求められる今、このような異常に強い独立性が排除されなければいけない。しかしこの回答を見ると、依然として、そのような方向に大学が動いておらず、学長、学部長といった大学幹部、あるいは評価に関わっている幹部教員が、苦勞している状況が読み取れます。次が、各教員の教育研究活動等に取り組む意識が向上したかという質問です。これに対してもまた非常にマイナス回答が多い。この最終報告書が機関全体のマネジメントの改善を促進したかという点については、残念ながら、否定的に捉えざるを得ないようです。

ではここで、結論に入りたいと思えます。現行の大学評価制度は全体として、対象校や評価担当者によって肯定的に受け止められているようです。しかしながら、この評価のメリットについて、大学教職員の理解を得るのは容易ではないようです。幹部の方々は現行の評価制度のメリットをととてもよく理解していただいているようですが、一般の

教職員にはまだそこまで理解されていないと言ってよいようです。また、大学の教育研究活動について、学生や社会の理解と支持を得るためには、評価の基準を改定する必要があることも分かりました。そして、将来、大学の教育研究活動を特色のある革新的なものに発展させるためには、将来というのがキーワードになりますが、評価機関各教育機関の一層の努力が必要であるようです。最後のスライドです。1998年に大学審議会は21世紀の大学像と今後の改革方策について、という報告書を提出しました。この中では、大学が競争的環境の中で教育研究活動の個性化を図っていく重要性が強調されています。私は、この個性化というのが、この国における大学改革のキーワードであると思っています。競争的環境の中でそれを目指すということが重要です。残念ながら、我が国においてははまだ競争的環境が存在しないと申し上げて良いと思います。現行の認証評価が、我が国の高等教育セクターに競争的環境を作り出すことに成功したであろうかという問いかけに対しては、否と言わざるを得ません。我が国の大学評価制度は、英国のモデルの導入から始まりましたが、国立大学の法人化の実施により、当初模索したシステムとは、全く異なったものとなってしまいました。英国でも、大学における教育の質の評価等が実施されていますが、研究評価にも多くの努力がなされています。そのため、英国と日本においては、競争的環境という意味では、大きな違いが出ています。我が国が取り組むべき次の課題は、競争的環境を作り出すような評価制度を生み出すことでもあります。これは私の個人的な考え方ですが、専門分野別の評価を導入することが一つの方策かもしれません。これは英国では既に実施されています。研究の質というものをより詳しく評価する必要がありますし、学生の学習成果を評価する必要があります。私が滞在しておりました時に、研究評価の結果が発表され、熱い議論になりました。オックスフォード大学は非常に伝統ある大学ですが、その近くにオックスフォードブルックスという規模の小さな大学があります。この大学は、社会的な順位で言えばオックスフォード大学よりもかなり低いのですが、歴史分野の研究評価において、オックスフォード大学よりも高い評価を得ました。この結果は英国中の新聞で大きく取り上げられ、大きなニュースになりました。私は、こういった評価こそが競争的環境を作り出すのではないかと考えています。我が国においても、そのようなことが起きることを心から望んでおります。御清聴ありがとうございました。以上で発表を終わらせていただきます。

現行の大学評価制度はどのように導入され、
大学によってどのように受け止められているか

文部科学省顧問
木村 孟

1

1. 大学改革に係る提言

- 大学審議会の答申に注目して -

- ・ 臨時教育審議会
 - ・ 「高等教育が我が国の弱点」
 - ・ 昭和62（1987）年 大学審議会の設置
昭和62（1987）年～平成12（2000）年

2

大学審議会 答申の概要

1) 「大学院制度の弾力化について」 昭和63（1988）年12月

- ・ 大学院教育に関する基本的事項
 - ・ 博士課程の目的の多様化
 - ・ 高度な専門的能力を有する人材の養成
(優れた研究者だけでなく)

3

2) 「大学教育の改善について」 平成3（1991）年2月 第1ハイライト

- ・ 大学教育の改善
 - ・ 初めて大学教育に焦点があてられた
- ・ 大学設置基準の大綱化
 - ・ 一般教育と専門教育の区分を廃止
 - ・ 各大学で特色のあるカリキュラム設計が可能となるよう
- ・ 大学の自己点検・評価システムの導入
 - ・ 大学評価の芽

4

- ・ 「学位授与機関の創設について」平成3年2月
 - ・ 生涯学習体系への移行
 - ・ 多様な高等教育機関の発展
- ・ 「大学院の整備充実について」平成3年5月
 - ・ 重点的整備

3) 「夜間に教育を行う博士課程等について」平成5（1993）年9月

- ・ 社会人学生の受け入れ

4) 「大学教員の任期制について」平成8（1996）年10月

- ・ 教員の流動性向上による教育研究の活性化


5

5) 「21世紀の大学像の今後の改革方策について」平成10（1998）年10月 第2ハイライト

- ・ 学部教育の再構築
 - ・ ファカルティ・デベロップメント
- ・ 大学院の教育研究の高度化
 - ・ 高度専門職業人養成に特化した大学院修士課程の設置
(法律分野の専門大学院から)
- ・ 責任ある組織運営体制の確立
 - ・ 学長・学部長のリーダーシップ
 - ・ 学長諮問会議の設置

6


- 多面的評価システムの確立
→第三者評価システム導入を要請
- 学位授与機構からの改組で
大学評価・学位授与機構(NIAD-UE)設置 平成12(2000)年
- 大学基準協会
- 日本高等教育評価機構
- 認証評価制度の導入 平成16(2004)年



7

中間まとめ

- 多様化・個性化・第三者評価
- 認証評価(7年周期)
 - 研究面の評価 - 義務ではない
 - その他(教育、経営面)
 - 各大学の目的・目標に沿った評価
- 多様化・個性化のみで日本の大学は国際的な競争力を向上させることができるのか - 否!
- 競争的資金の導入(COE、世界トップレベル研究拠点プログラム)



8

2. 認証評価の検証結果について (大学評価・学位授与機構)



9

1. 大学評価・学位授与機構が設定した評価基準及び観点について


- 教育研究活動等の質を保証するために適切であった。

	5	4	3	2	1	T
Y 05	0	4	1	0	0	5
Y 06	3	8	0	0	0	11
Y 07	5	29	5	0	0	39
Y 08	11	2	0	0	0	13
Y 09	2	31	4	1	0	38
Y 10	8	18	4	0	0	30
Y 11	0	7	0	0	0	7
T	29	99	14	1	0	143

5~強く思う
3~どちらとも言えない
1~全くそう思わない

- 教育研究活動等の改善を促進するために適切であった。


Y 05	1	3	1	0	0	5
Y 06	2	8	1	0	0	11
Y 07	5	31	3	0	0	39
Y 08	1	11	1	0	0	13
Y 09	4	28	5	1	0	38
Y 10	7	20	3	0	0	30
Y 11	0	7	0	0	0	7
T	20	108	14	1	0	143



10

- 教育研究活動等について社会から理解と支持を得るために適切であった。

	5	4	3	2	1	計
Y 05	1	3	1	0	0	5
Y 06	1	6	4	0	0	11
Y 07	4	19	14	2	0	39
Y 08	1	9	3	0	0	13
Y 09	1	21	15	1	0	38
Y 10	4	15	11	0	0	30
Y 11	1	5	1	0	0	7
T	13	78	49	3	0	143



11


2. 自己評価について

- 大学の総合的な状況が広く社会等の理解を得るために、分かりやすい自己評価書を作成することができた。

	5	4	3	2	1	T
Y 05	2	1	2	0	0	5
Y 06	0	8	3	0	0	11
Y 07	6	23	8	2	0	39
Y 08	11	0	2	0	0	13
Y 09	8	20	9	1	0	38
Y 10	4	17	9	0	0	30
Y 11	2	5	0	0	0	7
T	23	84	33	3	0	143

- 自己評価書の完成度は満足できるものであった。

Y 05	2	2	1	0	0	5
Y 06	0	9	2	0	0	11
Y 07	5	27	6	1	0	39
Y 08	0	11	2	0	0	13
Y 09	5	23	8	2	0	38
Y 10	6	20	4	0	0	30
Y 11	2	5	0	0	0	7
計	20	77	23	3	0	143



12

3. 大学評価・学位授与機構が作成した評価報告書について

●総じて、機構による評価報告書の内容は適切であった。

	5	4	3	2	1	T
Y 05	2	3	0	0	0	5
Y 06	6	5	0	0	0	11
Y 07	6	33	0	0	0	39
Y 08	2	10	1	0	0	13
Y 09	9	25	4	0	0	38
Y 10	11	17	2	0	0	30
Y 11	3	4	0	0	0	7
T	39	97	7	1	0	143

13

13

●教育研究活動等の質を保証するために十分なものであった。

	5	4	3	2	1	T
Y 05	2	3	0	0	0	5
Y 06	4	6	1	0	0	11
Y 07	5	30	4	0	0	39
Y 08	2	11	0	0	0	13
Y 09	6	25	7	0	0	38
Y 10	9	19	2	0	0	30
Y 11	2	5	0	0	0	7
T	30	99	14	0	0	143

●教育研究活動等の改善に役立つものであった。

	5	4	3	2	1	T
Y 05	4	1	0	0	0	5
Y 06	4	6	1	0	0	11
Y 07	6	28	5	0	0	39
Y 08	1	11	1	0	0	13
Y 09	5	26	7	0	0	38
Y 10	7	21	2	0	0	30
Y 11	0	6	1	0	0	7
T	27	99	17	0	0	143

14

14

●教育研究活動等について社会から理解と支持を得ることを支援・促進するものであった。

	5	4	3	2	1	T
Y 05	1	3	1	0	0	5
Y 06	2	5	4	0	0	11
Y 07	5	24	9	1	0	39
Y 08	3	7	2	1	0	13
Y 09	2	21	14	1	0	38
Y 10	8	14	8	0	0	30
Y 11	2	4	1	0	0	7
T	23	78	39	3	0	143

●教育研究活動等に関して新たな視点を得られた。

	5	4	3	2	1	T
H17	2	1	2	0	0	5
H18	1	4	6	0	0	11
H19	2	18	19	0	0	39
H20	1	7	5	0	0	13
H21	3	12	20	3	0	38
H22	7	17	6	0	0	30
H23	1	3	3	0	0	7
計	17	62	61	3	0	143

15

15

4. 自己評価を行ったことによる効果・影響について

●教育研究活動等について全般的に把握することができた。

	5	4	3	2	1	T
Y 05	4	0	1	0	0	5
Y 06	6	4	0	1	0	11
Y 07	10	28	1	0	0	39
Y 08	6	7	0	0	0	13
Y 09	14	22	2	0	0	38
Y 10	16	13	1	0	0	30
Y 11	1	5	1	0	0	7
T	57	79	6	1	0	143

●教育研究活動等の今後の課題を把握することができた。

	5	4	3	2	1	T
Y 05	4	1	0	0	0	5
Y 06	4	5	2	0	0	11
Y 07	3	34	2	0	0	39
Y 08	3	10	0	0	0	13
Y 09	11	23	4	0	0	38
Y 10	14	15	1	0	0	30
Y 11	0	6	1	0	0	7
T	39	94	10	0	0	143

16

16

●自己評価を行うことの重要性が教職員に浸透した。

	5	4	3	2	1	T
Y 05	0	4	1	0	0	5
Y 06	2	2	6	1	0	11
Y 07	0	20	16	3	0	39
Y 08	2	8	3	0	0	13
Y 09	3	10	21	3	1	38
Y 10	1	16	12	1	0	30
Y 11	0	3	4	0	0	7
T	8	63	63	8	1	143

●教育研究活動等を組織的に運営することの重要性が教職員に浸透した。

	5	4	3	2	1	T
Y 05	0	4	1	0	0	5
Y 06	1	5	3	2	0	11
Y 07	1	15	20	3	0	39
Y 08	1	6	6	0	0	13
Y 09	2	11	22	2	1	38
Y 10	1	13	15	1	0	30
Y 11	0	2	5	0	0	7
T	6	56	72	8	1	143

17

17

●各教員の教育研究活動等に取り組む意識が向上した。

	5	4	3	2	1	T
Y 05	0	4	1	0	0	5
Y 06	0	4	6	1	0	11
Y 07	0	19	16	4	0	39
Y 08	0	7	6	0	0	13
Y 09	2	10	22	3	1	38
Y 10	1	14	14	1	0	30
Y 11	0	2	5	0	0	7
T	3	60	70	9	1	143

●マネジメントの改善を促進した。

	5	4	3	2	1	T
Y 05	0	5	0	0	0	5
Y 06	0	9	0	2	0	11
Y 07	1	16	21	1	0	39
Y 08	0	11	2	0	0	13
Y 09	1	21	14	2	0	38
Y 10	3	16	11	0	0	30
Y 11	0	5	2	0	0	7
T	5	83	50	5	0	143

18

18

●教育研究活動等の改善を促進した。

	5	4	3	2	1	T
Y 05	2	1	2	0	0	5
Y 06	1	8	0	2	0	11
Y 07	3	25	11	0	0	39
Y 08	0	12	1	0	0	13
Y 09	7	17	12	2	0	38
Y 10	2	21	7	0	0	30
Y 11	0	5	2	0	0	7
T	15	89	35	2	0	143

●個性的な取組を促進した。

	5	4	3	2	1	T
Y 05	0	4	1	0	0	5
Y 06	1	4	5	1	0	11
Y 07	2	13	22	2	0	39
Y 08	1	8	4	0	0	13
Y 09	3	13	19	3	0	38
Y 10	3	13	14	0	0	30
Y 11	0	4	3	0	0	7
T	10	59	68	6	0	143

19

5. 機構の評価結果を受けたことによる効果・影響について

●教育研究活動等の質が保証される。

	5	4	3	2	1	T
Y 05	1	2	2	0	0	5
Y 06	2	6	3	0	0	11
Y 07	6	21	12	0	0	39
Y 08	1	11	1	0	0	13
Y 09	7	27	3	1	0	38
Y 10	9	17	4	0	0	30
Y 11	0	5	2	0	0	7
T	26	89	27	1	0	143

●学生(今後入学する学生を含む)の理解と支持が得られる。

	5	4	3	2	1	T
Y 05	1	1	3	0	0	5
Y 06	1	3	7	0	0	11
Y 07	1	11	26	1	0	39
Y 08	0	5	7	1	0	13
Y 09	0	12	23	3	0	38
Y 10	3	8	17	2	0	30
Y 11	0	3	4	0	0	7
T	6	43	87	7	0	143

20

●広く社会の理解と支持が得られる。

	5	4	3	2	1	T
Y 05	0	3	2	0	0	5
Y 06	1	4	6	0	0	11
Y 07	2	15	21	1	0	39
Y 08	0	8	3	2	0	13
Y 09	3	15	17	3	0	38
Y 10	5	10	14	1	0	30
Y 11	0	4	3	0	0	7
T	11	59	66	7	0	143

●個性的な取組を促進する。

	5	4	3	2	1	T
Y 05	2	2	1	0	0	5
Y 06	1	4	6	0	0	11
Y 07	1	21	16	1	0	39
Y 08	3	6	4	0	0	13
Y 09	3	13	20	2	0	38
Y 10	3	13	14	0	0	30
Y 11	0	5	2	0	0	7
計	13	64	63	3	0	143

21

まとめ

- 現行の大学評価制度は、全体として、対象校に肯定的に受け止められている。しかしながら、評価のメリットについて、大学教員の理解を得るのは容易ではないようである。
- 大学の教育研究活動について、学生や社会の理解と支持を得るためには、評価の基準を改定する必要がある。
- 大学の教育研究活動を、特色のある革新的なものに発展させるためには、さらなる努力が必要である。

22

大学審議会「21世紀の大学像と今後の改革方策について(答申)」平成10(1998年)では、大学が競争的環境のなかで、教育研究活動の個性化を図っていく重要性が強調された。

現行の大学評価制度は、我が国の高等教育に競争的環境を作り出すことに成功しただろうか。おそらく、そうは言えないだろう。我が国が取り組むべき次の課題は、競争的環境を作り出すような教育評価制度を生み出すことである。英国で取り組まれてきた専門分野別の評価を導入することが一つの方策かもしれない。重要なのは、研究(R)の質と学生の学習成果(T)である。

23

ご清聴ありがとうございました

24

講演 1

TUNINGコンピテンス枠組：
教育改革と学習のグローバル
な測定・比較に向けてフローニンゲン大学人文学部長・同大学院研究科長
ローベルト・ワーヘナール

はじめに

この20年間に、高等教育の領域は抜本的に変化した。グローバル化及び情報通信技術は高等教育の国際化を加速させた。高等教育プログラムに要求される質及び有効性の水準は、もはや地方又は国内のレベルのみで決定されることはなく、今日では国際的にも参照されている。しかし大多数の大学教職員がその認識を共有しているというわけではない。政府高官、高等教育機関の管理運営職者、学部長及び学術的なリーダーと、平均的な大学教員との間には、意見及び視点に著しい差が見られる。世界的には地域間で意見に相違があることも見て取れる。

しかし、好む好まざるにかかわらず、特に各国の名門大学では、教職員、若手研究者、博士課程並びに修士・学士課程の学生の獲得競争は、国内から海外に舞台を移している。教員・研究者並びに学生は、教育機関を比較して、自分にとって最も有利な選択をする。検索エンジン、ポータルサイト、ウェブサイト、そしてFacebook等のソーシャルメディアといった新しい形態のコミュニケーションによって、このプロセスは世界規模のものになっている。その結果として学生の性質が大きく変わり、教育機関も、教授言語を大幅に、若しくは一部でも、変更するようになってきている。教員・研究者は、教育プログラム並びに学習及び指導内容に関して学生の要求水準が上がってきたことを感じている。これは、学生が高等教育機関及び教育の内容、方法、指導者に関する情報により広くかつ容易にアクセス可能であるためだけでなく、さらに、ソーシャルメディアを介してこうした情報が簡単に交換されるためでもある。正しいもののみならず、間違っただイメージ及び評判までもが、非常に簡単に作られたり壊されたりするため、こうした状況は、全ての関係者にとってリスクを伴っている。それは提供される教育の質にとっても同様である。現在、この問題に、以前にも増して注目が集まっている。大学は、特に外からの圧力の結果として、いわゆる「質を問う」文化を発達させている。この点に関する主な課題は、「高品質な教育プログラムとは何か」ということを、何を基準にどの機関の誰が決めるのかということである。

上記の問題や人格の発達以外にも、高等教育プログラムの内容、実施及び提供方法に対する影響を強めている側面がもう一つ存在する。それは社会に対する適切性である。ここでいう適切性とは、市民性と雇用可能性を高めることを意味する。特に、雇用可能性の比重は一層高まっており、学生の実際の関心や能力と、重要性を競い合っている。これは、世界経済危機が大勢に影響を与えた状況では理解できるが、間違っただ選択につながり、結果としてますます多くのドロップアウトにつながる可能性もある。これは単なる経済面を越えて、高くつく結果となるかもしれない。学生が卒業後に適切な水準の職に就く確率は、あるプログラムで他のプログラムよりも高いことがあるかもしれない。しかしながら、このことは就職に不利なプログラムが、卒業生に対して社会への移行に関する義務を負わなくてよいことを意味するのではない。教育は、単に「芸術のための芸術」であることを意図するものではない。また、今日のダイナミックな世界において、学位プログラムも教員自身の学術的プロフィールを反映させたものであることを期待されているのではない。このことは、プログラムの設計とその実践、及び育成される学生のコンピテンス及び期待される学習成果の在り方にとって、重要な示唆を含んでいる。

新たなパラダイムの模索

「欧州教育制度のチューニング (the Tuning Educational Structures in Europe)」プロジェクト (以下チューニング) は、上記の考えを念頭に置き2001年に開始された。1990年末には欧州単位互換制度 (European Credit Transfer and Accumulation System: ECTS) の創設者及び開発者の間にある懸念が強まっていた。それは、域内単一の単位制度では、社会的課題に十分対応することはできないこと、そして学生の国内 (外) の流動性及び欧州内外の高等教育機関間との協力に対する唯一の万能薬でもないことへの懸念であった。そして出された結論は、学習過程の成果をより強調すべきであること、さらに学位プログラム内において、転移可能な一般的な技能にも注意を払うべきであるということだった。さらに、高等教育機関が社会の福祉に貢献する主要な主体としての役割を担っていることを明記する必要があると考えられた。これは、高等教育機関が説明責任を有しており、責任を果たす能力もあり、持続可能な組織であることを示すことを強いられている時代の流れにも合致していた。これは高等教育機関だけでなく提供されている個々の学位プログラムにも適用されなければならない。こうした分析から、いくつかの結論が導かれた。

- ▶ 社会的なニーズに対応できるよう、高等教育の仕組み、プログラム及び資格を大幅に改革すべきである。
- ▶ この過程において、教員・研究者は中心的役割を与えられるべきである。
- ▶ この改革過程では、専門分野・専攻のレベルにおいて、国際的に共有するための参照基準・基準の開発が必要となる。
- ▶ 全ての主要ステークホルダー (すなわち、教員・研究者、学生、卒業生、卒業生の (潜在的) 雇用主、専門団体) が理解できる用語を作り上げる必要がある。
- ▶ (潜在的) 雇用主及び専門団体を含む全てのステークホルダーのうち、特に卒業生は、カリキュラム設計及び教育改善の過程に (間接的に) 参加すべきである。
- ▶ プロフィールの明確化と柔軟性の促進を通して、学位プログラムの多様化を重視すべきである。
- ▶ 改革は、国内的・国際的な流動性と、学習期間 (プログラムの上位段階にアクセスするための資格要件を含む) の承認を容易にするようなものであるべきである。

「コンピテンス」の概念は、この用語が世界中で異なる意味に解釈されることを認識しつつも、あえて、ステークホルダーとのコミュニケーションの手段として選択した。基本的な理念は、過去も現在も変わらず、教育の第一の役割は、学習の過程の結果として、学生・学習者をより有能にするというものである。コンピテンスとは、チューニングの概念によれば、認知的・メタ認知的技能、知識と理解 (洞察)、対人的・知的・実践的技能、そして倫理的価値が有機的に結合したものと理解されるべきものである。つまり、高水準のコンピテンスを有している者は、社会の主導的役割において、自信を持って活動して成功しなくてはならない。コンピテンス以外にも、チューニングはその方法論の一部として学習成果の概念を導入した。チューニングにおける学習成果とは、科目、モジュール、学位プログラムを通して開発されるべきコンピテンスの水準を規定したものである。

チューニングでは、当初から全般的又は一般的コンピテンスと特定の専門分野のコンピテンスを区別してきた。これは特に一般的コンピテンスに対する認識を高めるためのものであった。既に学位プログラムの一部としては指導されていたかもしれないが、当時では、ほとんどの場合、科目に関する資料で明確には述べられていなかった。時間とともに幾分変化したかもしれないが、これはまだ多くの学位プログラムにおいて、重要な留意事項であると思われる。

とはいえ、実施された分析からさらに深遠な結論が引き出された。改革過程には、教授・学習・評価の過程におけるパラダイムシフトが必要であるというものである。開発されるべきコンピテンスという観点から学習過程の成果に注目した結果、教育インプット又は教育内容から、学習成果又は学習のアウトプットへのシフトがもたらされた。言い換えると、これは学習目的の概念に基づく教職員本位のアプローチから、学習成果の概念に基づく学生主導の学習の概念への移行である。この学習成果とは、学習者が学習の過程を修了した時点で習得していることが期待される知識、理解、技能を意味する。そして学習目的とは、教員が取り上げることが意図する資料又はクラスで検討する専門分野に関連した質問を概説したものである。後者のアプローチは、実際には学習過程ではなく教授プロセス、そして教員から学生に対する知識の移行が重視されていることを意味している。学生本位の学習とは、学習者の達成を重視し、一人一人の優先順位に対応し、学生の学習量 (すなわち学習プログラムの期間中に実施可能な作業量) に配慮した学習プログラムの設計を下支えするアプローチやシステムのことである。これは学習の内容、方法、速度及び場所

の選択に学習者が参加することを奨励するものでもある。

こうした観点から、チューニング構想によって二つの主要な行動計画が立てられた。一つは最新かつ高品質で社会的レバンスのある学位プログラムを学士、修士、博士の全てのレベルで開発するための方法論又はアプローチを開発すること、もう一つは学位プログラムの（再）設計、実施、提供及び教育改善を下支えする国際的な参照基準ないし専門分野固有のコンピテンス枠組を開発することである。本講演では、後者について述べる。

TUNINGコンピテンス枠組み

上述のように、学位プログラムはもはやそれ自体（のみ）で開発・提供されるものではなく、最低でも国単位、望ましくは国際的に確立された参照基準に基づいて開発・提案されるべき点を説明した上で、チューニングでは国際的なコンピテンス枠組の開発を主導することになった。ほぼ同時に、英国の高等教育質保証機構が、英国学位水準基準に係る文書の作成に着手した。どちらの場合にも、教員・研究者のグループに専門分野固有の枠組み、すなわちベンチマーク設定の責任を委託するボトムアップ・アプローチが適用された。二つのケースで採用されたアプローチはわずかに異なっていたが、議論と省察の過程に基づく点では一致していた。チューニングでは、主要コンピテンスのリストが策定され、ステークホルダー間の協議の基礎資料として使用された。当時チューニング・プロジェクトに参加していた全ての専門分野・専攻について、30項目を含む共通の一般的コンピテンスのリストと、専門分野固有の主要コンピテンスのリストが作成された（以下を参照）。ステークホルダーとの協議は、2001年から2002年に7分野で実施され、2008年には一般的コンピテンスのリストを微調整・改良して9分野で再度実施された。協議に参加したのは大学関係者、卒業生、雇用主であり、2008年には学生も参加した。

協議はいずれの場合も以下の異なる変数に基づいて行われた。

- 重要度：ステークホルダーの意見による（彼らの専門の仕事のための）コンピテンスの妥当性
- 達成度：当該大学の学位を取得した結果としてのコンピテンスの達成度

これらの二つの変数を評価するために、協議参加者は「1＝あてはまらない、2＝余りあてはまらない、3＝あてはまる、4＝大いにあてはまる」の尺度が採用された。加えて、協議参加者は、

- 一般的コンピテンスと専門分野固有のコンピテンスの重要性の順位付けをすることも求められた。最も重要な五つの一般的コンピテンス、及び五つの専門分野固有のコンピテンスについて、最も重要な場合には5ポイント、第2位は4ポイント、最下位には1ポイントが割り当てられた。重要なコンピテンスとして選択されなかったものについては、0点が与えられた。

協議の成果は非常に有益かつ意味深いものであり、TUNINGコンピテンス枠組を策定する資料として使われた。特に示唆が得られた結果は、一般的コンピテンスが高い水準で獲得されている卒業生を社会は要求していることであった。専門分野についての知識及びそれに対する洞察を得ることに関しては順調であっても、学生は抽象思考、分析・統合技能（協議参加者が最も重要なコンピテンスと考えていた）実践への知識の応用、問題解決、学習能力、書面及び口頭でのコミュニケーション能力に関して、必要と思われる水準まで発達していなかった。加えて、特に雇用主と卒業生の意見に基づくと、学習過程におけるリーダーシップとチームワークのコンピテンスにもっと注目すべきである。さらに、創造性及び企業家精神の発達により多くの注意を払うようにという助言もあった。同じ構成員から成るステークホルダーグループによるこの種の協議は、修正されたリストを利用して（状況、文化、地域、そのリストを作成した地域の大学関係者の意見に基づき）世界各地で繰り返されたが、おおむね同様の結果が得られた。

協議及びコンピテンス・リストは、コンピテンス枠組を開発するための基礎要素のうちの一つにすぎなかった。その他の要素としては、当時も現在も、学術分野の説明が含まれる。これは国ごとに記述され、地域レベルで統合されるものである。例えば欧州地域がその筆頭であり、世界の他の地域でも展開されている。さらに他の要素としては、提供される典型的な学位のマッピング（見取図づくり）、潜在的な雇用可能性分野を含む社会的・専門的ニーズや将来位方向性についての説明が含まれる。この最後の項目からは、国と高等教育機関に違いがあることが明らかになった。この事実からは、特に修士レベルでは、学位プログラムが多様なプロフィールを持つべきであること、さらに地域的に規定される部分が小さくないことが確認された。

専門分野固有のコンピテンス枠組みの構築

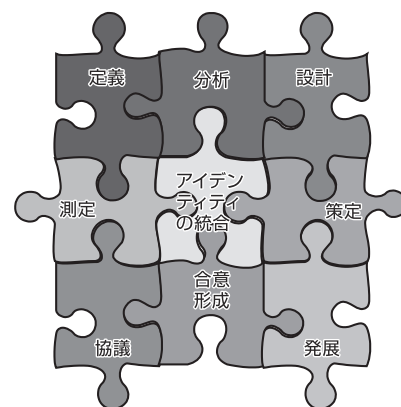
2001年以降、チューニングにより国内及び国際的なコンピテンス枠組みの構築に関して、多くの経験が蓄積された。欧州のチューニング・プロジェクトの一環として、現時点の多数の専門領域をカバーした「[専門分野の名称]における学位プログラムの設計及び実施のための参照基準」と呼ばれる文書が数多く作成されている。これらは、外部の研究者の認証と修正を経て発行されたものである。欧州の文脈では、例えば「参照基準 (reference point)」という用語が「基準 (standard)」の代わりに使われている (ロシア及びオーストラリアでは、この用語が同じ概念を表現するために使用されている)。チューニングは基準の設定を意図したものであるが、これらの参照基準が規範的なものと見なされることは、是非とも避けたいと考えている。今日のダイナミックな世界においては、参照基準を石に刻んで使い続けるわけにもいかないで、できればステークホルダー協議の新ラウンドに基づいて、6年ごとに改訂が必要である。参照基準は、関係する専門分野にかかわらず、共通の書式に基づいて開発されてきた。

開発される各コンピテンス枠組みにつき、12人から15人の国際専門家のグループが形成された。メンバーは様々な国から選ばれ、該当する専門分野について、関係する高等教育機関を代表していた。開始当時、チューニングは、伝統的な研究志向の純粋な専門分野に焦点を当て、方法論の開発に当たって不要な混乱を避けようとした。後になって、より複合的で学際的な学位プログラム並びに応用領域も対象に含まれるようになった。

このプロセスは、上述のマッピング・プロセスとステークホルダーとの協議から始まった。既に述べたように、協議プロセスに備えて、様々なグループの共同作業として、一般的コンピテンスの共通リストが作成されるとともに、各グループの独自の取組として、「専門分野を枠づける」いわゆる専門分野固有の主要コンピテンスのリストが作成された。後者のリストには平均25個の記述が含まれていた。そのリストは、専門分野の学位プログラムに関する自由な議論に基づく一連のアイデア及び予想に基づいて確立された。この長いリストをベースに、協議プロセスに利用される (主要コンピテンスの) 短いリストに関する合意が図られた。専門領域全体の中で育成されるべきコンピテンスに注目するものであるため、第1サイクル (学士) と第2サイクル (修士) の区別はなされなかった。協議プロセスに向けて、各大学ではその分野に該当する雇用主のリスト、この3から5年以内に卒業した卒業生の名簿、及び相談をする大学関係者のリストを作成した。第1回の欧州での協議ラウンド (2001年から2002年) においては、(成人) 学生は協議の対象に含めなかった。Tuning America Latinaプロジェクトは2004年に学生の参加を決定した。以降これが標準的な手続となっている。協議プロセス終了後、その結果は専門分野グループごとに分析された。これにより当初の専門分野固有のコンピテンスのリストは再定義された。このプロセスにおける次のステップは、それぞれの学位サイクルに対する学術的及び職業上の枠組みを設計することである。これは可能な限り記述の形で示され、それぞれについて最も適切なコンピテンスに関するコンセンサスを形成していくものであり、学術的承認を受けるために共通する部分と、専門分野固有の特徴を表す部分を組み合わせたものであった。専門分野の枠組みが当時最新の基準だった欧州高等教育圏資格枠組みに合致するものであるかどうかもチェックされた。欧州では、この枠組みがボローニャ・プロセスの一環として2005年に各国教育大臣によって承認されたのであった。以下では、こうしたいわゆる包括的又はメタ枠組みと専門分野固有の枠組みとの補完的關係については、後述する。

コンピテンス枠組み確立プロセスにおける次のステップは、これらの枠組みが現実的な学生の学習量に相当するかどうかをチェックすることであった。つまりこれは、望ましいコンピテンスの水準を達成するために必要な学習量を「測定」し、(学習成果の他に) 単位として換算できるようにすることである。このステップの後には、特定されたコンピテンスを開発することを目的とした教授・学習・評価の方法やアプローチとして、最も適切なものを見極めるため、自由な議論が行われた。こうした省察は、上述した学生本位の学習にパラダイム転換する提案に呼応して行われたものである。その成果は、各専門分野における様々な可能性や戦略の概要としてまとめ、各高等教育機関で学位プログラムの提供を担当する教員が、その中から選択して組み合わせられるものとなった。最後のステップは、コンピテンスを一貫性のある教育改善の仕組みに関連付けることであった。

このアプローチは、次のイメージのように視覚化することが可能である。



Tuning America Latina2プロジェクト（2011年から2013年）の一環として、コンピテンス枠組みの開発に関する概念はメタ・プロフィールの概念を導入することにより、更に改良され深化した。参照基準及びメタ・プロフィールの目的は、専門分野・専攻の中核（要素）を特定し記述することである。ただし、メタ・プロフィール・アプローチは若干異なる。元のアプローチでは、一般的コンピテンスと専門分野固有のコンピテンスの両方において中核的又は主要コンピテンスを特定することが重視されているが、2番目のアプローチでは、いわゆるメタ・コンピテンスを引き出すために、一般的コンピテンスと専門分野固有のコンピテンスを集積（関連の高い要素から構成されるグループに分類）することに焦点が置かれている。相互に関連する要素から構成されるメタ・コンピテンスの各グループは、専門分野のエッセンスをより一般的な言葉で表現するメタ・プロフィール（コンピテンス枠組み）を規定する基礎となる。メタ・プロフィールが定義されれば、個々の学位プログラムを構築するテンプレートとして利用することが可能である。

この新方式は、ロシア、アフリカ、中央アジアといった世界の各地域で既に展開されているチューニング・プロジェクトにも適用されるようになってきている。コンピテンスの長いリストをまとめ、より本質的なものへと煮詰めていくという既存のテンプレートは、十分な構造を欠いていたため、新方式は我々により洗練された新しい道を提供してくれる。

コンピテンスを分類する上で、主に二つの方法がとられている。1番目の方法は、関連の強いコンピテンスを扱いやすい数のグループ（最低数5，最大数8）に集積する方法である。集積した後に、内容や目的を最大限反映する名称が各グループに与えられる。二つ目の方法は、逆の順番をとり、グループの名称を決定した後に、コンピテンス集積の基礎として利用する方法である。メタ・コンピテンスと呼ばれる各グループは、どれも一般的コンピテンスと専門分野固有のコンピテンスを含むものである。これは、2種類のコンピテンスが同時に開発されることを要求するチューニング哲学に完全に合致している。一般的及び転移可能な技能又はコンピテンスと名づけられているものの、これらは単独ではなく、常に主な専門分野・専攻と関連付けて開発されており、またそうでなければならない。例を用いて説明すると、歴史学における抽象的思考、分析、統合のコンピテンスは、例えば物理学や数学におけるそれらとは異なる理論的、方法論的枠組みに基づいている。これは一般に、口頭発表及び文章表現能力、リーダーシップ、チームワーク、企業家精神といったものにも当てはまる。なぜなら、どの専門分野も独自の学術的文化及びパラダイムを有しているからである。

包括的なコンピテンス枠組み

コンピテンス枠組みは、カリキュラムを新規に設計したり、既存のものを改善したりするための基盤であり参照枠組みであるだけでなく、国内的・国際的文脈の中で学習を測定して比較するための基盤でもあり参照枠組みでもある。コンピテンス枠組みには様々なレベルのものが存在している。チューニングでは専門分野レベルに重点を置いてきたが、政府や質保証機関等では、学習の全て又は一部のレベルで捉えることを意図した包括的又はメタ枠組みを導入している。チューニングによると、これらの相互に補完的な枠組みは、高等教育とその資格の質を改革したり強化したりするための必要条件である。

資格枠組みは新しい現象でないが、近年その内容及び構造は変化している。原則として、どの国にも資格枠組み又はシステムがあり、過去には、あるいは今でも、それぞれの国で提供されている多様な種類の資格、相互の関係性、システム内で上進していく道筋を明らかにしてきた。こうした従来型のモデルは、現在多くの国で、学習過程の成果を重視するモデルに取って代わられている。その基盤として、指標が使われている。これは期待される・達成された学習を、コンピテンスと（その水準を表す）学習成果の形で説明するものである。こうした新型の資格枠組みは、国際レベル（欧州）でも、国内レベルでも開発されてきている。

欧州レベルでは、二つのメタ枠組みが開発されている。一つは、ボローニャ・プロセスの一部として手がけられた欧州高等教育圏資格枠組みであり、もう一つは欧州生涯学習資格枠組みである。ともに欧州各国の高等教育資格枠組みの関係性に透明性をもたらすための包括的枠組みとして定義することが可能である。前者は、いわゆるダブリン指標に基づいている。これらは、チューニングが最初のコンピテンス枠組みを開発した同じ時期（2001年から2005年）に開発された。2002年3月にアムステルダムで開催された「ヨーロッパにおける質的側面へ取組（Working on the

European Dimension of Quality)」と題されたボローニャ・公式セミナーで導かれた主な結論は、学位サイクルのための指標と専門分野レベルの参照基準は統合されるべきだというものだった。発行された報告から引用すると、「学士・修士課程全般の主要な成果を規定する『ダブリン指標』は有益であるという、幅広いコンセンサスが共有されている。これらの一般的指標は、チューニング・プロジェクトにおいて（中略）知識分野（専門分野）のレベルで開発されてきた成果と相互補完的である。」ダブリン指標は、学士、修士、博士というボローニャ・サイクルの完了につながる全ての学位プログラムの本質的な要素を概説している。これは以下の観点に基づいている。

- 知識と理解の獲得
- 知識と理解の応用
- 十分な情報に基づく判断と選択
- 知識と理解の伝達
- 継続的学習能力
- 独創的研究への貢献（博士号のみ）

「ボローニャ」欧州高等教育圏資格枠組みの他に、欧州連合は八つのレベルから構成される、全ての教育レベルの学習達成を捉えた欧州生涯学習資格枠組みも制定した。高等教育に関する限り、この枠組みの上位四つの水準（すなわち5, 6, 7及び8）は欧州高等教育圏資格枠組みに含まれる三つの学位サイクルと一つの短期サイクル（準学士）に相当する。その成果は、若干異なる用語で表現されているが、ダブリン指標で表現されたものに対応している。欧州資格枠組みは、成果ベースの指標を整理するために、知識、技能、コンピテンスという三つの主要なカテゴリを設けている。ここでのコンピテンスは広義で解釈されるべきであり、責任と自律性を意味している。

各専門分野・専攻・職業部門において、欧州高等教育圏資格枠組みのダブリン指標及び（又は）、欧州生涯学習資格枠組みの水準指標は、その部門固有の学習方法に応用して適用することができる。ダブリン指標及び欧州資格枠組み指標は、欧州レベルの一般的参照基準を形成しており、どの学位プログラムでもそこに位置付けることが可能である。

コンピテンスと学習成果に基づく国家資格枠組みの開発は、特にこの5年間に相当な気運を集めてきた。この枠組みを確立するため、各国ではその教育構造及び伝統に従って、独自の資格枠組みを策定している。国家資格枠組みは、国レベル又は教育システムのレベルの個別の説明であるが（スコットランド、イングランド、ウェールズといった）、国際的に理解されており、全ての資格及び他の学習達成が一貫した方法で相互に関連付けられている。優れた事例としては、オーストラリア、南アフリカ、タイ、欧州の多くの国々、そしてルミナ財団によって開発された米国のモデルが挙げられる。国家資格枠組みは、学生、雇用主、質保証・認証評価機関に学生の学習達成を「比較・照合」させ、他の国家・地域の枠組みに対してベンチマークすることを可能にする。

オーストラリア資格枠組みでは、高等教育レベルに関して、知識、技能、知識と技能の応用に加えて、学習目的と学習量（時間換算）の項目を識別している。すなわち、米国ルミナ財団学位資格プロフィールでは、欧州高等教育圏向資格枠組みと同様に、領域に基づいている。すなわち、幅広く統合された知識、専門知識、知的技能、応用学習及び市民学習である。この最後のモデルにおいては、学習過程が分類の原則として強調されている。

これら全ての包括的コンピテンス枠組みの共通点は、一般的で幅広く定義された指標に基づいていることである。これは、あらゆるタイプと志向性（応用、研究志向等）を持つ資格と学習達成を内包しているためである。

メタ・コンピテンス枠組みとTUNING専門分野固有のコンピテンス枠組みの橋渡し

チューニング・マネジメント・チームは、2008年に、学術的プログラムを領域や部門で括って中間レベルの枠組みを開発することによって、メタレベルと専門分野レベルの間のギャップを埋めることが有益であるという結論を出した。部門又は領域とは、おおむね類似した学習プロフィールを持つ関連した専門分野の組合せとして理解される。チューニングでは、人文科学、創作及び舞台芸術、工学、自然科学、保健及び社会科学の6部門に分類している。部門別資格枠組みを開発する基本的な目的は、各部門に含まれる全ての専門分野について、学生に期待される達成レベルに関する共通の記述を作成すること、そしてこれらの専門分野の共通性を定義することであった。この枠組みは、ある専門分野が他部門の専門分野とどのような点で異なっているのかを示すとともに、部門グループとしてどのよう

な特徴を持つのかを明らかにすることを意図している。プロジェクト開始時には、十分な共通点を見つけることができるのか、専門分野の専門家の間で多くの疑念がもたれていた。こうした躊躇は、プロジェクト期間中に、部門別資格枠組みが提供し得る機会に対する熱意へと徐々に変わっていった。

2008年から2010年に、最初のプロジェクトである「社会科学のためのTUNING部門別資格枠組みの開発」が手がけられた。社会科学部門は、経営学、欧州研究、教育科学、作業療法及び社会福祉、法学、心理学及び国際関係論から構成されていた。このプロジェクトでは、高等教育部門レベル5から8、並びにそれに先立つレベル3及び4に適用される部門別資格枠組みを設計した。この開発は、先駆的及び革新的な経験となった。そして、何よりも、欧州高等教育圏資格枠組み、欧州生涯学習資格枠組み、国家資格枠組み及びTUNING専門分野固有の参照基準という、様々な既存の構想を関連づけるための重要な第一歩となった。

社会科学の部門別資格枠組みプロジェクトの後に、人文科学と創作及び舞台芸術の二つの部門別資格枠組みを開発するプロジェクトが続いた。社会科学の部門別資格枠組みプロジェクト・グループが、知識、技能、コンピテンス（広義）という幅広いカテゴリに対応するコンピテンスに関する記述を定義することに留まっていたのに対して、その他の二つの部門別資格枠組みプロジェクト・グループはさらに一歩先に進んだ。コンピテンスに関する記述を整理するための次元（dimension）が、創作及び舞台芸術のグループによって特定されたのである。部門別枠組みを共有することの有効性について相互信頼を生成するきっかけとなったのは、創作及び舞台芸術において、それを構成する全ての専門分野における創造性の次元（制作、演出、設計、概念化）の果たす役割の重要性が認識されたことによる。人文科学でも、人間（人間の状況、経験、表現）の概念にこうした中心的次元を見いだすことができる。「工学において期待される／望ましい学習成果の概念枠組み」（2009年）の開発を目的としたOECD TUNING-AHELOプロジェクトの枠組みにおいても、コンピテンスを整理するための一連の次元が決定された。将来的には、他の3部門についても、次元に基づく部門別資格枠組みを開発する予定である。以下の表は、上述の3部門で特定されている次元の概要を示している。

表1：分野別資格枠組みの次元

人文科学の次元	創作及び舞台芸術の次元	工学の次元
人間	制作、公演、設計、概念化	
文化と社会	人間性に関する再考、熟考、解釈	工学基礎・専門
文書と文脈	実験的、革新的、調査的	
理論と概念	理論、歴史、文化	工学分析
学際性	専門的、環境的、状況依存的な問題	工学デザイン
コミュニケーション	コミュニケーション、コラボレーション、学際的研究	一般的技能
独創性と創造力	独創力と冒険心	工学実践
職能開発	職能開発	職能開発

次元に基づくチューニングの分野別枠組みの開発は、革新的だったと思われる。これは、欧州における二つのメタ枠組みと各国における国家枠組みだけでなく、専門分野レベルにおけるメタ・プロフィール（参照基準）を橋渡しするものである。これまでに開発された三つの部門別枠組みやプロフィールは、学位プログラム設計、実施、質保証及び改善、そして学位及び学習期間の承認のために必要な水準の正確さを提供する。

下の表では、チューニングの部門別資格枠組みのうち、人文科学及び創作及び舞台芸術について、最初の2次元の指標を表している。それぞれの部門に属する学位プログラムの枠組みの中で、何が実際に教授・学習されているのか（又はすべきなのか）に係る情報が示されている。

表2：欧州資格枠組みとTUNING部門別資格枠組みの次元アプローチの関係付け

欧州資格枠組 レベル6	知識	技能	コンピテンス
欧州資格枠組み	専門分野に関する発展的な知識。理論と原理に関する批判的理解を含む。	特定の専門分野における複合的かつ予測不可能な問題を解決するために必要な修習と革新の域に達した発展的的技能。	<ul style="list-style-type: none"> 複合的な技術・専門的活動やプロジェクトを管理し、予測不可能な環境において、責任を持って意思決定を行う。 個人やグループの職能開発の管理を、責任を持って行う。
創作及び舞台芸術の次元（CPD）：制作、公演、設計、概念化	自己の専門分野における制作及び（又は）公演の基礎をなすプロセス及び概念についての発展的知識を有する。	自己の独創的な概念を創出、具現化、表現するために必要な発展的の技能を有する。	学習を通して獲得した知識や技能に基づいて、多様な状況に対して創造的に行動・対応する。
創作及び舞台芸術の次元（CPD）：人間性に関する再考、熟考、解釈	自己の専門分野における実践及び（又は）制作が、我々の人間性から生み出され、それを形成するものであることを理解する。	創造的実践において、人間性について解釈し、省察する技能を発揮する。	学習を通して経験した事柄に基づいて、倫理的認識をもって行動し、他の個人やグループの幸福の増進に取り組む。
人文科学（HUM）：人間	多様な形態や環境における人間の状況、経験、表現について批判的に理解する。	現代の社会問題を理解して解釈するために、専門分野の知識を活用することができる。	個人をその個性、文化、社会的側面から理解し、尊重することができる。
人文科学（HUM）：文化と社会	人間の行動、制度、表現の形態が、思想、信条、価値観から生まれ、相互作用を起こすことを理解し、批判的に洞察する。	社会的・文化的な相互作用に関する重要な問題や課題を、指導助言を得ながら、専門分野の知識に基づいて、特定して定義することができる。	社会における人文科学及び人文主義的な視点の役割を自覚し、社会的結束と持続可能性を達成するために、それを活用するという倫理的使命感を発揮する。

こうした情報は、個々の学位プログラムの設計のために参照されることが望ましい。ここでは「参照」という表現を用いているのは、各学位プログラムがそれぞれ固有のプロフィールを有しており、かつ持つべきであることを明らかにするためである。学位のプロフィールは、コンピテンス枠組みだけでなく、高等教育機関の使命、卒業生が社会で果たすと見込まれる役割、学位を提供する学部の特定の専門領域という観点からの強み、そして利用可能な財政的手段によっても規定される。学生本位の学習の哲学、つまりチューニングの哲学によれば、学位プロフィールは、学位（資格）の提供に関係している大学関係者のチームによって開発されるべきである。このチームには1人以上の成人学生が含まれることが望ましいのであるが、このチームによって、学習成果、学位プログラムの概要と、個々の科目が策定される。こうした構造をとることによって、全ての学習成果を網羅的に開発・達成しつつ、学生の学習量の観点から現実的で実施可能なプログラムを策定することができる。

新たなプログラムの開発及び既存プログラムの改善に資するよう、チューニングでは10ステップ・アプローチを開発した。これは「学位プログラムプロフィール策定のためのチューニング指針（A Tuning Guide to Formulating Degree Programme Profiles）」（2010年）に掲載されている。

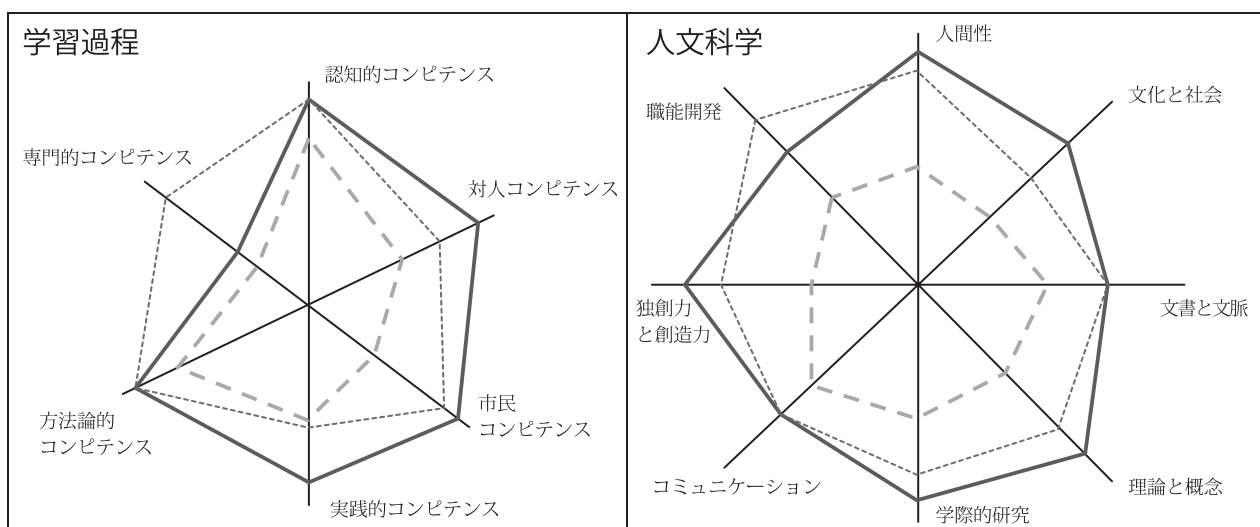
TUNINGコンピテンス枠組み：学習の比較と測定のためのツール

AHELOプロジェクトは、TUNINGコンピテンス枠組みが、特定の専門分野・専攻の学習成果の達成水準を国際的観点から比較することを可能にするアセスメント枠組みの基礎であることを明らかにした。アセスメント枠組みは、個々の学習者の達成度と同時にシステムの達成度を測定するものである。ここでいうシステムとは、高等教育機関中の学部、地方・地域、国家などの様々なレベルの制度を意味する。メタ・コンピテンス枠組みと部門別枠組みを組み合わせることで、TUNINGコンピテンス枠組みは資格や学習期間の承認のための重要なツールにもなる。

世界規模で開発されたTUNING-AHELO概念枠組みを除いて、TUNINGコンピテンス枠組みは、全て欧州、ラテンアメリカ、アフリカ、ロシアといった各地域で定義された。様々なチューニング・プロジェクトを通して、大きく2種類のメタ・プロフィール、又はこの2種類の組合せが開発されてきたことが分かる。一つ目は、学習過程に注目するものである。二つ目は、専門分野又は学術的部門に由来するものである。ともに次元に基づくものである。

「学習過程アプローチ」で定義されてきた次元として、例えば次のものが挙げられる。すなわち、認知的コンピテンス、対人コンピテンス、市民性コンピテンス、実践的コンピテンス、方法論的コンピテンス及び職業的コンピテンス。専門分野アプローチで定義されてきた次元は、上の表に示されている。

いずれのモデルも、個々の学位プログラムを自らのプロフィールに即して開発し、位置付けるために活用することができる。チューニングの専門分野固有のコンピテンス枠組みで合意されてきたコンピテンスの記述に照らして、学位プログラムのプロフィールをレーダーグラフとして視覚化することができる。それぞれのミッション、役割、学術的強みによって、特定される各次元は、メタ・プロフィールで示された水準よりも高かったり低かったりする。このレーダーグラフでは、学士、修士及び博士プログラムのプロフィールを、それぞれ視覚化している。



凡例：緑（破線）=第1サイクル/学士
赤（点線）=修士
青（実線）=博士

先に述べたとおり、チューニングは既に数多くの専門分野において、コンピテンス枠組みを開発してきた。これらの枠組みは全て高等教育の第1及び第2のサイクル、すなわち学士及び修士に対応しており、博士号が含まれることも少なくない。枠組みを開発する際には、複数のレベルに対応することが重要である。これにより、設計者（専門家）は、該当する各コンピテンス水準に対して、開発可能な達成度若しくは開発されるべき達成度を決定することを余儀なくされるからであり、バランスのよい定義が保証されるからである。下の表では、建築の専門分野の事例を示している。建築の専門分野では、部門別資格枠組みのうち、工学ではなく創作及び舞台芸術を参照しながら、コンピテンス枠組みを開発する決断をした。この選択をした理由は、欧州のチューニング専門分野グループの中で建築の専門家は、（工学に典型的な）モデリングの次元よりも創造性の次元に共感したためであった。そして、欧州資格枠組みのレベル6から8、学士号、修士号及び博士号を対象とした指標が開発された。

表3：コンピテンスの習熟度：建築の例

建築			
EQFの分類：	知識	技能	(広義の) コンピテンス
建築物の創作と建築上の創造性			
レベル／次元	建築分野の卒業生に期待される水準。		
6／建築物の概念化，設計，具現化	建築創作を導く過程，概念，文化的価値観について，発展的知識を有する。	自己の建築概念を批判的に考察し，創造的に精緻化し，3次元の形へ革新的に転換する発展的知識を有する。	学習を通して獲得した知識や技能に基づいて，多様な状況に対して批判的・創造的に行動・対応する。
7／建築物の概念化，設計，具現化	建築創作を導く過程，概念，文化的価値観について，分野での第一線の知識を含む，極めて発展的な知識を有する。	自己の建築概念を批判的に考察し，創造的に精緻化し，3次元の形へ革新的に転換するために，自己の建築家としての能力を高度に専門的な水準にまで発達させている。	学習を通して獲得した知識や技能に基づいて，複合的で予測不可能な，新たな戦略的アプローチを有する状況に対して批判的・創造的に行動・対応することのできる，よく発達した人格となる。
8／建築物の概念化，設計，具現化	建築分野の特定の専攻に関連するあらゆる手法と探究の技法について，深い知識を有する。	建築分野において独創的・創造的な洞察を示せるように，過去の経験を統合する。 完全なる創造的自主性を持って働く。	自己の研究能力の他の領域への転移可能性を理解している。 専門的，創造的，学術的な誠実さを示している。

上記から，各レベルを示すために，異なる用語が用いられていることが分かる。上の表が示すとおり，チューニングの経験から，レベルは複合性と範囲によって規定されることが明らかになった。

測定ツールを開発する場合，習熟又は達成の水準について，合意を形成することは極めて重要である。この水準は，その特定のコンピテンスに対する学習成果の形で表される。この学習成果もやや一般的なものとどまっているため，複数（例えば三つ）の指標に具体化する必要がある。これらの指標も，到達の水準を示す複数（例えば五つ）の評価基準に関連付けられる。こうした細分化が，評価・測定される一般的コンピテンス及び専門分野固有のコンピテンスのそれぞれについて行われる。この点に関しても，チューニングは良好な実践例を開発している。一般的コンピテンスに関しては，「コンピテンスに基づく学習：一般的コンピテンスの評価に対する提案 (*Competence-based learning. A proposal for the assessment of generic competences*)」(Aurelio Villa Sanchez & Manuel Pobleto Ruiz, eds. (2008). Bilbao) を参照していただきたい。個別の学位プログラムを開発又は強化する場合にも，本書の参照を勧める。

結論

本講演は，高等教育がグローバル化と情報通信技術といった社会のダイナミズムの挑戦を受けており，こうした状況が，今日の社会に貢献するために卒業生が習得する学習成果の内容と水準を巡る，高等教育提供者間の（グローバルな）競争を引き起こしていると想定してみることから始まった。

状況を分析した結果，高等教育機関が社会のニーズに対応するという責任を果たすべきであることを示している。こうしたニーズとは，とりわけ社会で上手く働くために必要な知識，技能及びコンピテンスを有する，十分な教育を受けた卒業生を指す。そうした卒業生を教育するためには現在使用されている教育モデルのパラダイムの転換が必要

である。このことは実際には教員本位のアプローチから学生本位のアプローチへの移行を意味する。学生本位のアプローチのモデルとは、学位プログラムの設計、実施及び質の改善にステークホルダーのグループが関係していることを意味する。同時にこれは、全ての関係者によく理解できる単一の用語が使われることを意味する。こうした用語は、チューニングにおいて、コンピテンスと学習成果の概念を用いて提案・開発してきた。さらにチューニングでは、学位プログラムの（再）設計の方法論を概説し、部門及び専門分野の両レベルにおいて、コンピテンス枠組みを開発してきた。個々の学位プログラムでは、これらを参照することができる。

これらのチューニングの専門分野固有のコンピテンス枠組み（メタ・プロフィール）のコンピだけでなく、TUNING部門別資格枠組み、欧州資格枠組み、国家資格枠組みといった包括的コンピテンス枠組みが、高等教育及び学位プログラムを改良するための重要なツールであることが本講演で明らかになった。特に、チューニングのメタ・プロフィールは学習（レベル）の比較を可能にする。しがたって、部門・専門分野レベルの学習を、地方、全国、地域、そして世界全体で比較するための測定ツールを開発する際の支柱となる。チューニングで開発された水準指標は、パフォーマンスの評価と学習の比較を公平に行うことを可能にする。

なぜ学位プログラムやシステムの学習成果を、相互に比較することが重要なのだろうか。何よりもまず、比較することによって相対的な長所・短所が明らかになる。透明性と説明責任が問われる今日の世界において、これは重要な情報である。弱点への焦点化を可能にし、洗練した形で学位プログラムの改善を図るための正確なデータを提供してくれる。目標達成の指標が、評価ツールの一部として市民性や雇用のニーズや要件と関連付けられているとき、それらに合わせてプログラムを調整することも可能になる。ただし、チューニングでは、それぞれの高等教育プログラムの内容と形態に対して責任を持ち続けるのは、高等教育機関と大学関係者であるべきだと考えている点は、強調しておく必要がある。その一方で、チューニングの経験が我々に示したように、社会に最善の貢献をすることは、大学関係者の役割であり、大学関係者は全責任を担ってそれを行うことができるし、そうすべきである。本講演で示してきた新しいアプローチによって、学术界は質の高い学位プログラムとは何であるかを示すことができるし、学术界がその役割を良く果たしていることを証明することもできる。

教育改革国際シンポジウム 2013
2013年12月10日 東京

TUNING

TUNINGコンピテンス枠組
教育改革と学習のグローバルな測定・比較に向けて

Robert Wagenaar

rijksuniversiteit groningen


1

TUNINGコンピテンス枠組

TUNING

発表内容

1. グローバルな課題
2. 新たなパラダイムの模索
3. TUNINGコンピテンス枠組
4. 包括的なコンピテンス枠組
5. メタ・コンピテンス枠組とTUNING専門分野別コンピテンス枠組の橋渡し
6. 学習の測定 : パフォーマンスの測定
7. 結論



2

1. グローバルな課題

TUNING

- グローバル化
- 情報通信技術
- 国際化された高等教育市場での競争激化
 - ・ 教員／研究者
 - ・ 学士、修士、博士課程の学生
 - ・ 教授言語
- 労働市場での機会
 - ・ 経済危機の影響

3

2. 新たなパラダイムの模索

TUNING

課題にどう対応するのか (1) :

- 社会的なニーズに対応できるよう、高等教育の仕組、プログラムおよび資格を大幅に改革すべきである。
- このプロセスにおいて、教員・研究者は中心的役割を与えられるべきである。
- 改革プロセスでは、専門分野／専攻レベルで、国際的に共有するための参照点／基準の開発を要件とすべきである。
- 主たるステークホルダー(すなわち、教員・研究者、学生、卒業生、卒業生の(潜在的)雇用主、専門団体)全員にとって理解可能なコミュニケーション言語を開発すべきである。

4

新たなパラダイムの模索

TUNING

課題にどう対応するのか (2) :

- 卒業生をはじめとする全てのステークホルダー((潜在的)雇用主や専門団体を含む)がカリキュラム設計と改善のプロセスに(間接的に)関与すべきである。
- プロフィールの明確化と柔軟性の促進をおとして、学位プログラムの多様化を重視すべきである。
- 改革は、国内的・国際的な流動性と、学習期間(プログラムの上位段階にアクセスするための資格要件を含む)の認識を容易にするようなものであるべきである。

5

新たなパラダイムの模索

TUNING

用語の共有: 「コンピテンス」と「学習成果」

チューニングにおけるコンピテンスの定義

- ・コンピテンスとは、認知的・メタ認知的技能、知識と理解／洞察、対人的・知的・実践的技能、および倫理的価値が有機的に結合したものを意味する。
- ・コンピテンスの涵養は、あらゆる教育プログラムの目的とするところ。
- ・コンピテンスは種々の科目で育成され、さまざまな段階で評価される。

[コンピテンスは学生が獲得するもの]

6

新たなパラダイムの模索

チューニングにおける学習成果とは何か？

コンピテンスの水準は、学習成果（要求・期待される達成度）で示される。

- 学習者が学習を修了した時点で、何を知り、理解し、実行できるようになっていると期待されるかを詳述したもの。
- 単一の科目やモジュールに対応する場合もあるし、第1、2、3サイクルのように、一定の学習期間に対応する場合もある。
- 学習成果により単位取得の必要条件を規定するものである。

[学習成果は教職員によって定義される]

7

新たなパラダイムの模索

高等教育の改革には
教員本位の学習から学生本位の学習への
パラダイム・シフトが必要！

学生本位の学習:
学生の達成度の重視、学習者ごとに異なる優先順位の調整、合理的な学習量(すなわち、学習プログラムの期間内で実施可能な学習量)の設定を旨とする学習プログラムの設計を推進する方法または仕組み。学習の内容、様式、速度および場所の選択に関する学習者の関与を強めるものである。

8

3.TUNINGコンピテンス枠組

学位プログラムの設計と実践に資する参照基準の開発

標準的方法
このとおりでなければならないわけではない

プロセス:
・12~15人の国際的な専門家グループを結成する。
・学問分野について説明する。
・一般的コンピテンスの一覧(30種)を定義する。
・分野固有のコンピテンスの一覧を定義する。
・典型的な学位をマッピングする。
・雇用可能性がありそうな領域をマッピングする。

9

TUNINGコンピテンス枠組

手順:
コンピテンスについて記述する
一般的コンピテンスと分野固有のコンピテンスを区別する
ステークホルダーと協議する: 教員、卒業生、雇用主(および学生)

協議の結果は、**TUNINGコンピテンス枠組**を定義するための情報

10

TUNINGコンピテンス枠組

異なる観点からの協議:

- 重要さの程度: コンピテンスの妥当性に関するステークホルダーの意見(専門的業務を行う上での)
- 達成度: この大学の学位を取得の結果として、このコンピテンスをどの程度達成できたか

これら2変数について、インタビューでは次の尺度を用いた。
1 = あてはまらない、2 = 余りあてはまらない、3 = あてはまる、4 = 大いにあてはまる

次に、

- 一般的技能の順位付け: 教員・研究者、卒業生、学生および雇用主が重要と考える上位5つのコンピテンスに基づく。

11

TUNINGコンピテンス枠組

欧州2001年 欧州2008年

12

TUNINGコンピテンス枠組

欧州2001年 欧州2008年

教員・研究者が重視する上位6コンピテンス

1 基礎的・一般的な知識	1 抽象的思考、分析・統合能力
2 分析及び統合能力	2 実際の状況に知識を応用する能力
3 学習能力	3 専門分野についての知識と理解、職業についての理解
4 創造力	4 問題を識別・設定・解決する能力
5 知識を実践に応用する能力	5 学習能力および学習を継続し最新の知識を維持する能力
6 批判的能力および省察力	6 新たなアイデアを生み出す能力(創造力)

13

TUNINGコンピテンス枠組

欧州2001年 欧州2008年

雇用主が重視する上位6コンピテンス

1 学習能力	1 知識を実践に応用する能力
2 知識を実践に応用する能力	2 抽象的思考、分析・統合能力
3 分析及び統合能力	3 問題を識別・設定・解決する能力
4 問題解決	4 専門分野についての知識と理解、職業についての理解
5 質へのこだわり	5 チームで協働する能力
6 チームワーク	6 計画をたてて時間を管理する能力

14

TUNINGコンピテンス枠組

欧州 アフリカ ラテンアメリカ

全グループ: 上位5コンピテンスの順位付け

共通する2つのコンピテンス:

- 知識を実践に転換する能力
- 概念的思考、分析・統合能力

15

TUNINGコンピテンス枠組

参照基準からメタ・プロフィールへ

- 参照基準のアプローチ: 学問分野の中核的または主要コンピテンスを特定する
- メタ・プロフィールのアプローチ: 一般的コンピテンスおよび分野固有のコンピテンスをクラスター化して、メタ・コンピテンスを導く

グループ化するための2つのアプローチ:

- ラベリングに続いてクラスター化する
- クラスター化に続いてラベリングする

メタ・コンピテンスは一般的コンピテンスと分野固有のコンピテンスの組み合わせだったもの

クラスター数: 最低5つから最高8つまで

16

5. 包括的なコンピテンス枠組

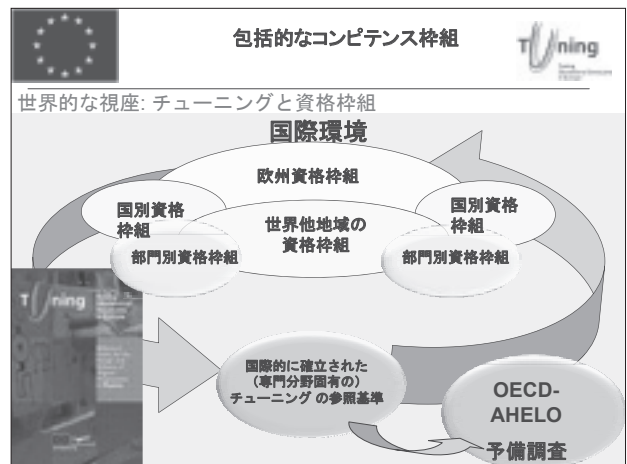
相補的ツール:

- 包括的な資格枠組 (欧州資格枠組 / 欧州高等教育圏資格枠組)
- 国別資格枠組
- TUNING部門別資格枠組: 社会科学、人文科学、創作・舞台芸術
- TUNING専門分野別メタ・プロフィールまたは学際的領域のための参照基準
- ディプロマ・サブリメント (学位プログラムの内容および成果を公表するための文書)

整備中:

- TUNING部門別資格枠組: 他部門

17



18

包括的なコンピテンス枠組

生涯学習のための欧州資格枠組: レベル7 (修士)

知識	- 独創的思考と研究の基礎となる、研究分野における第一線の知識を含む、高度な専門知識。 - 研究分野あるいは異なる研究分野間の学際的な知識問題に対する批判的認識
技能	- 新たな知識と手続きの開発、異なる分野の知識統合を目的とする研究と革新において必要とされる専門的な問題解決技能
(より広範な) コンピテンス	- 複雑かつ予測不可能で新たな戦略的アプローチを必要とするような研究状況をコントロールし変える能力 - 責任をもって専門的知識と実践に寄与する能力および/または責任をもってチームの戦略的パフォーマンスを見直す能力

19

6. 枠組の橋渡し

コンピテンス枠組のピラミッド

20

枠組の橋渡し

21

枠組の橋渡し

次元にもとづくコンピテンスと学習成果の構造化

■ 一般的分類/次元の例:

欧州高等教育調査資格枠組	米連ルミナ(Lumina USA) 学位プロフィール
知識と理解の獲得	幅広い統合的な知識
知識と理解の応用	専門知識
十分な情報に基づき判断と選択を行う	知的技能
知識と理解の伝達	応用学習
継続的学習能力	市民学習
独創的研究への貢献(博士号のみ)	

22

枠組の橋渡し

部門別/分野別の次元の例

人文科学の次元	創作および舞台芸術の次元	工学の次元
人間	制作、公演、デザイン、概念化	
文化と社会	人間性に関する再考、熟考、解釈	工学基礎・専門
文書と文脈	実験的、革新的、調査的	
理論と概念	理論、歴史、文化	工学分析
学際的研究	専門的、環境的、状況依存的な問題	工学デザイン
コミュニケーション	コミュニケーション、コラボレーション、学際的研究	一般的技能
独創性と創造力	独創力と冒険心	工学実践
職能開発	職能開発	職能開発

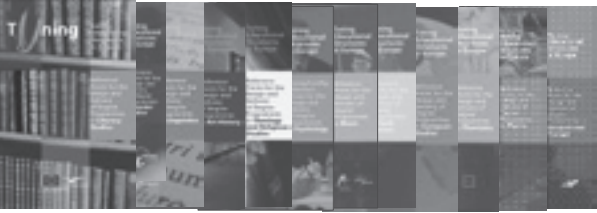
23

枠組の橋渡し

欧州資格枠組レベルにおける部門	知識	技能	コンピテンス
欧州資格枠組	理論と原則に関する批判的理解を含む、研究分野についての上級知識	特定の調査分野における複雑かつ予測できない問題の解決に求められる専門および革新を示す上級の技術	<ul style="list-style-type: none"> 複雑な技術・専門的活動またはプロジェクトを管理し、責任をもって予測不能な研究環境における意思決定を行う。 責任を持って個人またはグループの職能開発管理を行う。
創作および舞台芸術の次元(CPD): 制作、公演、デザイン、概念化	特定分野における製作および/または公演の基礎をなすプロセスおよび概念についての上級知識	独自のクリエイティブな概念を創出、具現化、表現するために必要な上級の技能	研究で得られた知識・技能を用いて、異なる状況において創造的に行動・対応することができる。
創作および舞台芸術の次元(CPD): 人間性の再考、熟考および解釈	自身の専攻分野の成果であるプラクティスおよび/または創作物がいかにして人間性から生じ、人間性を形作るのかを理解する。	創造的実践において解釈技能と人間的側面の反映を示す。	研究で得られた経験を用いて、道徳的認識をもって行動し、成長の促進と他の個人・グループの厚生に取り組む。

24

**TUNING専門分野別
メタ・プロフィールの刊行物**



他分野のパフレット: 経営学、教育科学、ジェンダー研究、歴史学、数学、看護学、など。
最近の出版物: 美術史、言語学、文学および文化、神学および宗教学
経済学および工学のためのチューニング AHELO 概念枠組 (第1サイクル) も刊行されている。

25

枠組の橋渡し

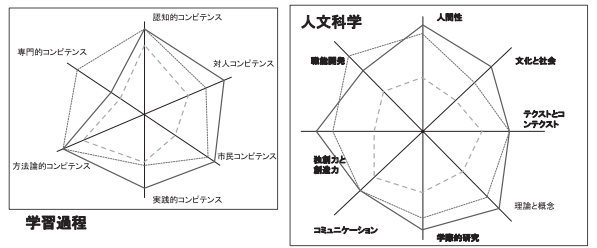
学位プログラムのプロフィール:

- > 独自の特徴
- > 包括的 / メタ枠組に合致
- > 部門別資格枠組に依拠 (利用可能な場合)
- > 分野別 / 領域別のメタ・プロフィール / 参照基準に依拠

26

枠組の橋渡し

プロフィールは学習過程と専門分野という2種類の次元にもとづいて定義することが可能。例:



学習過程
凡例: 学士(破線), 修士(点線), 博士(実線)

27

枠組の橋渡し

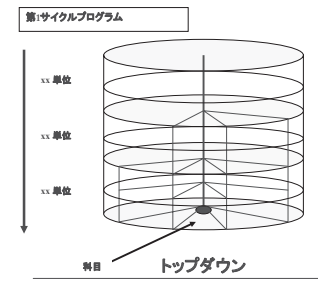
学問的・職業的プロフィールの主な要素

- アイデンティティ** → 学位取得者について「あるべき」とされるアイデンティティの本質を集める
- 機能** → 卒業生が携わる可能性がある職業と業務を割り出す
- 文脈** → 卒業生が有利に仕事をするのできる環境に焦点を当てる
- 教育** → コンピテンス(一般および分野固有)の観点から主に期待される学習成果を規定する

28

枠組の橋渡し

学位プログラムの設計と改善の結果



チューニングの方法に基づく学位プログラム:

- ・プロフィールに基づくプログラム、獲得すべき一連のコンピテンス、達成が望まれる学習成果、学習量に基づく単位授与
- ・プログラムの設計は、協議、議論、協力に基づくチームワークとして行う
- ・開発すべき学習成果 / コンピテンスを単位配分の根拠とする
- ・教授、学習、評価のアプローチは単位配分を尊重している: 実施可能性が重要な要素

29

6. 学習の測定: パフォーマンス測定

学習達成度の比較:

- 地方別: 同一機関のクラス内あるいはクラス間比較
- 国別: 機関間比較 / 外部のレビュー
- 地域別: 国数は限定的、例えば、欧州、東アジア、ラテンアメリカ、北米
- 世界的、例えばAHELO

成果: 成果達成にむけた学習戦略と方法の有効性を検証するうえで信頼できる方法 (自校の学位プログラムの目的とプロフィールを考慮)

30



31

パフォーマンスの測定

コンピテンス達成度を測定する

方法:

- **コンピテンスを規定する**
- 当該コンピテンス**達成度**(習得度)を定義する(一般的な学習成果)
- 各達成度に対する**指標**(進行ルート)を規定する(細分化された学習成果)
- 各指標について、**デスクリプター**あるいは測定基準として1から5段階で示された達成度の記述を**マトリックス**に示す

一般的コンピテンスの例: 批判的思考

32

パフォーマンスの測定

達成度: 批判的思考

- 第1レベルの達成度: 周囲の生活について自身に問いかけ、それについての議論に積極的に参加する。下された判断を分析し、自身および他者の出した結論について熟考する。
- 第2レベルの達成度: 自身および他者の判断の論理を分析し、その個人的・社会的影響を評価する。
- 第3レベルの達成度: 下された判断の妥当性を論議し、根底にある原理と価値に基づき自身の行動の一貫性について分析する。

33

習得レベル	指標	デスクリプター				
		1	2	3	4	5
第1レベルの習得: 周囲の生活について自身に問いかけ、それについての議論に積極的に参加する能力。さらに下された判断を分析し、自身および他者の出した結論について熟考する能力。	批判的精神を示す	自身の生活環境・状況について疑問を持たない。	自身の特定の生活状況について疑問を持つ。	自身の生活状況に対して批判的な態度を示す。	生活について熟考しつづ、リアリティを問い求める。	リアリティについての系統的な熟慮を元に自身の判断を形成する。
	他者の論議のうえで、意見、解釈、評価などから事実を区別する。	意見や評価などに基づく判断または決定を客観的な事実であるかのように自身のものとして受け入れる。	意見や評価などに基づく判断または決定を客観的な事実であるかのように疑問を持つことなく受け入れる。	意見や評価などに基づく判断または決定に疑問を持つ。	意見と評価から客観的事実を区別する。	意見や評価などに基づく判断または決定を正確に分析する。
	積極的に議論に参加する。	議論中も受け身であり続ける。	議論の状況に加わるのは難しいと感じる。	積極的に議論に参加する。	建設的に議論に参加し、厚みのある共通の意見の構築に貢献する。	議論において、他者に対する建設的な参照基準の役割を果たす。
	決定とアプローチの実験的な影響を予見する。	決定と提案の影響に気付かない。	決定と提案の実際的な影響を無視する。	決定と提案の実際的な影響を予見する。	提案された決定の影響に関する賛否両論を分析する。	決定と提案に関する賛否両論を適切に評価することが重要だとする。
自身の決定が他者に与える継続的影響を熟考する。	自身の行動の結果を考えない。	自身の行動に関する他者の所見と批判を受け入れる。	自身の決定が他者に与える継続的影響を認める。	自身の誤りを認め、自己を修正する。	自身の行為に関する他者のフィードバックについて尋ね、考察・考慮する。	

34

パフォーマンスの測定

学問分野別・専攻別コンピテンスに対して同一のアプローチを適用することができる。

- 分野別コンピテンスを規定する
- 次の単元に進むための望ましい習得(達成)レベル(進行ルートまたは経路を規定する)(一般的な学習成果)
- 規定された(各)レベルの達成指標を定める(細分化された学習成果)
- デスクリプター/測定基準を規定する

成果: 専門分野別コンピテンス/学習成果を測定するための枠組


建築分野固有の例: 概念化、設計、具現化

35


パフォーマンスの測定

EOEの分類	建築		
	知識	技能	(より広範な)コンピテンス
建築専攻の卒業生に期待される水準...			
6/建築物の概念化、デザイン、具現化	建築創作につながるプロセス、概念、文化的価値についての知識を有する。	自身の建築上の概念を3次元の形へと批判的に考察し、創造的に精緻化し、革新的に転換する上級技能を有する。	建築研究で得られた知識と技能を用いて、様々な状況に対して批判的・創造的に行動・対応することができる。
7/建築物の概念化、デザイン、具現化	建築創作につながるプロセス、概念、文化的価値について、分野での第一級の知識を含む、極めて上級の知識を有する。	自身の建築上の概念を3次元の形へと批判的に考察し、創造的に精緻化し、革新的に転換する上級技能を有する。	十分に発達したバーナリチアを持ち、建築研究で得られた知識と技能を用いて、複雑かつ予測不可能、新たな戦略的アプローチが求められる状況において、批判的・創造的に行動・対応することができる。
8/建築物の概念化、デザイン、具現化	建築の特定分野に関連した研究に関して、関連するあらゆる手法と技術について深い知識を有する。	これまでの経験を結合し、建築分野における独自の・創造的な洞察を示す。	他の分野への自身の調査能力の可搬性を理解している。 専門的、創造的、学問的な誠実性を示して働く。

36



7. 結論



- 包括的なコンピテンス枠組とチューニングによるメタ・プロフィールは高等教育とその学位プログラムの改革のために極めて重要なツールである。
- チューニングのメタ・プロフィールにより学習(レベルの)比較が可能となる。
- チューニングのメタ・プロフィールは、国内・国際的環境における部門/専門分野別の学習成果の測定ツールの礎となる。
- 学習レベルのデスクリプターと指標により公平なパフォーマンスの測定と学習の比較が可能になる。
- 学位プログラムは現代・将来の世界の課題に対処できるようモデル・チェンジされる必要がある。

37




ご清聴ありがとうございました。



<http://www.unideusto.org/tuningeu/>

38

講演2

AHELOフイージビリティ・スタディ: 調査結果と技術諮問グループ(TAG)の結論

全米高等教育経営システム研究所副所長
ピーター・ユーウェル



この論考の目的は、2009年から2012年にかけて経済協力開発機構（OECD）が実施した、高等教育における学習成果調査（Assessment of Higher Education Learning Outcomes, AHELO）フイージビリティ・スタディ（以下、AHELO-FS）の設計と調査結果について説明し、技術諮問グループ（Technical Advisory Group, TAG）としての見解に基づき、AHELO-FSから得られた知見に関する結論を導くことである。TAGはAHELO-FSへの技術的な指針を提供すること、及びその経験に基づいて今後見込まれているAHELO本調査実施に関する結論を導くことを目的として、調査の初期段階に設置された¹。

AHELOの背景と初期の展開 OECD教育局は、2007年春に高等教育における学習に関する国際的なアセスメント調査の計画に着手した。最初の「専門家会合」は同年春にワシントンDCで開催され、6か国から招へいされたアセスメント専門家及び政策代表者が参加した。会合の協議内容は一般的なものとどまったものの、参加国にとってそうしたアセスメント調査は有意義と思われること、OECDはこうした取組の課題を見極めるために、数箇国を対象とした探索的研究を実施すべきことが合意された。この会合を受けて7月初めにパリで、10月後半に韓国で、より大規模な専門家会合が開催された。

第3回会合が終わるまでに、AHELO構想の基本的な骨子と計画表が作成されていた。すなわち、このアセスメント調査では、複数の専門分野を対象とし、ベンチマークとしての利用を目的としており、個別の高等教育機関のレベルで調査結果を出すものとする。また、OECDが既に実施しているPISA調査と同様に、AHELOも自由回答式の記述式問題をその主たる特色とする。調査全体のうち二つのセクションでは、特定の専門分野の成果を国際的な妥当性をもって検証することを目指す。何度かの協議を経て、候補となる分野として経済学と工学が選ばれた。提案された調査のもう一つのセクションでは、全ての専門分野に共通すると考えられる批判的思考や問題解決能力といった「一般的技能」を直接検証することを目指す。後者が推進された理由の一つは、この分野では、大学生学習調査（Collegiate Learning Assessment, CLA）のような、「有効かつ信頼できる」ツールが既に利用可能だったからである。専門家の間では、一般的技能が個人の成功にとって決定的に重要であることは広く認められているが、同時に、そうした技能の評価は個々の専門分野の文脈の中で適切に行われるべきであると認識されていた。一方、歴史、文学、法律といった多くの分野はあまりにも文化的に規定されているため、国際的なベンチマーキングがほとんど意味を持たないであろうことも認識されていた。最後に、カリキュラムの構造、教育を巡る価値観、基盤にある教育学といった高等教育の置かれた状況が国によって全く異なるという認識のもとに、「背景情報」を重要なセクションとして導入することとなった。これは教職員や高等教育機関の管理運営職者が回答する背景調査、並びにそれぞれの認知的調査（テスト）と同時に実施される学生調査から構成される。

¹本報告書の筆者はTAGの議長を務めた。本報告の内容の大部分は、これまでにTAGが公表してきたフイージビリティ・スタディの報告書に依拠している。

フィージビリティ・スタディの詳細な設計を作成する前に、AHELOの目的と運営方針に関する判断が更に求められた。OECDによると、大学卒業生の流動性が増大する現代において、主たる専門分野について、高等教育の学習成果に国際的整合性をもたらすことが喫緊の課題となっている。これは欧州ボローニャ・プロセスの主な目的であると同時に、複数の専門分野において様々な国の高等教育機関の間で学問的基準を調整する試みである「チューニング・プロジェクト」のような構想に反映されている。個々の機関が参加する理由は、CLAのような全国的な標準テストに参加する場合と同じである。つまり、それによって得られる情報を戦略的計画に役立てること、外部の質保証審査の根拠資料として活用すること、さらに、教育現場で開発している学生の達成度評価の実績のベンチマークとして活用することが可能だからである。

組織体制の面では、AHELOはOECD内部の高等教育機関経営プログラム（Institute for Management in Higher Education IMHE）に設置されている。IMHEは、世界の高等教育機関に対して、戦略的計画やアセスメントといった共通課題に関するサービスを提供することを主たる目的とするユニットである。このユニットの方針は、OCED事務局とともにAHELO構想に関して広範な責任を負う教育政策委員会（Education Policy Committee, EdPC）によって制定されている。AHELOに対するより詳細な指針を提供するため、参加国・地域の代表及び不参加ではあるもののAHELOに関心を表明した「オブザーバー」諸国の代表から構成される各国専門家会合（a Group of National Experts, GNE）が設置された。GNEは、政策・行政代表者や国際的なテストに関する専門家等、多様な委員から構成されている。AHELO-FSは当初、数箇国による2、3年間の取組として構想されていた。この取組を通じて答えを見いだそうとした主な課題は以下のとおりである。a)国によって状況が極めて異なる中で、複数の言語による有効かつ信頼できるテスト問題を開発することは可能なのか、b)そうした取組をOECDの複雑かつ複数の国家が関わる組織体制の下でうまく管理することができるのか、c)テスト実施プロセスは国内コーディネーターや高等教育機関コーディネーターをたてたり、調査運営マニュアルを準備したり、各国で研修プログラムを実施することで効果的に遂行できるのか、d)調査結果データは分析可能であり、参加国と参加機関にとって有益なものとなるだろうか。

AHELO-FSの設計

フィージビリティ・スタディの各参加国は、自国の高等教育システムの多様性を反映する高等教育機関のセット（6～10機関）を選考するよう要請された。次に、各高等教育機関では、3年制又は4年制の高等教育で卒業を控えた学生200人が抽出された。フィージビリティ・スタディの主な目的は、特定のアセスメント方法を試行し、実施についての結論を導くことにあったため、調査結果データが各高等教育機関や各国を比較するデータとして活用されることを意図していたわけではない。また、効率性を最大化するため、アセスメント調査は全て「スパイラル化」され、それぞれの学生がテスト問題全体の異なる部分に解答することとなった。AHELOの参加学生の全員が、それぞれの分野で90分から2時間のテストに解答したほか、学生の背景情報と教育経験の情報収集のために設計された短い調査票にも回答した。アセスメント調査は全て保護されたコンピュータ網を通じてオンライン実施された。記述式問題（CRT）項目は、各国の訓練を受けた採点者によって、特別に設計された採点ルーブリックを用いて採点された。

AHELOの一般的技能分野の核となったのは、ニューヨークに拠点を置く教育支援審議会（Council on Aid to Education, CAE）が運営する自由回答式のテスト問題であるCLAの修正バージョンである。CAEの代表者は、AHELOの計画のための初期の会合に全て出席した。CLAアセスメントは当時、記述式問題を使用する一般的技能テストとしては唯一のものだったため、CAEがAHELO-FS委託事業の主要な受託組織の一つに選ばれた。一般的技能分野への参加を検討している国が招待されたニューヨーク会合の後、改善を図っていくためのCLAの問題が2問選択された。一つは、「湖から川へ」という問題で、安全面の影響が不確かなダム建設計画を進める上での賛否両論を扱う問題である。もう一つは、「ナマズ」という問題で、観察された生態異常が、産業汚染に起因する可能性を扱う問題である。学生は、どちらかの問題一つに解答するとともに、オーストラリア教育研究所（Australian Center for Educational Research, ACER）が開発した卒業生能力調査（Graduate Skills Assessment Test）に基づく一般的コンピテンスに関する多肢選択式問題（MCQ）群にも解答した。

AHELOの経済学アセスメント調査は、教育試験サービス（Educational Testing Service, ETS）によって開発された。それは、経済学分野の既存の多肢選択式テスト問題と、チューニング・プロジェクトの成果や英国高等教育質保証機構が制定した分野別学位水準基標（Subject Benchmarks）といった経済学分野の知識に関する様々な国際的取組の成

果を組み合わせたものである。テストの開発は、専門家グループによって作成されたアセスメント枠組みに沿って行われた。この専門家グループは、テスト問題案が開発されると、その検討にも携わった。経済学分野のテスト問題の焦点は、専門分野の知識を厳密に問うというよりも、学生の「考える」能力、専門分野の概念や用語を効果的に使用する能力にあてられており、記述式問題のセットと大学院進学共通試験（Graduate Record Examination, GRE）の経済学の問題に手を加えた多肢選択式問題群から構成されていた。

工学分野のアセスメント調査は、ACERが主導する複数のアセスメント開発機関による国際的な連携に基づいて開発された。それは各国の既存のコンピテンス枠組み、日本の土木学会の認定土木技術者資格試験及び日本技術士会の技術士第一次試験、欧州チューニング・プロジェクトの成果に基づくものであった。工学には多様な専攻があるため、調査で対象とする専攻を最初に決める必要があった。工学分野の専門家グループは、概念的に最も明快であることを主な理由として、土木工学を選んだ。加えて、分析、設計、それを下支えする基礎的な科学概念といった、全ての専攻に共通するいくつかのより一般的な技能についても調査した。テスト開発はまた、工学専門家によって作成されたアセスメント枠組みにのっとなって行われた。経済学と同様、工学分野のテスト問題の焦点は、単なる知識の習得にとどまらず、学生の「技術者のように考える」能力を測定することにあてられている。テスト問題の設計は、日本の技術者資格試験の影響を大きく受けている。経済学分野と同様に工学分野でも、テスト問題は多肢選択式問題と記述式問題の組合せであった。

フィージビリティ・スタディの当初の計画では、参加国は5, 6箇国のみを想定していた。最終的には全17か国が25分野（一般的技能9か国、経済学分野7か国、工学分野9か国）に参加した。どの国も背景情報調査に参加した。それはアセスメント調査と併せて実施する10分間の学生調査、教職員調査、学部長（経済学部、工学部）調査、機関管理運営職者調査といったものであった。さらに、追加の背景情報として、各国の調査コーディネーターが整理して提供した高等教育システムやカリキュラムに関する記述的な資料も集められた。2012年春に始まったテスト実施は5月末に終了し、採点と分析は2012年夏の終わりまでに完了した。

TAGの設立とその役割 OECD事務局と関係各国は、AHELO構想の開始当初より、AHELOフィージビリティ・スタディの技術的基準の審査と維持に責任を持つ諮問機関の必要性を認識していた。この必要性は、2007年にワシントン、パリ、ソウルで開催された三つの専門家会合において確認された。

TAGは2010年に、アセスメントや政策に係る世界的な専門家8人をメンバーとして、正式に発足した。TAGはOECD事務局とAHELO各国専門家会合（GNE）の双方に報告を行った。TAGの委任事項は幅広く指定されているものの、主な役割としては、a)フィージビリティ・スタディの全側面に関する草案資料を検討し、中間修正の提案を行うこと、b)AHELO本調査を将来的に実施することについて提言を行うこと、c)調査終了時点でAHELOの実施可能性について最終提言を行うことが挙げられている。四つ目の役割は、フィージビリティ・スタディの第二フェーズにおいて、TAGが一般的技能分野及び背景情報調査に関する専門家グループとしての任務を与えられた際に追加された。また、委任事項にはGNEが技術的あるいは「その他の事項」についてTAGに助言に求めることができると規定されていたため、フィージビリティ・スタディが進展するに伴って、取組の方針や実施に係る問題についてTAGが検討する頻度は増加した。

フィージビリティ・スタディの実施期間中にTAGは8回の会合を開いた。そのうち3回は対面式の会合で、残りは電話会議だった。ほとんどの会議の内容は、書類の検討に基づく進捗状況の確認であり、アセスメント枠組み、テスト開発、サンプリングの方法、国家間の調整、テスト実施の手続、記述式問題の採点に係る手順、分析の計画、報告の在り方といった調査のあらゆる側面を含んでいた。フィージビリティ・スタディの中間指針の提言は、こうした検討の過程でTAGによって策定された。各会合後には、TAG座長が報告書案を作成し、各国専門家会合と（GNE）とOECD事務局に送達した。また、各対面会議の後、TAG座長はGNEと対面して課題と提言について報告するとともに、GNEの議長もまた2012年10月のTAG対面式会合にオブザーバーとして出席した。

フィージビリティ・スタディ全体に関するTAGの評価 AHELO-FSは高等教育レベルにおける複数国でのデータ収集という過去に例を見ない取組である。自由回答を重視した記述式問題（CRT）と予め用意された選択肢から選んで答える多肢選択式問題（MCQ）の両方を含むテスト問題を用いて、17の異なる国・地域において、三つの分野で学

習成果に関するデータが収集された。学生、教職員、機関所属のコーディネーター、各国プロジェクトマネージャーが回答する調査からは、様々な背景的要因に関するデータが収集された。翻訳、適正化、サンプリング、テストのオンライン実施、記述式問題の解答の採点、データクリーニング、統計分析、報告といった実施上の数々の課題は、ほぼ全て適切に対処され、うまく解決されている。確かに、一部の国・地域は相対的に多くの困難を経験し、その結果として成功のレベルは一律ではなかった。しかしながら、全ての参加国がその経験から何らかを学習し、意義のある取組であったと報告した。またこれに劣らず重要な点として、AHELO-FSは高等教育レベルでの学生の学習に関して様々な重要な結果を導き、AHELO本調査の実施の在り方について、数々の教訓を与えるものとなった。

もっとも、AHELOのフィージビリティ・スタディには非常にうまくいった点がある一方、それほどうまくいかなかった点もいくつかある。これらの多くは、今後のAHELO本調査への教訓を示している。

うまくいった点 フィージビリティ・スタディの際立った長所としては以下のものがある。

- **アセスメント調査の運営** フィージビリティ・スタディにとって、世界規模での複数言語・地域におけるアセスメント調査のオンライン実施は大きな挑戦となったが、見事に達成することができた。数百もの高等教育機関で実施された膨大な数のテスト・セッションのうち、運営上の重大なミスは一つにとどまった。この達成を可能にした技術的基盤、機関コーディネーターに対する綿密な研修体制、確立された運用手続の遵守が調査実施の成功に寄与したといえよう。
- **データ分析の技術的側面** 多数の異なる回答者に対して六つの異なるテストや調査を実施した結果、フィージビリティ・スタディから得られたデータは大量かつ複雑なものとなった。この問題を前にして、技術的な立場から言えば、信頼に足る分析を提供する取組は模範的なものであった。分析計画は堅実なものであり、採用された統計的技法は適切かつ巧みに進められ、分析上の問題（欠損値や機能不全の項目など）が生じた場合には、適切かつ効果的な「回避策」が採られた。
- **AHELOのために作成されたテストの設計** 工学分野と経済学分野の多肢選択問題と記述式問題、及び三つの背景情報調査といった、フィージビリティ・スタディのために特別に設計されたツールは全て、技術的な質の観点から模範的なものだった。いずれも開発にあたって、適切かつ有益なアセスメント枠組みを参照しており、豊富な知識を有する専門家グループ（工学及び経済学の場合）や、相当な経験の蓄積（背景情報調査の場合）から情報提供を受けたものだった。しかも、これらのテストや調査票は、迅速に用意され、ほとんど手直しを必要としなかった。高い技術水準にかなうように設計され、短期間に期待できる最良の方法でパイロット調査が実施されている。
- **全体の調整** AHELO-FSのように複雑な企画の運営と調整には、5大陸、17の固有の文化的・政治的文脈、そして多数の時差が生じる異なる時間帯において、一貫した手続を保持するという大きな課題があった。AHELO-FS運用に当たって設けられた仕組みによって、これらの課題は、明確な方向性と最小限の混乱の下で対処された。やむを得ない問題に遭遇しても、大半は迅速かつ容易に解決することができた。

余りうまくいかなかった点 同時にAHELO-FSにはそれほどうまくいかなかった側面もある。したがって、これらの点は今後のAHELOの進展において特に検討が必要な領域を示している。

- **財源と時間** AHELO-FSは深刻な財源不足に見舞われ、極めて短いスケジュールで実施された。財源と時間をもっとあれば、次の重要な取組を手掛けることができただろう。例えば、グループ・インタビューや新しく開発したテスト問題を試行するためのパイロット調査を増やしたり、テスト実施や採点の手順を試行するための大規模実査を行ったり、調査結果の説明やそれについてのディスカッションにより多くの時間を割いたりすることができただろう。
- **学生参加率のばらつき** どんなアセスメント調査においても、その妥当性は、各機関においてサンプルとして抽出された学生の十分な参加を得られるかどうかにかんして依拠する部分がある。AHELO-FSでは、国や機関によって学生の参加率に著しい相違が見られた。国によっては（例えば、コロンビアとメキシコ）、ほぼ全ての学生がアセスメント調査に参加した。他の諸国（例えばノルウェーやオランダ）では参加学生の数に余りに少なかったため、得られた結果は分析する妥当性を欠くものであった。今回、学生の参加率にばらつきが生じたことを

考えると、AHELO本調査では必要とされる水準の学生参加を確保するために、通常以上の注意が必要であることを意味している。

- 記述式問題の難しさと適正化 工学分野と経済学分野のアセスメント調査で採用された記述式問題は技術的に高品質であったが、多くの学生にとって難しすぎて、効率的に取り組み好成績を残すことが困難なものだった。同時に、CLAに基づく一般的技能調査で採用された記述式問題は、国際的な場面で使用するには過度に「アメリカ的」であることが判明した。上述のように、テスト問題のパイロット調査や実査により多くの時間を費やすことができたのであれば、より早い段階でこうした状況が明らかになり、実際の調査実施までに修正できたと思われる。
- 契約上の取決め AHELO-FSはOECD事務局と二つの主たる受託組織（ACERとCAE）の間の別個の契約から開始した。個別の契約関係は、受託組織間のコミュニケーションの悪さと作業の重複をもたらす場合があった。しかも、一般的技能分野のテスト問題の調達・開発は、入札手続を経ずに行われた。この事態に、CAEをACERが率いるAHELOコンソーシアム傘下の下請組織となるよう契約上の取決めを再編する形で対処したが、こうした経緯が真のパートナーシップ文化を確立することを難しくした。

AHELO-FSから得た具体的教訓 AHELO-FSの経験から、今後のAHELO本調査に向けて考慮すべき次の教訓が得られた。

- アセスメント調査の設計及び調査結果の分析の過程で、ステークホルダーの参加する機会を拡充する必要がある。フィージビリティ・スタディの多くの局面において、国・機関レベルで実施に携わった人々の知恵を、調査の改善に向けてより巧みに収集・利用できたのではないかと考えられる。テスト問題や運営手順を適正化する取組は秀逸なもので、ほとんどの場合成功したが、より多くのステークホルダーを巻き込む連携のアプローチを採用していれば、更に優れた成果を生んだかもしれない。
- 全てのテスト問題とテスト実施の手順を試行するための大規模実査ができれば、「設計して構築する」プロセスにステークホルダーが参加できるようになる。このことによって、設計を試行するとともに、その改善により多くのステークホルダーが関わるができる。これはとりわけ、国や機関間に調査結果を報告し、データを共有する場合に重要である。高等教育機関の参加者の多くは、調査結果から導かれる情報への自分たちのニーズに注意が払われていないことに、いささか失望していた。学生人口のグループ別分析を含む高等教育機関レベルの報告や、より詳細に説明された機関・国別のデータファイルがあれば有益だったと思われる。プロジェクト全体の品質管理者、さらに、参加した国・地域別の品質管理者も設けるべきである。これらはこうした調査を実施する場合の国際標準に適合するものである。
- AHELO参加国・機関に対して、費用対効果に関するより多くの情報を提供する必要がある。AHELO本調査への参加を検討している国・地域は、次の二つの基本的な質問をされると思われる。すなわち、「負担するコストはどれくらいになりそうか」、そして「何を学べそうか」という問いである。AHELO-FSは終了したばかりであるため、現時点では二つ目の問いについて語れることはほとんどないが、一つ目のコストについてはある程度情報を示すことができる。OECDの契約記録を見れば、それぞれのテスト問題の開発、適正化、実施に掛かる直接的なコストがわかる。サンプリング、学生の募集、テスト実施、採点及び調査結果の報告、調整／監督といった活動に対して高等教育機関や政府で発生したコストの多くについても、同様に知ることができる。だが、例えば高等教育機関や政府の職員がAHELOのプロジェクトに従事した時間など、多くははっきりと把握するのが比較的難しいコストであるため記録されていない。結論として、将来の本調査では、直接的コスト、間接的コストの両方に関するデータ収集について、体系的な取組が必要である。
- 高等教育機関・政府が実際にAHELOを遂行して利益を得ることができるか否かを判断することを可能にする「準備状況に係る基準」のようなツールを開発し、参加を検討している高等教育機関・政府やOECDが利用できるようにすべきである。AHELO本調査における高等教育機関の参加を管理するため、国・高等教育機関の準備状況に関する明確な基準が規定されるべきである。この基準には、サンプリングのための学生母集団に関する情報、十分なコンピュータ設備、コンピュータを使ったテストを支援するIT人員、研修参加へのコミットメント、効果的な内部管理といった規定を含めるべきである。基準には、調査実施の手順に従い、AHELO技術基準を遵守す

るという正式な誓約も含めるべきである。

- 複数の専門分野及び国家／文化的文脈において、一般的技能を効果的に調査する方法について、検討を深める必要がある。AHELO本調査の設計において、独立した分野として一般的技能分野を設けるかどうか、重大な設計上の選択となる。調査対象の専門分野・専攻から独立した一般的技能コンピテンスの存在は、高等教育アセスメントの分野において争点となっている。一般的技能の中には専門分野間での転移が比較的容易なものもあれば、特定の研究分野の文脈の中でよりよく開発、応用、調査されるものもある。フィージビリティ・スタディの一般的技能の記述式問題の結果からは、これらの課題を、文脈に即してよりよく適正化したならば、よりよい成績に結びつくかもしれないことを示唆している。フィージビリティ・スタディで使われた一般的技能に関する二つの記述式問題は、「実社会」での問題解決という文脈に設定されたものだった。しかしながら、調査結果を見ると、こうした課題がどのように文化的に位置付けられ、認識されるかは、国・地域によって著しく異なっていた。今後のAHELO本調査において、一般的技能をいかに文脈に即して適正化するかという問題には更なる検討が求められる。選択肢の一つは、各専門分野別のテスト問題において「専門分野特有の一般的な」要素を含めるといふ方向性を継続することである。これは工学分野のフィージビリティ・スタディで、またある程度は経済学でも採用された方法である。この方向性での更なる発展を目指す場合、各「専門分野特有の一般的な」コンピテン스가共通の内容となるように、一定の整合性を持たせる必要がある。独立した一般的技能を設けるという方法の継続が決定された場合、科学、社会科学、人文科学、芸術といった幅広い専門分野別グループの文脈の中で、記述式問題を設定する方法も考えられる。
- 教授・学習の改善に資する意味のある情報を提供するためには、高等教育レベルの国際的なアセスメント調査において、テスト問題の形式を組み合わせる必要がある。テスト問題の設計に係るもう一つの選択は、AHELO本調査に自由回答式の記述式問題を含めるのかどうかを巡るものである。記述式問題が、多肢選択式問題だけから構成される問題群と比較して、信頼性の観点からは、より優れた結果をもたらすことはない点は、何十年もの研究から明らかにされてきたことである。フィージビリティ・スタディの結果からも、三つの分野の全てにおいて同様の結論が得られている。AHELO本調査で問題となるのは、記述式問題を採用することによって高まる妥当性が、信頼性の低下という代償を埋め合わせることができるほど十分なものかどうかという点である。この点に関して、工学分野のフィージビリティ・スタディの結果は、教授・学習の改善を促進し得る最も重要な情報の多くは、記述式問題から得られることを示唆している。記述式問題をテスト問題に含める場合の主な弱点は、コストの大幅な上昇である。AHELOの主な目的が教育の改善にあるのなら、記述式問題を含めることで調査結果の妥当性は間違いなく向上する。一方、AHELOの主な目的が、学生の学習成果に関する教育機関ごとの成績について、最も信頼できる国際的なベンチマークを提供することにあるのなら、多肢選択式だけのアプローチを採用することによって、信頼性を高めコストを抑えるのが望ましいだろう。
- 高等教育機関間の結果の比較可能性を保証するためには、正確な確率抽出に基づくサンプルと許容できる水準の回答率が必要である。フィージビリティ・スタディでは、約4分の3の参加機関が学生の確率抽出又は全数調査を実施した。残りの機関でそうした方法が採られなかった理由は、明らかでない。AHELO本調査では、参加機関は対象となる学生（又は学生グループ）の一覧をまとめ、確率抽出又は全数調査を実施することが求められる。その一方で、確率抽出法の選択について、ある程度の柔軟性も確保されるべきである。例えば、対象となる学生の数が多い場合、層別の集落抽出が合理的と考えられる。AHELOが国・高等教育機関レベルで回答率の最低値を設定し、その値を下回ったデータは、分析から除外するのも合理的であろう。また、AHELOで新たにデータ収集を行う前に、回答率を高める方法について積極的に調査する必要がある。
- 学生の学習成果や学習改善支援のための政策・実践を検討している国際的な学術共同体の中にAHELOを適切に位置付け、統合するべきである。過去10年間に高等教育における学習実態の改善に対する政策的・学術的な関心は著しく増大した。その証拠を、欧州におけるボローニャ・プロセス及びチューニング、米国におけるスペリング委員会（Spellings Commission）とアクレディテーションに対する関心、多数の国での資格枠組みの策定、U-Map及びU-MultiRankといった世界大学プロフィール化事業の登場に見ることができる。AHELOはこうした高等教育の質に関する新たな学術的・政策的な議論を適切に整理する格好の機会といえる。
- これらのどの課題に対処するにも、より多くの時間と十分な資源が必要である。AHELO-FSは深刻な財源不足

を経験し、それにより実施の過程で多くの側面に悪影響が生じた。財源不足は時を追って深刻になり、いわゆる「フィージビリティ・スタディ」の参加国数を上回る数の諸国が参加したことで影響は複雑になった。今後のAHELO本調査では、財源不足という類似の事態は許されない。OECDと参加諸国は今後の進展に向けて十分な財源を確保する必要があるだろう。この点が保証されない限り、実施は時期尚早といえる。

今後に向けての展開 OECDはフィージビリティ・スタディがAHELO本調査の実施に関する十分に肯定的な結果を得たと結論し、本調査の主な特徴の説明と諸国への参加を呼びかける内容の文書を既に配布している。とはいえ、この判断によってAHELOがより広い範囲で展開されることが保証されるわけではない。今後の展開を左右する真の問題は、こうしたアセスメント調査の実施が可能かどうかではなく、この取組が費用対効果の観点から効果的といえるかどうかである。費用面については現時点で既に明らかである。AHELO-FSの実施費用は900万ユーロを超え、その大半は参加国の政府と高等教育機関が負担した。一方、参加国・機関が得た利益は一律ではなく、データ分析や結果の普及に地域・機関がどれだけの労力を費やしたかに大きく規定されてきた。OECDがAHELO本調査に期待するとおり、参加機関は、国際的ベンチマーク情報を伴うアセスメント調査の結果を、戦略的計画を策定する上で有用と認識するかもしれない。一方、参加国は、高等教育の実績を評価する上で、多様な調査結果をどのように利用することができるかを模索する機会を得ることになる。フィージビリティ・スタディに参加した高等教育機関や参加しなかった高等教育機関や国が、本調査に投資する価値があると信じるようになるかどうかは、今後見極めていかなければならない。

AHELO フィージビリティ・スタディ: 調査結果と技術諮問グループ(TAG)の結論

Peter T. Ewell

全米高等教育経営システム研究所
(National Center for Higher Education
Management Systems, NCHEMS)
国立教育政策研究所教育改革国際シンポジウム
東京(日本)
2013年12月10日

1

AHELOの背景と初期の展開

- PISAとPIACCIにおける同様の取組
- 2007年、専門家会合(ワシントン、パリ、韓国)
- 調査分野: 経済学、工学、一般的技能、背景情報
- 「フィージビリティ・スタディ」を実施することが決定された

2

AHELO フィージビリティ・スタディのデザイン

- 17か国(地域)
- 各国6~10の高等教育機関が参加
- 学位プログラム修了間際の学生、各機関200人を対象
- テスト時間は90分から2時間
- 多肢選択式・記述式問題を組み合わせて実施
- 学生ごとに異なる問題セットに解答:「スパイラル化」

3

AHELO 一般的技能の調査

- 9か国で実施
- 大学生学習調査(CLA)に基づいて作成された記述式問題2問
- 多肢選択式問題は豪州卒業生能力調査から採用
- 各国において適正化および翻訳
- 各国において採点

4

一般的技能とは何か

- 認知的側面:
 - ・ コミュニケーション能力(筆記および口頭)
 - ・ 批判的思考力と問題解決能力
- 非認知的側面:
 - ・ 共感と倫理
 - ・ 他者と協同する
 - ・ 責任を担う

5

卒業生の属性: 雇用者が重視するもの

- 「とても重要」「ある程度重要」と回答した雇用者の比率:
- 倫理的判断力と誠実さ- 96%
 - 多様性に対する寛容さ- 98%
 - 専門的職能開発の余地- 94%
 - 企業サービスを地域社会に還元することへの意欲- 71%

6

一般的技能の教授: 通常のアプローチ

- 「一般教育」科目群
- 学問分野別の教育に「一般的技能」を組み込む
- 特殊な教授法:
 - ・ 協調学習とグループ学習
 - ・ インターンシップとフィールドワーク
 - ・ 卒業研究とプロジェクト学習
 - ・ 「正課外」活動の役割

7

AHELO 経済学分野の調査

- 7か国で実施
- チューニングおよび英国分野別学位水準基標(経済学)に基づいて作成された記述式問題
- ETSが実施するGRE試験(経済学)に基づいて作成された多肢選択式問題
- 学生の「考える」力に焦点を当てた問題
- 各国において採点

8

AHELO 工学分野の調査

- 9か国で実施
- 土木工学に限定
- チューニングに基づいて作成された記述式問題
- 日本の土木学会の認定土木技術者資格試験および日本技術士会の技術士第一次試験に基づいて作成された多肢選択式問題
- 学生の「技術者のように考える」能力に焦点を当てた問題
- 各国において採点

9

AHELO 背景情報調査

- テスト得点における各国間の差異を説明(調整)するための背景情報を収集することを目的として設計
- 背景情報調査の要素
 - ・ 学生(受験者)調査
 - ・ 大学教員調査
 - ・ 機関の背景情報調査
 - ・ 各国に関する背景情報

10

調査結果

- どの分野における調査でも、期待される水準以上の信頼性が確認された
- 記述式問題は多肢選択式問題よりも、信頼性が低かった[予想通り]
- 記述式問題は明らかに学生の知識の独立した側面を捉えていた
- どの分野における調査でも、一定程度の収束的妥当性が確認された

11

技術諮問グループ(TAG)

- AHELOフィージビリティ・スタディの全局面において、活動の適切さを技術的観点から審査する役割を担った:
 - ・ 測定する能力の枠組とテスト問題
 - ・ サンプル抽出とテスト実施の手順
 - ・ 分析と報告
 - ・ AHELO調査の実施可能性、および将来的に本調査を実施するかどうかに係る最終提言

12

TAGの見解: うまくいった点

- 調査の運営
- データ分析の技術的側面
- AHELOのために作成されたテストの設計
- 全体の調整

13

TAGの見解: あまりうまくいかなかった点

- 財源と時間
- 学生の参加率のばらつき
- 記述式問題の適正化と難易度の設定
- 契約上の取決め

14

AHELO本調査に向けた教訓

- 調査設計と分析の過程で、ステークホルダーの参加する機会を拡充する必要がある
- 本調査と同等の規模で、テスト問題と実施の手順を試行する必要がある
- 費用対効果に関する情報を、初期段階から提供する必要がある
- 国や大学が参加の有無を決断する際に参照することのできる「準備状況に係る基準」を開発する必要がある
- 一般的技能の調査については、検討を深める必要がある

15

AHELO本調査に向けた教訓(続き)

- 問題形式を組み合わせる(多肢選択式および記述式問題)方針は維持する
- どの機関においても、サンプルの確率抽出または全数調査を実施する
- AHELOを質に関する学術的・国際的な議論の中に位置づける
- どの課題に対処するにも、より多くの時間と十分な資源が必要

16

今後に向けて

- OECDはAHELOが実施可能であると結論づけ、次の段階に進むことを計画している
- 国や機関が参加するかどうかは、費用対効果に関する個々の判断に委ねられる
- 費用については明らかになっているが、効果については未だ明らかではない
- AHELOが次の段階に進むかどうか、どのような形で進むかは、1年以内に判明するだろう

17

講演3

日本におけるAHELOの 取り組みと今後の展望

東京工業大学 大学院理工学研究科工学系長
岸本 喜久雄



それでは前半最後の講演になりますけども、東工大の岸本の方から、日本におけるAHELOの取組と今後の展望ということで、お話しさせていただきたいと思います。先ほどからのお話のように、AHELOの取組については、我が国は工学分野で参加するということになりまして、いろいろな活動を行ってきました。まず、最初に、取組の概要ということで、工学分野ではどういうことをしてきたのかということをお紹介したいと思います。そのあとに、具体的にどういうテスト問題を作って実施したのかということをお紹介したいと思います。その上で、参加したことの意義がどこにあったのか、これからの展望がどこにあるのかということについて、お話しして参りたいと思います。まず、取組の概要でございますけども、ここに書いてあるのは、先ほどからの説明のとおりですけども、ユーウェル先生が御紹介になったように、大学教育の学習成果を世界共通のテストを用いて測定するということを目指して活動して参りました。最終報告書につきましては、現在、邦訳中ですが、2014年の3月発行予定であります。原文はこちらに書いてあるサイトからダウンロードできますので、御覧になっていただきたいと思います。それで、どういうものを測るかというときに、TUNING-AHELOの考えで、その能力枠組みをまず作るということの作業から始めまして、それを具体的に問題に落とし込むという作業になりました。工学、及び経済学で同じようなやり方を取ったわけでございます。それで、日本の取組の体制ですが、国立教育政策研究所がAHELOナショナル・センターであり、妥当性検証作業の実施組織になりますが、具体的な作業は、本学、東工大が代表を務めさせていただきましたけども、文科省の先端的大学改革推進事業の中で実施して参りました。上側がOECDに関する組織ですが、OECDがAHELOのコンソーシアムに委託して、委託事業として活動しました。我が国は、実際にテストを使って、学生たちに問題を解いてもらって採点するというのもしました。もう一つは、コンソーシアムの中に入って具体的に問題を作り、解答例を作るというような作業に参加しました。ということで、具体的に問題作りから取り組んだということで、AHELOの全貌を我々として経験できたということになりました。

AHELOの取組は、全体で二つのフェーズで行いました。最初のフェーズでは、問題を作って、小規模でテストを実施しました。その中で、具体的に有効な問題、あるいは改善が必要な問題、採点基準等々を見直しまして、そのあとで第2フェーズということで、大規模な調査行って、実際に実施可能かどうかということを検証しました。それを踏まえて、最終会合ということで、今年の3月ですけども国際会議が行われ、先ほどにありましたようなレポートが出たということでございます。

それで、工学分野の取組ですけども、最初にどういう枠組みでテストを作っていくのかということで、2009年の5月の4日から5日かけて、ブリュッセルに専門家が集まって討議をいたしました。その内容がここに書いてあります1から9までということになっております。このときは、どの分野で実施するかということがまだ決まってない段階ですので、工学の全ての分野について議論が行われました。工学分野における学術専門領域にどのようなものがあるとか、典型的な学位はどうだとか、職業等々。正にチューニングのやり方でレポートを作ったということになります。そんな中で、5番目の期待する学習成果の定義ということが大事になります。学習成果というのは何かということについては、最初のワーヘナール先生のところにもありましたが、ここに書かれたようなかたちで定義さ

れるということで、ではこれを具体的にどういうふうを選択していくのかというのが、具体的な課題になるわけです。そのときに工学分野におけるLearning Outcomes, すなわち学習成果としては、このようなかたちで決めるということを行いました。一つは、ABETのEngineering Criteria, これはアメリカの工学分野の認証に使われる基準になりますけども、その枠組みを一つ考えると、これについてはWashington Accordということで、英語圏の国、さらにはアジアの国々が入っているわけですが、日本としてはJABEEがこの枠組みの中に入っておりますので、日本でいうとJABEE, すなわち日本技術者教育機構の基準と対応したものになっています。そういった枠組みと、もう一つはヨーロッパの方ではEUR-ACEということで、ヨーロッパの技術者教育の認定の組織がありまして、そちらが考えているLearning Outcomesがあります。ヨーロッパはFirst Cycle, Second Cycle, Third Cycleになっておりますけども、First CycleのBachelorのDegreeを参照するということになりました。細かく言うと、こちらは3年教育で、Washington Accordの方は4年教育なので、具体的にその水準だとかいったところは、突き詰めて議論するというような問題が出てきますが、ここではfeasibility studyだということで、同者は同等であろうということで、こういった枠組みの中で検討いたしました。これらを基に、AHELOで考えるときのTUNING-AHELOということで、Conceptual Frameworkを考えたということになります。具体的には、こちらの表のようにまとめられるわけですが、一番左にありますのが、EUR-ACE, ヨーロッパの基準になりますが、知識と理解, 工学的分析, 工学デザイン, 調査研究, 工学の基礎, 汎用的能力というかたちに大きくくりにされています。ABETの方は、それに対応させると、例えばa)のところでは、数学, 科学, 工学に関する知識を応用する能力というようなかたちで、カテゴライズされるわけです。それで、今回はできるだけ共通部分を取り出してやろうということから、TUNING-AHELOということで、こちらに書いてあるような枠組みを設定することになりました。ということで、全てのことを取り入れて考えるというよりは、ここでは共通部分を取り出して考えていこうというようなかたちで進めたということでもあります。その中で特に基礎になるのが、基礎科学だとか工学に関する知識, 理解ですけども、そのほか、工学的な分析だとか、工学的なデザインだとか、工学の実践といったようなエンジニアとしての能力がこの中に入っているということでもあります。そういった枠組みを設定したうえで、では問題としてはどんなふうに作っていくかということが、今度はコンソーシアムの方の仕事になるわけです。

それでは、テスト問題作成に際してどのように具体的化したかということですが、こういうかたちでいたしました。先ほどのこちらの枠組みを、少し順番並べ替えてやっておりますけども、一番下支えするのが、工学分野における一般的技能ということで、分野にとらわれず、習得すべき技能ということになります。そのうえで、工学的な基礎能力、さらには専門的な能力ということで、今回、土木分野でやろうということになったわけですが、その中では、今度は土木分野におけるそれぞれの能力というのが、ここに記載されている項目になります。それを下支えとして、エンジニアは実際の問題を分析し、デザインし、実践をしてもつくりをしていくことになります。このことを踏まえてどういうふうの問題を作るかということですが、こちらの項目について、すなわち、基礎的な知識, 技能については多肢選択式の問題で測ろうということになりました。それに対して、実践的な活動に関係するところについては、記述式の問題ということで考えていくということになりました。具体的には多肢選択式の問題については、30問の中から20問を解いてもらい、記述式の問題については3問の中から1問を解いてもらうということで、問題セットの中から学生たちは与えられた問題を解いていくというスタイルで実施いたしました。

では、具体的にはどのような内容なのかというものについてお話ししたいと思います。まず、概念枠組みに基づく問題の作成ということになりますけども、コンソーシアムが原案を作成いたしました。先ほども紹介がありましたけども、多肢選択式の問題は日本側が提案するということにいたしました。一つは、技術士の1次試験の問題を提供していただいて、それを英訳して提案したりしました。もう一つは、土木学会の中で、学生が卒業するレベルでどの程度習得しているかということ、認定試験というかたちでされていらっしゃるんですけども、その問題を提供していただきました。コンソーシアムでは、その中からAHELOのfeasibility studyにふさわしい問題を抽出して実際に使っていくということを行いました。一方で、記述式の問題はオーストラリアのチームが提案をし、その中からふさわしい問題を選んで実際に使いました。こういった提案を受けて、国際専門委員会の会合で、提案された問題の中の取捨選択であるとか、採択された問題について国際的な通用性を高める方向で問題を改善することをいたしました。具体的にどんなことがあるかと申しますと、例えば技術士の試験や土木技術者の認定資格の問題の中には、我が国の特有の技術的な問題が含まれています。そういった問題については、国際的通用性の観点からはふさわしくないだろうと

ということで、取り除くという作業をいたしました。また、ほかの国では余り教えてないような課題については、そういった問題については除くということを行いました。例えば、地震に対する対策ですが、東日本大震災のあとであれば、我々はかなり強く主張できたんですが、この作業が行われたときはそれ以前のことでしたので、我々の国では地震は関係ないと言われ、その問題は取り除くというようなことをいたしました。また、記述式についてもいろんな問題が提案されましたが、例えば、鉄道の問題は非常にいい問題だったのですが、全ての国に鉄道が重要であるわけではないことから、そういったものから取り除くということで、なかなか実際に問題を作っていくという作業についても難しさがあるというふうに感じました。具体的な問題ですが、多肢選択式の問題では、例えばこういった内容の問題が出題されました。このようなトラスについて、ここに力がかかると、この部材にはどんな力が作用するかというような問題です。これは構造力学の基本的な問題になるわけですが、このABCDの中から正解を選択していくということになります。ちなみに、技術士の試験は5択になっています。五つ解答肢がある場合と四つの場合では難易度が少し変わってくるのですが、今回は4択とすることになりました。こういった問題については、我が国においては多くの経験がありますので、かなりの部分、基礎的な能力があるかどうかというのは測れるという感じがしております。

記述式の問題については、例題はこのような内容です。上の方にホームページのURLが掲載してありますので、こちらからダウンロードしていただくと原文もお読みになっていただくことができます。フーバーダムというダムがありますけども、このダムに関する問題です。このような写真と、建設前の候補地になったところでの写真、あるいはどうかたちで貯水池を作るかというようなスケッチだとか、発電所の建設計画だとかといったような資料を合わせて提示いたしまして、それに対してここにあるような設問を設定しています。設問は、ここに示されている四つだけではなく、もう少し数が多いですけども、例えば1番目の問題は、この場所がこのダムを作るのに適している理由を説明しなさいというものです。回答は少なくとも、二つの側面について述べなさいとなっています。実際には、答えは多くあるわけですけども、その中から採点官は正答なのか、これは三角ぐらいの答えなのか、答えになっていないのかということ、共通の採点基準に基づいて採点をするというようなことをいたします。2番目は、フーバーダムの構造上の強度と、安定性を高めている設計上の主な特徴を二つ挙げて説明しなさい、というような問題であります。ほかの設問としては、「このダムを作っている途中で不具合が見つかりました。あなたならどういうふうに対応しますか」、というようなものがあります。このようなことを問うことによって、実際に技術者になったときにどう振る舞うかということ、きちんと学んでいるかどうかを問うことになっております。さらに、このような設問を少し詳しく見ていきますと、この1問目の方の例題ですけども、測定しようとしている学生成果というのは、このダム式水力発電に適した場所の特徴を理解しているかということですけども、きちんと答えが出てくるためには、専門的な知識もあるだろうし、分析する能力というのが必要になるということでもあります。答えの観点ってというのはここにありますように、一つではありませぬので、AからFまでありまして、これに関するような答えが出ていると1個答えが書いてあれば1点、二つ書いてあれば2点、それ以上書いてあっても2点にしかならないことになります。2番目の設問については、フーバーダムの構造上の強度と安定性に寄与する要因を理解しているかということで、これはダムを作るときに工学的なデザインに関して、どういうことをきちんと身に付けて考えることができるかということで、アーチ構造をしているというようなことだとか、周りの谷が岸壁である、堅固であることとか、といったことが正答になります。先ほどの一つ前の設問に関して言うのを忘れましたけども、この回答としては社会的影響が低く抑えられていることがあります。これを具体的に言うと、この周りに人が住んでないことがあります。当初はこの設問の正答にはこのような内容のものが入っていませんでした。1回目の試験でやってみると、日本の学生が、たくさんそのような答えを書いてきたので、これはやはり正解にしてもらわなければ困るだろうなという、正解にするということで、実際に問題を作って何が正解になるかというようなことも、専門家同士で話してみるというようなことが、実際問題としては必要だということがこの例としてあります。そんなような問題に対して、学生たちはどのような反応を示したのかということですけども、多肢選択式の問題については、「ほとんどが大学で教わっている問題であり、解けてしかるべきだと思う。問題が全部単発で、広く浅くという印象。もっと深く掘り下げた方がよいのでは」というような感想でした。ただ、問題としては平易であるという回答ではありましたが、実際には、学生の成績がよかったかということ、必ずしもそうではなく、印象と成績が違うというようございませぬ。次に、記述式の問題ですが、これについては、第1フェーズでは、紙に答えを書いてもらう方法で実施いたしました。このときは、学生た

ちは非常にたくさん文章を書いてくれました。しかし、第2フェーズのときはコンピュータでやらせることになったので、非常にそっけない答えが大分出てきました。すなわち、手書きで回答を書かせるのかコンピュータでやるのかによって反応が違ってくるというのは、また課題があるように思います。第1フェーズの際に、記述式の問題についての学生たちの反応ですが、彼らが言ったのは、「大学では理論的、抽象的な問題が多い中、現実問題を扱っている点が非常に面白い。何が原因だったかを考え、どうするかという流れだったので面白い問題だし、こういう問題を解くことによって自分たちの実務に役立つような技術が上がってくるのではないか」というようなことです。また、「こういった問題が重要であるならば、もっとグループワークやケーススタディの授業が必要である。」とか、「こういった問題に対する解くアプローチを学ぶ機会がなかったのは、是非そういうようなことを授業でやってほしい」というような、学生たちからの非常にポジティブな反応がありました。ということで、feasibility studyをやった結果もそうですが、学生たちがこういうような反応を示したことは、これからの教育のことを考えるうえで大事であると思います。

次に、AHELOに参加したことに対する意義について、簡単にまとめておきたいと思います。一つは、最初の枠組みのところの議論から参加できましたので、学習成果の枠組みの構築に参加できたことです。我が国では日本技術者教育認定機構というのがあって、このようなことについて随分議論してきているわけですが、そういったことが実際に世界的な標準として使われていることを経験することができ、さらには、基準や教育の実態と整合的なものとなるように、我々からも意思表示ができたということでもあります。もう一つは、問題作成のプロセスから参加することで、我が国の工学教育の視点から、我々はどういうふうの問題を作っているというふうなことで、国際的に積極的に情報発信することができたというふうに思います。それで、我々の観点がこうであるということについては、コンソーシアムの仲間、あるいは専門家会議でもかなり尊重して意見を聞いていただくことができました。それは恐らく、我が国の科学技術が世界の人たちからリスペクトされていて、それを我々の卒業生であるエンジニアが支えているということから、工学教育に対する世界からの我が国への関心も高いということを感じました。そういう意味で、国際的な貢献を、ますますこういった分野でも我々はしていかなければいけないのではないかなというふうな考えた次第であります。もう一つは、この取組をするに当たって、工学の仲間だけではなくて、教育学の専門家の皆さんとも一緒に仕事ができたとということで、非常に我々としても勉強になったということから、やはりこういう国際活動をきちんとやっていくためには、特定の分野の工学なら工学ではなくて、やはり教育学の専門家だとか、そういう方々と一緒に仕事をするのが非常に大切に感じた次第であります。そういったことから、我々は工学分野でAHELOに参加したわけですが、こういったところで得られた内容を丁寧にいろいろな、今日の会合もそうですが、場面で説明していく必要があるというふうな考えております。

次に、このような経験から得られた知見についてですが、多肢選択式については、日本が提案した問題の国際的通用性が確認されました。というのは、ある対象分野の設問が少ないからということで補充した問題も検討してみましたが、それは結局、最終的には適切な設問になっておらず使えませんでした。我々が長年かけて作っている問題が、基礎能力を測る上でよく練られている問題だということでもあります。一方で、記述式の問題については、作成する難しさがまだまだありまして、採点も非常に難しいということを感じました。また、問題を日本語に訳すときに、日本語と英語との枠組みが違うので、逐語訳をするのか、内容が伝わるように訳すのかどうかというのは非常に難しかったです。例えば、日本語は主語を余り言わないが、言わないと先ほどの倫理の問題ですけども、倫理上問題があったときに、要するにダムに不具合があったときどうしますかといったときに、他人ごとのような答えがたくさん出てくるのですよね。やはり、技術者としての立場から答えなければいけないことをきちんと明文化して問題にしないと、そのような答えが出てしまうのではないかなということでもありました。いずれにしても、記述式の問題というのは、正答率からいうと多肢選択式の問題の成績がいい人と、ある程度相関があるようですけども、こういった問題は作り続けていくことが非常に大切だなというふうに思いましたし、こういったことが大学の教育にも影響を与えていくというふうな考えることができると思います。

今後の展望については、このようにまとめました。一つめは潜在的なインパクトということですが、繰り返しくなりませんが、世界的な流れの中で、日本の工学教育がおかれている状況について再確認できたということでもあります。AHELOの教育改善の効果を最大限に生かすということが、これから必要になるだろうと思いますので、今回作成した問題を公開し、採点の方法を公開していくことで、学修達成度を測る測り方がもっともっと改善されていくのでは

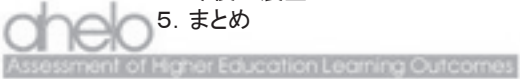
ないかと考えられるかと思います。そういう意味で、今後の展望になりますけれども、一番上は繰り返しになりますけれども、こういった取組をいろんなところでつなげていくことが大切であるということでもあります。ただ、一番下にありますように、先ほど指摘もありましたけれども、マンパワーやコストの観点から、やはり持続的な方法を工夫していくことが何よりも大切で、余り性急に成果を求めず、継続的にじっくりやってくということが大切だというふうに感じております。時間がなくなりましたので省略いたしますけれども、全体的なことをまとめというかたちで提示しておきましたので、読んでいただければと思います。以上でございます。

平成25年度教育改革国際シンポジウム
2013年12月10日(火)13:00~17:30@ 文部科学省講堂

日本におけるAHELOの取り組みと 今後の展望


岸本 喜久雄
東京工業大学 大学院理工学研究科工学系長
AHELO Engineering Expert Group

1. 取り組みの概要(工学分野)
2. テスト問題
3. 参加したことの意義
4. 今後の展望
5. まとめ



1

1. 取り組みの概要(工学分野)

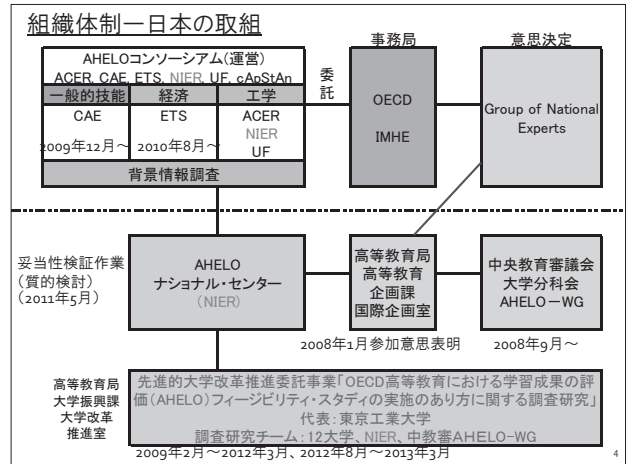


2

OECD-AHELO 高等教育における学習成果調査 Assessment of Higher Education Learning Outcomes

- 目的：大学教育の学習成果を世界共通のテストを用いて測定すること。
- 2008～2012年度にかけて実施されたのは、そうした国際的な学習成果アセスメントが実施可能であるかどうかを検証するためのフィージビリティ・スタディ。
 - 専門分野(工学・経済学)と一般の技能の3分野(+背景情報調査)で実施
 - 日本は工学分野で参加(12大学、504人の学生、196人の大学教員)
 - 最終報告書(邦訳2014年3月発刊予定)
 - AHELO Feasibility Study Report Volume 1 – Design and Implementation (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>)
 - AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf>)
 - AHELO Feasibility Study Report Volume 3 – Further Insights (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf>)
- 採用された能力枠組はTuning-AHELO(工学・経済学)

3



4

OECD-AHELOフィージビリティ・スタディ 実施可能性を探る試行的研究-3つの段階

第1フェーズ 2010年1月～ 2011年6月	Initial proof of concept 【妥当性検証作業-質的検討】 ●各国の多様性と特殊性をふまえて、学習成果について信頼できる結論を導くアセスメント・ツール(測定するための道具=テスト)を作成することが可能なか。 ⇒問題および採点ルーブリックの作成、小規模の実査、ヒアリング ＜問題の内容と翻訳・採点ルーブリックの内容の修正＞
第2フェーズ ～2012年12月	Scientific feasibility and proof of practicality 【妥当性検証作業-量的検討】 ●アセスメント・ツールの妥当性と信頼性は確保できているか。 ●大学と学生の参加を促し、アセスメントを適切に実施することが、実質的に可能なか。 ⇒大規模の実査と採点
最終会合 2013年3月	①②の結果を踏まえて、AHELOを本格的に実施するかどうか、改めて検討される。

- 3つの分野(一般の技能、経済学、工学+背景情報)

http://www.oecd.org/document/22/0,3746,en_2649_39263238_40624662_1_1_1_1_00.html

5

AHELO工学分野におけるチューニング

ブリュッセル専門家会議(2009/5/4～5)の主要議題

1. 工学分野における学術・専門領域
2. 工学分野における典型的な学位
3. 典型的な職業
4. 学習成果の記述の分類と順序付け
5. 期待する学習成果の定義
6. 各分野における期待される学習成果
7. 水準を示す指標の定義
8. 工学士の一般的記述
9. 学習、指導、評価への取り組み

6

学修成果 (Learning Outcomes) とは何か?

"Learning outcomes are statements of what a learner is expected to know, understand and/or be able to demonstrate at a completion of a process of learning."

AHELO: 工学分野におけるLOの選択

ABET Engineering Criteria 2000 (Washington Accord 加盟国も同等の規準; 日本はJABEE基準) と EUR-ACE Learning Outcomes for First Cycle Bachelor Degreesを参照して決定

=> "A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering"
OECD Education Working Paper No.60
(<http://dx.doi.org/10.1787/5kghtchn8mbn-en>)

7

7

チューニング・AHELO工学分野で期待される学習成果の概念枠組
A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering pp.28-29

技術者教育認定基準の特長	ABET 技術者教育認定基準	Tuning-AHELO 学習成果の特長
知識と理解 Knowledge and Understanding	a) 数学、科学、工学に関する知識を応用する能力	基礎科学・工学 Basic and Engineering Sciences
工学的分析 Engineering Analysis	b) 実験をデザインして実行し、データを分析して解釈する能力 e) 工学の課題を同定、整理、解決する能力	工学的分析 Engineering Analysis
工学デザイン Engineering Design	c) 経済、環境、社会、政治、倫理、健康、安全、生産可能性、持続可能性などの現実的な制約のもとで、定められた要件を満たすシステム、要素、工程をデザインする能力	工学デザイン Engineering Design
調査研究 Investigations	-	(「工学-分析」に統合)
工学の実践 Engineering Practice	f) 職業的・倫理的責任に関する理解 j) 現代的問題に関する知識 k) 工学の実践に必要な技法、技能、現代的な工学の道具を活用する能力	工学の実践 Engineering Practice (「汎用的技能」の一部を含む)
汎用的技能 Transferable Skills	d) 学際的なチームの一員として、役割を果たす能力 g) 効果的にコミュニケーションをとる能力 h) 工学による解決法のインパクトを、グローバル、経済、環境、社会的文脈のなかで理解するために必要な幅広い教養 i) 生涯を通じて学習に取り組む必要性を認識し、実際に取組む能力	一般的技能 Generic Skills (「知識と理解」の一部を含む)

(http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en)

8

8

Tuning AHELO 工学分野の中核概念の一覧

工学プロセス	
工学分析	実験をデザインして実行し、データを分析して解釈する能力。 工学の課題を同定、整理、解決する能力。
工学デザイン	経済、環境、社会、政治、倫理、健康、安全、生産可能性、持続可能性などの現実的な制約のもとで、定められた要件を満たすシステム、要素、工程をデザインする能力。
工学実践	専門職としての倫理的責任に関する理解。現代社会の問題に関する知識。 工学の実践に必要な技術手法、最新工学のツールを活用する能力。
工学基礎・専門	
工学専門 (土木工学)	材料と建設、構造工学、地盤工学、水工水理学、都市計画に関する知識を応用する能力。
工学基礎	数学・科学に関する知識を応用する能力。
一般的技能	
工学分野における 一般的技能	学際的なチームの一員として役割を果たす能力。効果的にコミュニケーションをとる能力。工学による解決策の影響を理解し、経済、環境、社会的文脈のなかで理解するために必要な幅広い教養。生涯を通じて学習に取り組む心構えと能力。

A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering
(http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en) (邦訳2014年3月発刊予定)

ENGINEERING ASSESSMENT FRAMEWORK (2011年11月19-19日第8回(開催))
(<http://www.oecd.org/dataoecd/28/44/49702022.pdf>)

9

9

2. テスト問題



10

概念枠組にもとづく問題の作成

- コンソーシアムが原案を作成
 - 多肢選択式問題は日本側が提案
 - 日本技術士会技術士第1次試験
 - 日本土木学会認定土木技術者資格試験
 - 記述式問題は豪州側が提案
- 国際専門委員会合
 - 提案された問題の取捨選択
 - 採択された問題について国際通用性を高める方向で検討

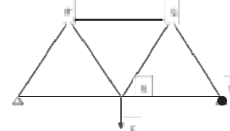
11

11

学習成果の範囲と水準の規定

多肢選択式問題の例

<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf> (pp.265-266)



下図のワーレントラスに荷重Pが作用している。下の記述のうち、正しいものを選びなさい。ただし、部材の自重は考えないものとする。

- 圧縮力が上弦材 (P-Q) と下弦材 (R-S) の両方に作用する。
- 引張力が上弦材 (P-Q) と下弦材 (R-S) の両方に作用する。
- 圧縮力が上弦材 (P-Q) に作用し、引張力が下弦材 (R-S) に作用する。
- 引張力が上弦材 (P-Q) に作用し、圧縮力が下弦材 (R-S) に作用する。

注目しているコンピテンス：工学専門（専攻する工学分野の重要事項や概念に関する系統的理解）

※多肢選択式問題は、日本の土木学会の認定土木技術者資格試験、日本技術士会の技術士第1次試験をベースに、国際専門家チームが国際通用性を高めるための検討を加えた。

12

12

学習成果の範囲と水準の規定
記述式問題の例

<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf> (pp.252-264)



フーバーダムは、アメリカ合衆国コロラド川のブラック峡谷にある、高さ221mのアーチ重力式コンクリートダムである。このダムは、灌漑用水の供給、洪水調節、およびダム基部の水力発電所への水の供給を目的として建設された。

図1. フーバーダム
図2. 1921年ごろに提案されたダム建設候補地
図3. 提案された貯水池のスケッチ
図4. ダムと発電所の建設計画
図5. ダムと発電所の建設計画

- この場所がダム式水力発電に適している理由を説明しなさい。少なくとも2つの側面について述べなさい。
- フーバーダムの構造上の強度と安定性を高めている。設計上の主な特徴を2つ挙げて説明しなさい。
- フーバーダムのタービンで発電される最大電力は $2.08 \times 10^9 \text{W}$ である。この発電所が90%効率で稼働している場合、この出力でタービンを流れるおおよその水量はいくらか。最も近いものを選びなさい。
A) 10^9sec , B) 10^8sec , C) 10^7sec , D) 10^6sec (注) $1 \text{W} = 1 \text{J/sec}$, $1 \text{J} = 1 \text{Nm}$
- 現在、この場所とは異なる別の場所で、新しいダムの建設が計画されていると想定します。技術者が環境影響評価書のなかで検討すべき、ダムが環境におよぼす影響(上流でも下流でもよい)を2つ挙げて、簡単に説明しなさい。

13

13

学習成果の範囲と水準の規定
第1問目の例

- 測定しようとしている学習成果
 - ダム式水力発電に適した場所の特徴を理解している。
- 背景にあるコンピテンス
 - 工学専門(専攻する工学分野の重要事項や概念に関する系統的理解、応用する能力)
 - 工学分析(既存の方法を用いて工学課題を同定し、解決策を考案し、解決する能力)
- 解答の観点
 - a) ダムの高さ・位置エネルギーの高さ
 - b) 川の流量の豊かさ
 - c) 貯水地域の広さ・安定した水量
 - d) 社会的影響が低く抑えられていること
 - e) 岩盤が堅固であること
 - f) 峡谷が狭いこと

それぞれの観点について、どの程度書けていれば、何点を配点するかについて、採点トレーニング(国際・国内)のなかで合意を形成

14

14

学習成果の範囲と水準の規定
第2問目の例

- 測定しようとしている学習成果
 - フーバーダムの構造上の強度と安定性に寄与する要件を理解している。
- 背景にあるコンピテンス:
 - 工学デザイン(現実的な制約のもとで、定められた要件を満たすシステム、要素、工程をデザインする方法を理解し、応用する能力)
- 解答の観点
 - a) アーチ式の形状をしていること
 - b) 峡谷の岸壁が堅固であること
 - c) コンクリートの重量が大きいこと
 - d) 堤体の重心が低いこと
 - e) 洪水吐きと排水トンネルが設けられていること

15

15

学生のコメント抜粋(第1フェーズより)

【多肢選択式問題について】

- ほとんどが大学で教わっている問題であり、解けてしかるべきだと思う。
- 問題が全部単発で、広く浅くという印象。もっと掘り下げた方がよいのでは。

【記述式問題について】

- 大学では理論的、抽象的な問題が多いなか、現実問題を扱っている点が面白い。倫理的な問題が面白い。
- 何が原因だったかを考え、どうするかという流れだったことから、よい問題、面白い問題だったと思う。実務に役立つ。
- 記述式のような問題が重要なら、グループワークやケーススタディの授業が必要だが、今まではなかった。解くアプローチを学ぶ機会がなかった。

16

16

3. 参加したことの意義



17

AHELO-FSに参加したことの成果

- 学習成果の枠組構築に参画し、日本の技術者教育認定で使用されている基準や教育の実態と整合的なものとなるように、意見表明することができた。
- 問題作成のプロセス(テーマ選択・項目の作成・採点ルーブリックの作成と修正)に参画し、日本の工学教育の視点から、国際的に積極的に情報発信することができた。
 - 日本の科学技術を支える工学教育に対する世界の関心と期待は強く、それに相応しい国際的な貢献を、工学分野での学習成果調査においても期待されていることを認識した。
- 国際的・国内的枠組みの双方において、工学と教育学の専門家が協働して教育改善に取り組むことの重要性を認識した。
 - 工学分野の取り組みが、他の専門分野にどのような示唆をもたらすかを丁寧に検討し、説明・還元していく必要がある。

18

18

テストの作成・実施・採点の経験から得られた知見

- 多肢選択式問題(工学基礎)-日本が提案した問題の国際通用性を確認
- 日本技術士会技術士試験・日本土木学会認定土木技術者資格試験
- 各国の平均点・分散に大きな乱れは認められなかった(第1フェーズ)
- 記述式問題(工学プロセス)-作成する難しさ・採点する難しさを実感
- 経験を蓄積することで克服できる難しさ
 - 適切な難易度
 - テスト時間に処理できる情報量
 - 翻訳の技術(語順、受動態へのなじみの薄さ、主語の省略、複数形)
- 「考える力」を統一的に測定することの難しさ
 - 問題の趣旨の伝わりにくさ:明確にしよとすると、解答が限定され、思考を方向づけることになる。
 - 採点基準の統一をはかろうとするほど、許容される正解の幅が狭まってしまふ。

19

テストの作成・実施・採点の経験から得られた知見

- 概念枠組「工学プロセスは工学基礎の習得に下支えされている」
 - 記述式問題と多肢選択式問題の間には高い相関がある(第1フェーズ)。
 - 記述式問題をわざわざ実施する必要があるのか。→ある
- 記述式問題の意義
 - 「記述式問題をもちいて工学プロセスのコンピテンス測定すること」自体に重要な試行的研究上の意義があるが、「工学プロセスのコンピテンスが重要である」という明確なメッセージを発信した点にも、イニシアチブとしての極めて重要な意義がある。国際的取り組みであるから、影響力が期待できる。
 - 学生にそのようなコンピテンスを獲得させるためには、どのような教育プログラムを履修させる必要があるのかを、各大学の教員が共同作業(科目横断的)として検討し、教育改善に着手することを促す効果が期待できる。
 - コンピテンス獲得にむけた教育プログラムの構造化
 - 教育方法の工夫(PBL、科目間連携、インターンシップ等)

20

4. 今後の展望



21

AHELOの潜在的インパクト

- 工学教育でどのような学習成果の習得が期待されるのかについて、異なる国や大学の専門家が共同作業として検討することをとおして、国際的に共通認識が醸成されていることを、具体的な場面で確認することができた。また、それを具体的なテスト問題に集約することができた。
 - 日本の工学教育の国際通用性について検討するための重要な資料
- AHELOの教育改善効果を最大化するために
 - 問題と採点基準に関する情報を公開し、教育プログラムとの関連づけを行うこと。
 - 各国から幅広い層の教員が多数、取り組みに参画し、国際的な学習成果の枠組みについての理解を深め、その更新に貢献できる仕組みを作ること。また、有効な教育アプローチについて情報共有できるように工夫すること。
 - AHELOから得られる情報が、各大学にとって妥当性の高いものになるように、調査の枠組み(調査条件、情報公開のあり方など)を工夫すること。

大学が国際的文脈のなかで教育改善に取り組む契機になることが期待される。

22

22

今後の展望

- 同様の取り組みを継続させていくことで、より多くの専門家にテスト問題や採点ルーブリックの作成に携わる機会を提供するとともに、その成果を広く情報公開することで、国際的な学習成果アセスメントを教育改善に役立てていく努力が求められる。
- 国際的な学習成果アセスメントのあり方を専門家チームが長期的展望をもって検討していく必要性がある。
- 大学教育の改善に資する調査デザインや情報公開のあり方を明らかにし、それを実現する工夫を行っていく必要がある。
- マンパワーやコストの観点から持続可能な方法を工夫するとともに、大学にとって参加する意義のある国際的な学習成果調査のあり方を検討する必要がある。

23

23

5. まとめ

- 工学教育における学習成果の内容について、国際的に共通認識が醸成されていることを具体的な場面で確認することができた。
- 国際的・国内的枠組みの双方において、工学関係者と教育学関係者が協働して教育改善に取り組むことの重要性を認識することができた。
- 日本の科学技術を支える我が国の工学教育に対する世界の関心と期待は強く、それに相応しい国際的な貢献を、工学分野での学習成果調査においても期待されていることを認識できた。
- 学習成果調査のあり方に関しては、何を、どのように測定し、どのように比較するのかについては、長期的展望をもって取り組む必要があり、国際的な専門家チームによるさらなる検討が求められる。これに我が国も積極的に関わっていくことは大きな意義がある。



24

ご清聴ありがとうございました。



25

第2部 パネルディスカッション

報告

深堀 聰子

国立教育政策研究所高等教育研究部総括研究官

「大学の教育改善に資する情報提供にむけてー日本・豪州・カナダにおけるAHELOフィージビリティ・スタディの取組」

パネルディスカッション

コーディネーター

金子 元久

筑波大学大学研究センター教授

パネリスト

ローベルト・ワーヘナール*

オランダ フローニンゲン大学人文学部長・同大学院研究科長

ピーター・ユーウェル*

全米高等教育経営システム研究所副所長

岸本 喜久雄

東京工業大学 大学院理工学研究科工学系長

メアリーキャサリン・レノン*

オンタリオ州高等教育質保証カウンシル上級研究アナリスト

ダニエル・エドワーズ*

オーストラリア教育研究所主任研究員

深堀 聰子

国立教育政策研究所高等教育研究部総括研究官

- *印が付いているパネリストは、当日英語で発言しており、本原稿は仮訳です。
- 報告書作成に当たり、当日の発言内容に修正を加えていることがあります。
- 所属団体、職名は2013年12月10日現在のものです。

報告

大学の教育改善に資する 情報提供にむけて —日本・豪州・カナダにおける AHELOフイージビリティ・スタディの取組

国立教育政策研究所高等教育研究部 総括研究官
深堀 聡子



1. はじめに

経済協力開発機構（OECD）による「高等教育における学習成果調査（Assessment of Higher Education Learning Outcomes, AHELO）」とは、大学教育の成果を世界共通のテストを用いて測定することを目的とする国際事業である。そして、2008年から2012年にかけて実施されたのは、そうした国際的な学習成果アセスメントが実施可能であるかどうかを検証するための調査研究としてのフイージビリティ・スタディ（以下、AHELO-FS）である。AHELO-FSでは、17か国の参加のもとに、分野横断的な「一般的技能」及び専門分野「経済学」「工学（土木工学）」で調査が展開された。さらに、各分野のテストに加えて分野共通の「背景情報」調査も実施された。日本は、豪州やカナダを含む9か国の一つとして、工学分野に参加した（図1参照）。

ここでは、このAHELO-FSの背景・目的・研究デザインを概観したうえで、日本における組織体制と実施手続を整理する。さらに、調査結果の試行的分析を行い、国際的な学習成果アセスメントの意義について考察する。

一般的技能	専門分野 —経済学—	専門分野 —工学—
コロンビア	ベルギー（Fl.）	アブダビ
エジプト	エジプト	オーストラリア
フィンランド	イタリア	カナダ
韓国	メキシコ	コロンビア
クウェート	オランダ	エジプト
メキシコ	ロシア	日本
ノルウェー	スロバキア	メキシコ
スロバキア		ロシア
米国3州※		スロバキア
背景情報		
※コネチカット、ミズーリ、ペンシルバニア		

図1. AHELO-FSの参加国（参加分野別）

2. AHELO-FSの背景・目的・研究デザイン

学士の学位取得者に期待される知識・技能・態度（学士課程教育の成果）とは、どのようなものであり、大学は、それを学生に習得させることに成功しているのだろうか。大学進学人口が拡大し、大学入学者の資質や卒業後の進路先が多様化する中で、大学教育の質を学習成果の観点から問い直すトレンドは、世界同時進行で顕在化してきている。そして、学生や労働者の国境を超えた移動が活発化する中で、単位や学位の等価性や国際通用性を保証する仕組みへの国際的関心も高まってきている。

そうした情勢の中でAHELO-FSは構想され、次の二つの目的を掲げて着手された。第1に、国際通用性をもつ学習成果アセスメントを作成することは可能なのか。大学教育を通して学生が習得すべき知識・技能・態度について国際的な合意を形成し、その達成度を妥当性と信頼性をもって測定することのできるテストを開発することは可能なのかを明らかにすることが目指された。

第2に、大学と学生の参加を促し、学習成果アセスメントを実施することは実質的に可能なのか。AHELOは大学教育の画一化や標準化を志向するものではなく、大学ランキングや政府による資源配分の根拠資料として使用されることを意図するものでもない。AHELOは大学教育の質向上に資することを旨とする取組である点について、大学の理解を獲得して調査への参加を得ることは可能なのか。また、調査の学術的・政策的意義について、学生の理解を獲得して、成績評価に関連付けることなく実施しても、テストに真剣に取り組むよう説得することは可能なのかを明らかにすることが目指された。

こうした研究目的にむけて、2段階の研究デザインが構想された。第1フェーズでは、アセスメント・ツールの開発とその妥当性の検証が目指された。すなわち、「一般的技能」「経済学」「工学(土木工学)」のそれぞれの分野において、大学卒業間際の学生を対象とするテスト、及び「背景情報」に係る調査票を開発し、それらの妥当性を小規模の実査に基づいて質的に検証することが計画された。小規模の実査では、各国の任意の10大学から任意の学生100人程度を抽出し、テストに解答するとともに、テストの妥当性に関するアンケートに回答し、大学教員と学生によるグループ・ディスカッション(フォーカス・グループ)にも参加するよう依頼するデザインがとられた。

第2フェーズでは、第1フェーズの結果に基づいて修正されたテストや調査票を用いて大規模の実査を行い、それらの妥当性と信頼性を量的に検証するとともに、テストを実施することが実質的に可能かどうかを検証することが目指された。大規模の実査では、各国の任意の10大学から無作為に学生1,000人程度を抽出し、テストに解答するとともに背景情報調査票にも回答するよう依頼するデザインがとられた。さらに、大学への機関調査、及び教員調査も併せて企画された(OECD, 2012, pp.77-95)。

こうしたサンプル・デザインから明らかのように、AHELO-FSのサンプルは、各国の大学システム全体を代表するものではないため、「国」を分析の単位とした比較を行うことを目的として設計されたのではない。AHELO-FSは「大学版PISA」と称されることもあるが、各国の15歳生徒人口が在籍する学校を母集団として、二段階層化抽出法に基づいて学校と生徒が無作為に抽出されたPISAのサンプル・デザインが採用されたわけではない点を強調しておく必要がある。AHELOは、「大学」を分析の単位として、大学の教育改善に資する情報を導くことを将来的な目的として設計されている。そして、AHELO-FSでは、そのために用いるテストを作成して、妥当性と信頼性を検証することが直接の目的とされている。

3. 日本における組織体制と実施手続

3.1 国際的・国内的な組織体制

日本はAHELO-FSに工学分野で参加したが、そうした政策決定と事業の遂行を進めるために、図2に示す組織体制が整備された。ここでは、国際的・国内的な組織体制を合わせて整理する。

はじめに、国際的な組織体制に注目する。AHELO-FSには、OECD教育政策委員会(Education Policy Committee, EDPC)、高等教育機関経営プログラム運営理事会(Programme on Institutional Management in Higher Education, 現在はOECD Higher Education Programmeに名称変更; IMHE Governing Board, IMHE-GB)、専門家会合(Group of National Experts, GNE)の協議に基づく重層的な意思決定の仕組みがとられ、その下で事務局(Secretariat)が事業運営にあたった。EDPCとは、各国の教育政策に係る課題を検討するために設置されている委員会で、OECD全加盟国から構成されている。それに対してIMHE-GBが代表するIMHEとは、大学等の高等教育機関及び高等教育に関する課題を扱う非営利組織を会員として、高等教育の課題について検討するためにOECDに設置されているフォーラムである。一方、GNEとは、AHELO事業の運営方針について検討するために設置された組織であり、AHELO-FS参加国を中心とする各国の政府関係者や高等教育関係者から構成されている。AHELO-FSは、高等教育をめぐる極めて多様なステークホルダーによる合議の下に推進された。

テスト・調査票の開発、各国におけるテスト実施(実査)に係るマネジメント、実査に基づく妥当性と信頼性の検証作業に係る委託事業は、豪州教育研究所(Australian Council for Educational Research, ACER)を代表機関とし、次の四つのグループから構成されるAHELOコンソーシアムが受託した。すなわち、「一般的技能」の取組を分担した米国のCAE(Council for Aid to Education)、「経済学」分野の取組を分担した米国のETS(Educational Testing Service)、「工学(土木工学)」分野の取組を分担したACER、日本の国立教育政策研究所(NIER)、及びイタ

リアのフィレンツェ大学 (UF)¹, そして「背景情報調査」の取組を分担したオランダのCHEPS (Centre for Higher Education Policy Studies) 及び米国のCPR (Indiana University Center for Postsecondary Research) である。さらに翻訳 (cApStAn Linguistic Quality Control Agency), データベース構築 (International Association for the Evaluation of Educational Achievement, IEA; Data Processing and Research Center, DPC), オンライン・テスト・システム開発 (SoNET systems), サンプリング (Statistics Canada) に係る技術専門機関も, AHELOコンソーシアムのメンバーとして事業に参画した。AHELOコンソーシアムでは, 定期的にテレカンファレンスを実施し, 原則として事業の基本方針や方法を共有しながら, それぞれの分担する事業を遂行した。

さらに, AHELO-FSでは, 教育調査及び高等教育研究の専門家から構成される技術諮問グループ (Technical Advisory Group, TAG) が設置され, 専門的・客観的な立場からGNEとAHELOコンソーシアムに助言する役割を担った (OECD, 2012, pp.96-100)。

次に, 国内的な組織体制に注目する。日本では, 中央教育審議会大学分科会AHELOワーキンググループの助言の下に, 文部科学省高等教育局高等教育企画課国際企画室がAHELO-FS事業を所管し, 国立教育政策研究所がナショナル・センターとしてテスト問題の翻訳・適正化, 及びテストの実施を担当した。

さらに, 文部科学省は東京工業大学に先導的の大学改革推進委託事業「OECD高等教育における学習成果の評価 (AHELO) フィージビリティ・スタディの実施のあり方に関する調査研究」を委託し, 工学分野におけるAHELO-FSの望ましい在り方, 日本がAHELO-FSに参加する際の留意点, AHELO-FSの参加経験を今後の各大学における教育改善に生かすための方法等を検討するための調査研究が推進された。この調査研究を通して, AHELO-FSという国際的な学習成果アセスメントを大学教育の質保証という政策枠組み及び教育改善努力の一環として位置付け, その意義と課題について組織的に検討できたことは, 非常に意義深いことであったと思われる。また, この調査研究での検討と専門的助言を得ることによって, 国立教育政策研究所が国際コンソーシアムのメンバーとしてテスト問題の作成に積極的に貢献したり, ナショナル・センターとしてテスト実施の在り方について活発に発言したりすることが可能になったことも強調しておく必要がある (東京工業大学, 2013年)。

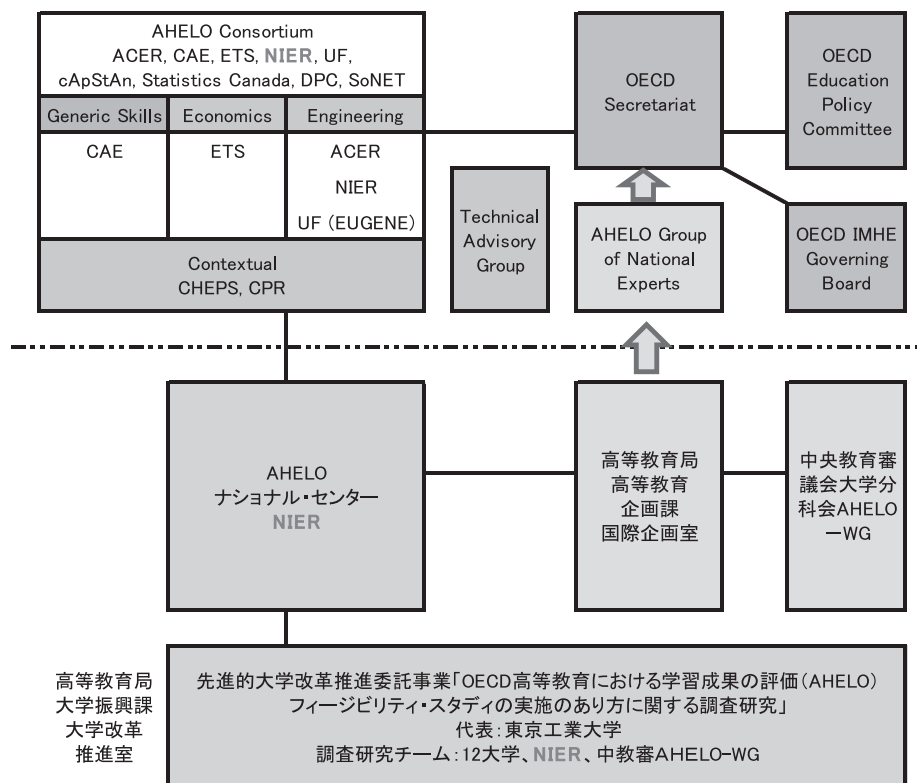


図2 AHELO-FSの組織体制 (国際的・国内的)

¹ フィレンツェ大学は, 76か国が加盟する欧州グローバル工学教育学術ネットワーク (European and Global Engineering Education academic network; EUGENE) の幹事校を務めている。

3.2 工学分野におけるコンピテンス枠組みの構築

学習成果アセスメントを通して何を測定しようとするのか。AHELO-FSでは、テストで測定する学習成果を体系的に捉えるためのコンピテンス（能力）枠組みとして、「Tuning-AHELO工学分野における期待される学習成果の概念枠組」がチューニング協会によってチューニングの方法に基づいて策定され（OECD, 2011a）²、GNE工学専門家の審査・助言・承認をうけて確定された（OECD, 2012, pp.121-128）。

チューニングとは、専門分野ごとに学生に身に付けさせようとするコンピテンスの枠組みを定義し、大学の自律性や多様性を尊重しながらコンピテンス枠組みに則した学位プログラムを構築する「方法」や「手続き」であり、それを採用する大学の取組を指す。コンピテンス枠組みを定義するに当たっては、教員が卒業生や雇用主と協議しながら、専門分野の意義を学術的観点からだけでなく、学生の進路先である社会的観点からも、分かりやすく説明しようとする点に特徴がある。また、学位プログラムを構築するに当たっては、大学間で共有する抽象性の高い「コンピテンス」の体系の中から、各大学の固有のミッションや学生ニーズに対応したコンピテンスを選び取り、重みづけをして、科目の履修を通して達成可能であり、成績評価をするために測定可能な「学習成果」に具体化させていく方法を示している点に特徴がある。さらに、大学間で共有するコンピテンス枠組みに即して各大学が学位プログラムを構築し、学位プログラムを構成する各科目の中で設定された学習成果の習得を単位認定の要件とすることで、外的基準に基づく大学の内部質保証を推進している点に特徴がある（ゴンサレス・ワーヘナール編（深堀・竹中訳）、2012年）。

チューニングは、欧州高等教育圏の確立を目指すボローニャ・プロセスの実質化を図る目的で、欧州の大学によって2000年に発案され、欧州委員会補助金を受けながら自主的に推進されてきた。チューニング第1・2期（2000～2004年）に九つ専門分野（経営、化学、教育科学、地球科学、歴史学、数学、物理学、看護学、欧州学）でコンピテンス枠組みが定義されたのを皮切りに多様な専門分野で援用され、北南米、ロシア、アフリカ、豪州、中央アジア、タイ、中国へと拡大してきた（Tuning Association, 2013）。

工学分野では、AHELO-FSが着手される以前から、既に多くの国々において、大学等の工学教育プログラムの内容が適格であるかどうかを認定する仕組みがかなりの程度確立されており、その国際通用性を保証するための取組も先駆的に進められてきていた。例えば、米国や英国をはじめとする加盟国の間で、各国の工学教育プログラムの適格性を相互に承認するためのワシントン協定が1989年に発足し、日本も技術者教育認定機構（JABEE）が2005年より加盟している（IEA, 2013）。欧州でも、欧州圏内の工学教育プログラムの適格認定を行う制度（EUR-ACE）が2008年に発足した（ENAE, 2013）。こうした動きの背景には、国境を越えて活躍する技術者が増加しており、各国の技術者としての専門職資格の等価性・同等性を相互に承認する仕組みが不可欠になってきていることが挙げられる。そうした中で、技術者としての専門職資格の基盤をなす工学教育にも、国際通用性が要求されるようになってきているのである。

AHELO-FSでは、このワシントン協定とEUR-ACEの基準として掲げられている分野別要件の内容を比較対照し、その共通性を抽出する方法でコンピテンス枠組みが定義された。そこでは、表1に要約したとおり、工学分野におけるコンピテンス枠組みが「工学基礎・専門」「工学分析」「工学デザイン」「工学実践」「工学一般的技能」という五つのコンピテンス・クラスターに分類されて定義された。

AHELOコンソーシアムでは、このように定義された工学コンピテンス枠組みに基づいてテスト問題を作成した。その際、「工学基礎・専門」は、主として基礎的な知識・技能の習得度を測定することに適した多肢選択式問題、「工学分析」「工学デザイン」「工学実践」「工学一般的技能」は、主として「考え方」を測定することに適した記述式問題で捉えることとした。多肢選択式問題の原案は、日本の土木学会の認定土木技術者資格試験（土木学会技術推進機構、2013）及び日本技術士会の技術士第一次試験（日本技術士会、2013）の問題の提供を受けて作成した。また記述式問題の原案は、実在する構造物の構造や機能の特徴を分析したり課題を解決したりする能力を問う問題を豪州チームが作成した。これらの原案を、各国（豪州、日本、イタリア、ドイツ、スウェーデン、米国、メキシコ）の工学教育の推進に中核的役割を担う専門家から構成される委員会でも精査する方法で確定した。いずれの問題についても、測定しようとするコンピテンスと学習成果、難易度、回答の観点と水準、配点を明記した採点ルーブリックを合わせて開発した（OECD, 2012, pp.252-268）。

²同様に、経済学分野でもチューニングの方法に基づいてコンピテンス枠組が定義された（OECD, 2011b）。

表1 Tuning-AHELO 工学分野におけるコンピテンス枠組（概要）

コンピテンス・クラスター	コンピテンスに関する記述
工学基礎・専門 Basic and Engineering Sciences	数学，科学，工学に関する知識を応用する能力。
工学分析 Engineering Analysis	実験をデザインして遂行し，データを分析して解釈する能力。工学の課題を同定，整理，解決する能力。
工学デザイン Engineering Design	経済，環境，社会，政治，倫理，健康，安全，生産可能性，持続可能性などの現実的な制約の下で，定められた要件を満たすシステム，要素，工程をデザインする能力。
工学の実践 Engineering Practice	専門職としての倫理的責任に関する理解。現代社会の問題に関する知識。工学の実践に必要な技術手法，最新の工学のツールを活用する能力。
工学一般的技能 Engineering Generic Skills	学際的なチームの一員として役割を果たす能力。効果的にコミュニケーションを取る能力。工学による解策の影響を国際社会，経済，環境，社会的文脈の中で理解するために必要な幅広い教養。生涯を通じて学習に取り組む心構えと能力。

3.3 テストの実施

前述したとおり，AHELO-FSの実査は2段階で実施した。第1フェーズ（2010年1月～2011年6月）では，開発された問題によって，各国の学生が大学で学んだ内容が問われているか，各国語に翻訳したテスト問題によって同等の学習成果が測定されているか，すなわち翻訳によって測定されている内容，設問の明示性，問題の難易度に違いが生じていないかを，小規模の実査に基づく質的検討を通して明らかにすることが目指された。

テストは，問題プールの中から多肢選択式問題20問，記述式問題1問を組合せ，大学ごとに異なる問題セットを出題し，60分間で実施した。さらに，テスト終了後に，テスト問題の妥当性に関する学生アンケートと，大学教員と学生によるディスカッションを行った。日本では，ナショナル・センターから参加を依頼した10大学の任意の学生75人を対象に，2011年5月16日から5月25日の期間に実施した。

第2フェーズ（2011年7月～2012年12月）では，第1フェーズの結果に基づいて修正したテスト問題によって，測定しようとしている学習成果を実際に測定することができたか（妥当性），また何度測定しても同等の結果が得られるか（信頼性）を，大規模の実査に基づく量的検討を通して明らかにすることが目指された。さらに，大学と学生の理解と協力を得て，学生サンプリングを適切に実施し，一定の回答率を確保することができるかを検証することも目指された。

テストは，問題プールから多肢選択式問題25問と記述式問題1問を項目反応理論に基づいて組み合わせた18種類のパターンの中から，学生ごとに異なる組合せを出題し，90分間で実施した。日本では，ナショナル・センターから依頼した12大学（全国の国立大学8校，私立大学4校）の学生504人（土木工学プログラムの学生全数，回答率12～100%・平均回答率65%）の参加の下に，2012年4月23日から5月25日の期間に実施した。学生調査と合わせて，大学調査（12大学が参加）及び教員調査（196人が参加）も実施した（OECD, 2012, pp.147-172）。

3.4 テストの採点

採点者がテストで測定しようとしているコンピテンスと学習成果について理解を共有し，共通の観点から同等の水準でテストを採点できているかどうかは，テストの妥当性と信頼性を確保する上で極めて重要な要件といえる。AHELO-FSでは，記述式問題の採点ルーブリックの観点と水準の規定及び採点者のトレーニングに特に重点的に取り組んだ。

第1フェーズの小規模実査では，AHELOコンソーシアム側で準備した採点ルーブリック（第1版）に基づいて，各国で採点を実施した。日本では6人の工学専門家が，75名の学生の解答を採点する過程で明らかになった採点ルーブリックの問題点を整理してコメントした。例えば，正答とみなされ得る解答の観点が採点ルーブリックに含まれていない場合には加筆を要求したり，採点ルーブリックの観点や水準が不明瞭な場合には明確化を要求したり，観点の重要性に相応した点数配分（重みづけ）を要求したりした。AHELOコンソーシアムでは，こうした各国からのコメントに基づいて修正した採点ルーブリック（第2版）を取りまとめた。

第2フェーズの大規模実査の採点に備えて、国際的な採点トレーニングが2回（のべ4日間）実施された。ここでは、各国の採点リーダーが一堂に会し、第1フェーズの小規模実査から得られた日本と豪州の学生の解答を採点ルーブリック（第2版）に基づいて実際に採点し、共通の結果が得られるまで協議を重ね、必要に応じてさらに修正を加えた（採点ルーブリック第3版・確定版）。

この際、採点ルーブリックの観点や水準に関する国際的な共通理解を確立する過程で、日本の学生の解答が採点資料として多用されたことには、日本の学生の学習状況が適切に反映される上で、大きな意義があったと思われる。国際的な採点トレーニング終了後には、各国において、採点リーダーの主導の下に国内的な採点トレーニングが実施された。AHELOコンソーシアムでは、採点が可能な限り共通の手順で進められるよう、採点トレーニングと実際の採点のためにオンライン・プログラムを開発した。

日本では、採点リーダーの主導の下、12人の工学専門家がのべ3日間にわたって採点トレーニングと採点に従事した。採点トレーニングでは、大規模実査で回収された学生の解答の一部について、採点リーダーが事前に採点し、その結果と採点者が採点した結果が一致しない場合、一つ一つについて採点チーム全員で不一致の理由を検討し、採点ルーブリック（確定版）の観点と水準に関する共通理解の確立が図られた。

こうしたトレーニングを経て遂行した採点は、結果的に高い信頼性が確保された。採点プログラムでは、学生の解答の2割余りについて、自動的に無作為に二人の採点者が重複して採点するように設定されていたが、日本では、この重複採点の一致度は89.1%に上った。異なる採点結果については、採点リーダーの判断で修正するとともに、採点チームとして、随時採点の観点と水準を確認する機会が採られた（OECD, 2012, pp.173-180 ; OECD, 2013a, pp.90-95）。

4. 調査結果の試行的分析

4.1 調査結果の分析についての考え方

AHELO-FSの大学サンプルはナショナル・センターの依頼に基づく便宜サンプルであり、各国の大学システム全体を代表するものではないため、「国」を単位とした調査結果の分析を行うことは適切ではない点は、先に確認したとおりである。

さらに、AHELO-FSの大規模実査では、項目反応理論に依拠して学生が異なる問題セットに解答しているため、テスト問題開発中の現段階では、「学生」を単位とした調査結果の分析を行うことも厳密には難しい。すなわち、項目反応理論では、過去の調査実績に基づいて問題の難易度や学生の得点分布があらかじめ既知の場合、異なる問題セットに解答した学生の得点を標準化して比較可能にすることができるが、AHELO-FSではこうしたテストの特性に関する情報が整っていない。AHELO-FSで項目反応理論に基づく調査デザインが採られたのは、大学や学生の負担をできるだけ抑制しながら、1度の調査にできるだけ多くのテスト問題の妥当性と信頼性を検証することを優先した政策判断によるが、結果的に調査結果を正しく理解するためには心理統計学の高度な知識が必要になり、工学専門家や教育学専門家にとって調査結果を分析するのが難しくなった点を指摘しておく必要がある。

AHELO-FS大規模実査は、国際通用性をもつ学習成果アセスメントを作成して実施することが可能かどうかを量的に検討することを目的として手がけられた。調査結果の心理統計学に基づく分析から、OECDはテストの妥当性と信頼性は基本的に検証されており、国際的な学習成果アセスメントは実施可能であると結論づけ、AHELOの本調査実施にむけた議論を開始した（OECD, 2013b）。しかしながら、AHELOが大学の教育改善に資するいかなる情報を提供し得るのかについては、AHELO-FSの枠組みの中では具体的な提案がほとんどなされなかった。AHELO本調査が実施可能となるためには、政府と大学の協力が不可欠であるが、調査に協力することが政府や大学にいかなるメリットをもたらすのかについての説明責任は、いまだ果たされていない。

AHELO-FSの調査デザイン上、「国」を分析の単位とした調査結果の分析を行うことは適切ではない。また「学生」がAHELOのテスト得点を個人の能力指標として利用できるようになるまでには、一定程度の調査実績の蓄積を待たなければならない。それでは、多大なコストをかけて、国際的な学習成果アセスメントを実施することの意義は果たしてどこにあるのだろうか。

先述した先導的の大学改革推進委託事業「OECD高等教育の学習成果の評価（AHELO）フィージビリティ・スタディ

の実施のあり方に関する調査研究」では、「大学」が国際的な学習成果アセスメントから期待する教育情報に関する検討が行われ、大きく次の3点が議論された。

第1は、国際的なベンチマーキング情報である。自校の学生の学習成果の水準は、国際的な水準と照らし合わせてどの程度なのか。その際、大学の問題関心は、日本の大学グループが「グループ全体」として世界で何位なのかということではなく、自校の学生の学習状況が、所定の大学グループの中でどの程度なのかという点にある。ベンチマーキングする大学グループが国の大学システム全体を代表しているかどうかは問題とされず、どの大学から構成されているのかという具体的な情報こそが重視された。

第2は、自校の強みと弱みを客観的に理解するための情報である。AHELO-FSでは、「工学基礎・専門」「工学分析」「工学デザイン」「工学実践」「一般的技能」の五つのコンピテンス・クラスターから構成されるコンピテンス枠組みが採用されたが、自校の学生はこのコンピテンス枠組みのどの領域で秀でており、どこに課題を抱えているのか。例えば、図3に示すような自校の学生のコンピテンス・プロフィールを国際的にベンチマーキングした情報を大学は求めているのである。

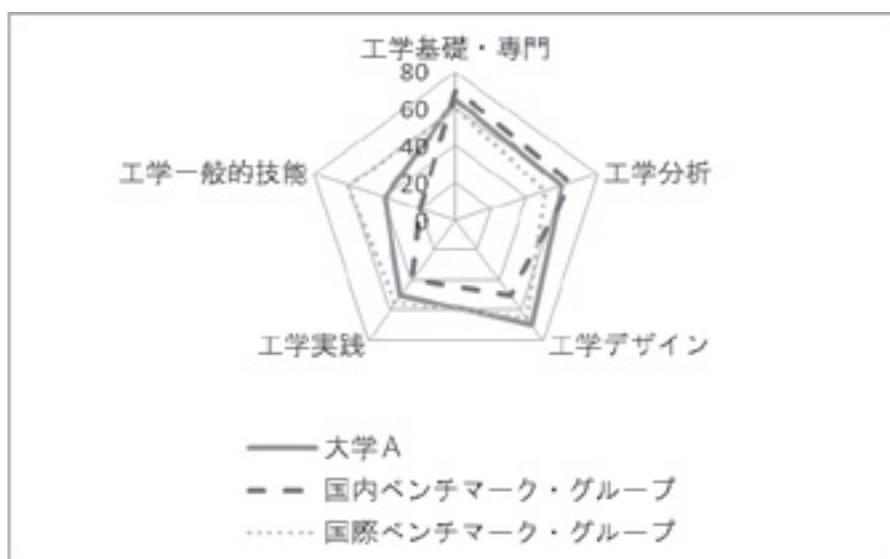


図3 コンピテンス・プロフィール (架空データ)

第3は、学生の学習実態や大学の教育環境の在り方とテスト得点との関係を理解するための情報である。大学教育の在り方にどのような課題があり、どう改善すれば学習成果の効果的な習得を期待することができるのか。学生による大学での学びをデザインする際に参考となるエビデンスを大学は求めている。

国際的な学習成果アセスメントに対する大学の期待に、AHELOはどの程度応えることができるのだろうか。繰り返しになるが、AHELO-FSはAHELOの実施可能性を検証するための調査であり、AHELOの目的である大学の教育改善に資する情報を導くことまで視野に入れて設計されたものではない。

特に、コンピテンス・クラスター別の分析については、残念ながらそれが可能となる十分なデータが今回の調査からは得られていない。したがってここでは、AHELO-FSから「何が明らかになったのか」を厳密に追究するのではなく、「どのようなことを明らかにできそうなのか」という探索的な観点から、試行的に手がけた分析結果を参考資料として提供したい。

以下のセクションでは、学生のテストや学部教育に対する印象、及びテスト結果との関係から「テストの妥当性」を確認した上で、学生の「時間の使い方」とテスト得点、大学の「授業形態」と学生のテスト得点について整理する。分析に使用するテスト得点データは、多肢選択式問題と記述式問題の総合点である。この総合点は、学生が共通して解答した多肢選択式問題に基づいて各テスト項目の正答確率を推定して導いた標準得点である（平均500点、標準偏差100）（9か国、70大学、学生6,078人）。

4.2 テストの妥当性

はじめに、テストの妥当性について確認しておこう。AHELO-FSのテストでは、学生が大学で学習した内容を適切に測定することができたのだろうか。ここでは、学生がテストの妥当性についてどのように感じたのか、そしてそうした感じ方がテスト得点とどのような関係にあるのかに注目する。

まず、図4が示すとおり、テストが「学部教育で学習した内容に合致していた」と回答した学生は9割近く（「非常によくあてはまる」14%、「よくあてはまる」37%、「あてはまる」37%）に上り、合致していたと感じている程度が強い学生ほど、テスト得点が高かった。同様に、テストが「将来の仕事で必要とされる内容に合致していた」と回答した学生は75%（「非常によくあてはまる」7%、「よくあてはまる」24%、「あてはまる」44%）に上り、合致していたと感じている程度が強い学生ほど、テスト得点が高かった。したがって、AHELO-FSのテスト問題は、学部教育の学習成果を測定するツールとして、おおむね妥当だと学生によって受け止められているようであり、大学教育に適応している学生の能力を測定する上で有効であったといえそうである。

なお、テスト得点は、学生の成績とも矛盾しない関係にあり、「学部教育を通しての成績（同じ専攻の他の学生と比較して）」（上位13%、平均より上25%、平均24%、平均より下22%、下位16%）が上位の学生ほど、テスト得点が高かった。一方、「解答するのにどれほどの労力が必要だったか」（「この上なく必要」15%、「かなり必要」45%、「少し必要」40%、「まったく必要なかった」1%）という問いに対して、より多くの労力が必要だったと回答した学生、すなわちテストが難しかったと感じた学生ほど、テスト得点が低かった。こうした結果も、AHELO-FSのテスト問題が大学教育に適応している学生の能力を測定する上で有効であったことを裏づける結果とみなすことができる。

学生の学部教育に対する評価と、テスト得点の関係についても確認しておこう。学部教育は「将来の職業や仕事で

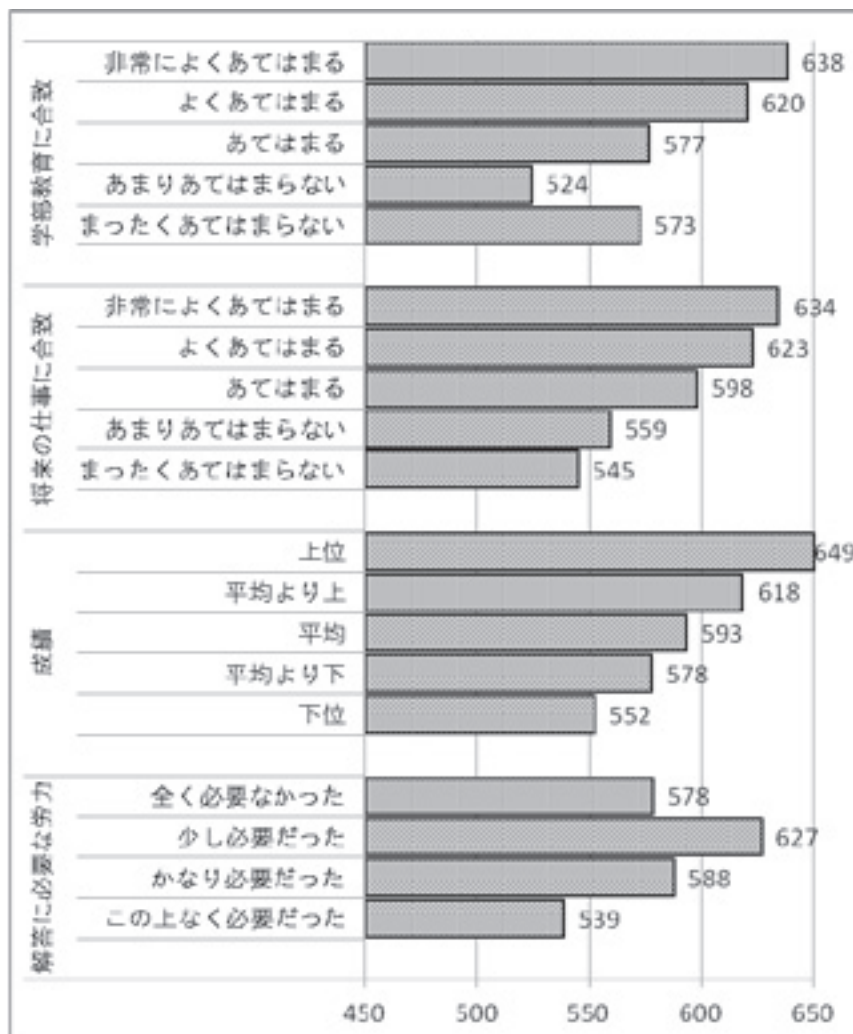


図4 テストの妥当性：学生の印象とテスト得点

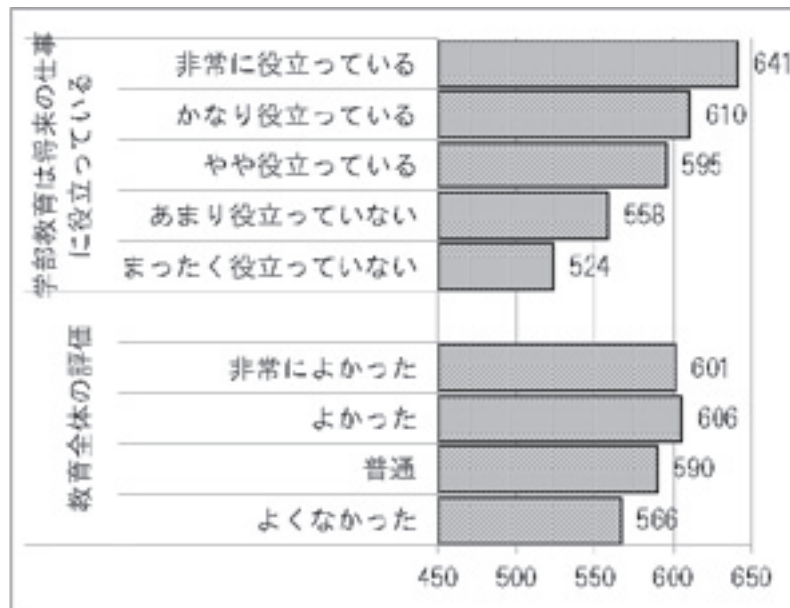


図5 ストの妥当性：学部教育の評価とテスト得点

必要だと思われる知識・技能を習得するうえで「役立っている」と回答した学生は9割近く（「非常に役立っている」6%、「かなり役立っている」30%、「やや役立っている」50%）に上り、役立っていると感じている程度が強いほどテスト得点が高かった。「学部教育をとおして経験してきた教育全体」の評価では、よかったと回答した学生は約半数（「非常に良かった」9%、「よかった」44%、「普通」41%、「よくなかった」6%）に留まったが、よかったと感じている場合の方が、テスト得点が高かった。これらの結果も、AHELO-FSのテスト問題が大学教育に適応している学生の能力を測定する上で有効であったことを支持している。

4.3 時間の使い方

それでは、AHELOは大学教育の改善に資するいかなる情報を大学に提供し得るのだろうか。ここでは試行的な分析として、学生の生活時間の使い方とテスト得点の関係に注目してみよう。

日本の高等教育政策では、学生の大学での質の高い学びを実現するために、「事前の準備、授業の受講、事後の展開を通じた主体的な学びに要する総学修時間」を確保する重要性が強調されている（中央教育審議会、2012年）。その背景には、日本の学生の「総学修時間」が、一日平均4.6時間（東京大学大学経営・政策研究センター（CRUMP）「全国大学生調査」）に留まっているという問題意識がある。

AHELO-FSでは、学期中（3年生前期）の典型的な1週間（7日）に学生が諸活動に費やす時間を尋ねている。ここでは「正規の授業への出席（講義・個人指導・ゼミ等）」18.9時間（平均値、以下同様）、「実験・実習・実技に取り組む（実験室での実験やフィールド調査等）」5.7時間、「授業の準備（予習・宿題・発表の練習等）」6.8時間、「専攻に関連したアルバイト」0.5時間、「専攻に関係のないアルバイト」10.5時間、「大学の課外活動への参加（大学新聞、自治会、クラブ活動等）」7.7時間であった。すなわち、学生が授業、実験・実習・実技、授業の準備に費やしている平均時間の合計は31.4時間であり、1日平均4.5時間ということになる。これは上述の調査とほぼ同値である。

一方、専攻に関係のないアルバイトや課外活動に費やす時間の平均は、18.2時間（一日平均2.6時間）であった。なお、専攻に関連したアルバイトには、日本の大学生はほとんど携っていない点が特徴的である。

こうした時間の使い方は、学生のテスト得点とどのような関係にあるのだろうか。「正規の授業への出席」では、一週間に10時間以下の学生（15%）は、テスト得点が低い傾向があり、学習上の支援を要するグループとして注視する必要があるのかもしれない。「授業の準備」では、全体の56%を占める「0-5時間」グループのテスト得点が低く、「11-15時間」（10%）グループのテスト得点が最も高かった。16時間以上の時間を費やす効果はテスト得点には表れていない。したがって、「0-5時間」グループの学生の自律的な学習時間を「11-15時間」程度まで引き上げる教育上の工夫が必要なのかもしれない。

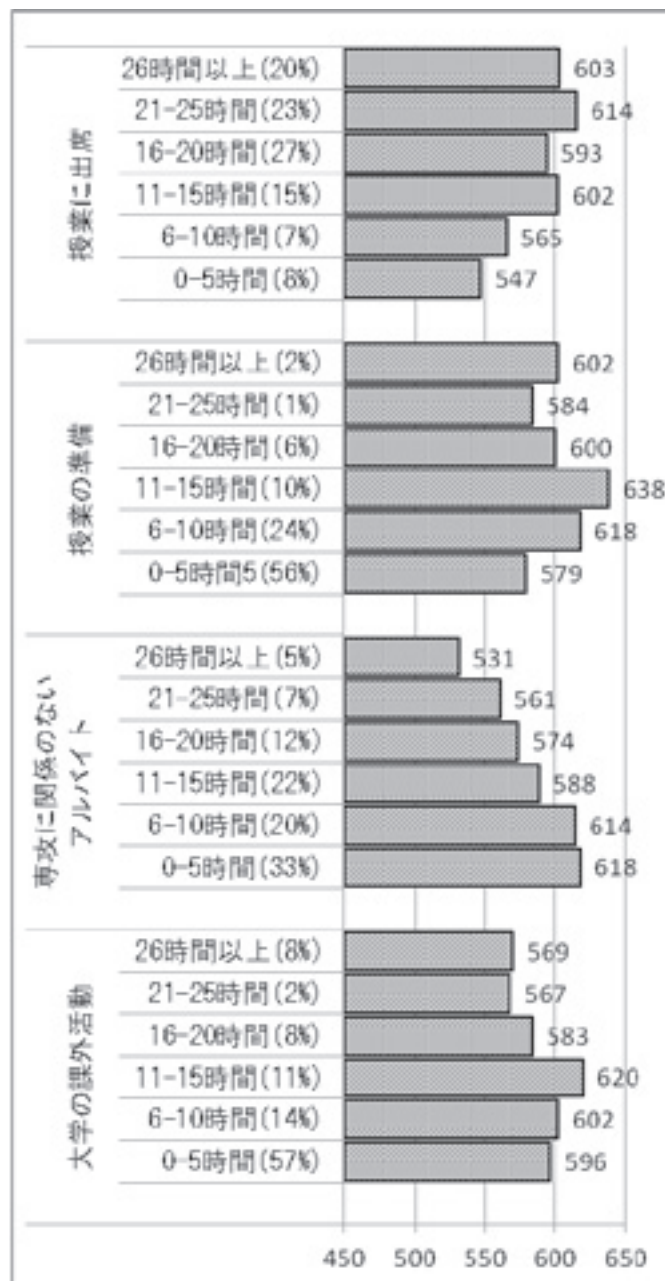


図6 時間の使い方とテスト得点

「専攻に関係のないアルバイト」は、学生の学びに負の効果をもたらしている。学生が専攻に関係のないアルバイトに費やす時間が長ければ長いほど、テスト得点は低下しており、特に1週間に10時間を超える場合に顕著である。学生がなぜ、専攻に関係のないアルバイトに過剰な時間を費やすのかを調査して、学生支援の拡充などの適切な措置を取る必要があるだろう。また豪州では、「専攻に関連したアルバイト」が学生の学びに正の効果をもたらすことが確認されていることから、そうした学習と結びついたアルバイトの機会を開拓していく工夫の余地があるのかもしれない。

最後に、「大学の課外活動」には、過半数の学生がほとんど参加していないが、テスト得点が最も高いのは、「11-15時間」グループ(11%)である。大学は学びの場であるとともに、社会的な交わりの場でもあることを考えると、課外活動への適度な参加は、大学生活全体への統合的参加を促し、学びにも望ましい影響をもたらすという仮説を立てることができる。

4.4 授業形態

学生が受講している授業の形態とテスト得点の関係に注目してみよう。日本の土木工学プログラムの学生が4年生

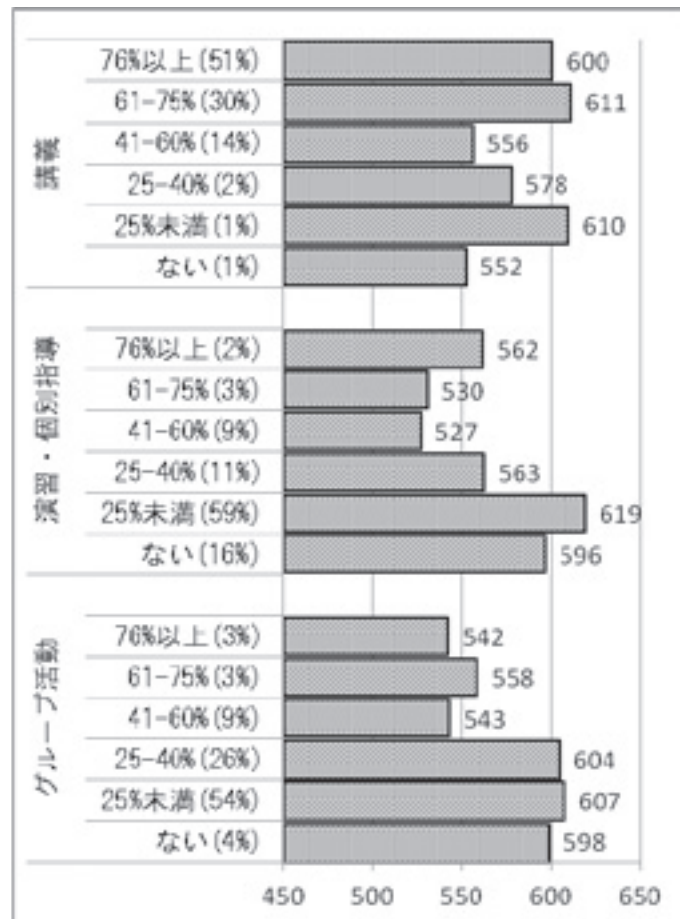


図7 各授業形態の割合別テスト得点 (全体)

の4月までに受講した「学部教育の正規の授業」全体に占める「講義（主に講師が話す）」「演習・個別指導（テーマについて学生が積極的に議論することが期待されている）」「グループ活動（教員の監督や支援のもとで、学生が協力して作業に取り組む）」の割合（平均値）は、それぞれ74%、20%、25%であった。学生は講義中心の授業を受講しており、演習やグループ活動などの参加型の授業はそれほど普及していない。

各授業形態が占める割合別にテスト得点を整理してみると、図7にみるとおり、「講義」形式の割合が高いグループ（「76%以上」51%、「61-75%」30%）、「演習」形式の割合が低いグループ（「25%未満」59%、「ない」16%）、「グループ活動」形式の割合が低いグループ（「25-40%」26%、「25%未満」54%、「ない」4%）ほど、テスト得点が高いことがわかる。

この結果から、次の二つの仮説をたてることが可能だろう。一つ目の仮説は、学生の学業成績の高い大学では、授業形態に係る工夫が余りなされておらず、伝統的な「講義」形式が踏襲されているのに対して、学生の学業成績に課題のある大学では、「演習」「グループ活動」などの参加型の授業を積極的に取り入れ、付加価値を高める努力が展開されているというものである。

二つ目の仮説は、「演習」や「グループ活動」などの参加型の授業の教育効果は検証されにくいというものである。学生による自律的な学習を促すためには、教員の側の周到な授業設計や教材準備に裏打ちされた高い教育力量が求められる。「講義」形式の授業が主流の日本では、いまだ十分な条件を整備するに至っていないのかもしれない。あるいは、参加型の授業の効果はテスト得点という形で顕著に捉えるのは難しいのかもしれない。

図8は、大学A（n=45）と大学B（n=23）について、各授業形態が占める割合別にテスト得点を整理した結果である。サンプル数が少ないため、極めて探索的な分析であるが、上述の二つの仮説を支持する結果とみることができる。

すなわち、大学Aに比べて学生の学業成績が優秀な大学Bでは、「講義」形式の授業の割合が高く、「演習」「グループ活動」形式の授業の割合が低い。また、学生もそうした授業形態に比較的よく適応しているようである。それに対

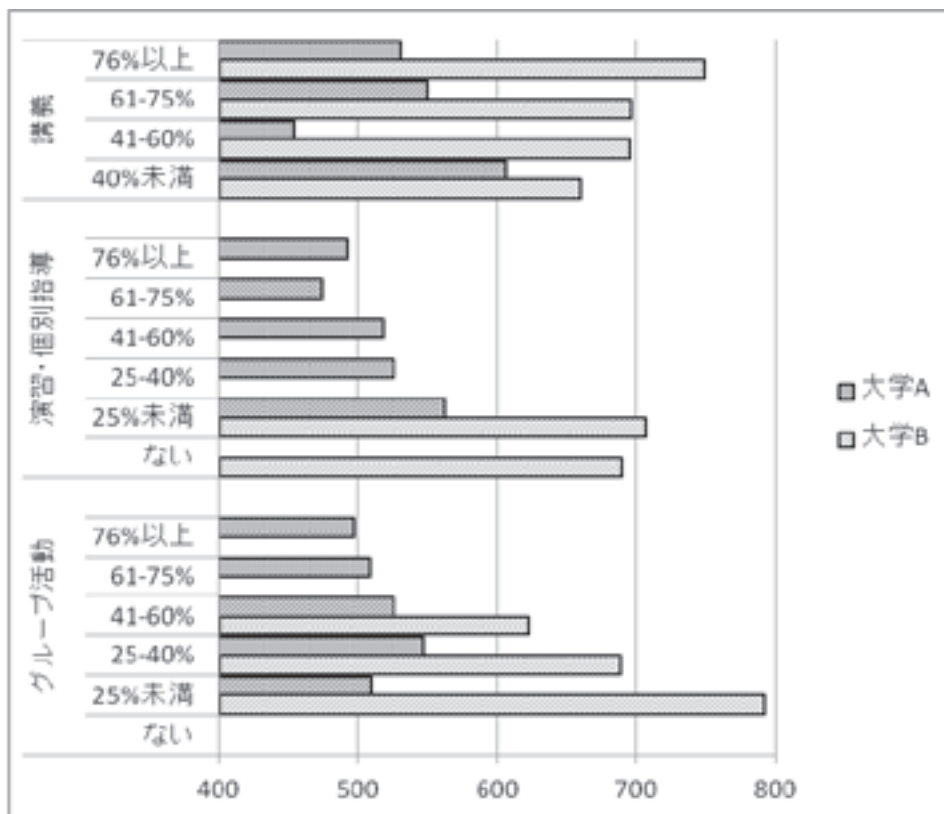


図8 各授業形態の割合別テスト得点 (大学A・B)

して大学Aでは、「講義」形式の授業が低く、「演習」「グループ活動」形式の授業が高い割合で導入されている。そして、学生のテスト得点は、受講している授業全体に占める「講義」形式の授業が4割未満の学生の場合に相対的に高いことを確認することができる。ところが、「演習」「グループ活動」の授業を多く受講している学生のテスト得点が高いかという点、そうした明確な傾向を確認することはできない。

4.5 大学の教育改善に資する情報提供にむけて

このように、AHELO-FS調査結果の試行的な分析から、学生の学習実態や大学の教育環境の在り方とテスト得点との関係についての理解を深めるうえで有益ないくつかの情報を導くことができた。大学が自校の教育課題を見極め、学生の学びをデザインするために参考となる情報を提供する可能性を、AHELOは確かに含んでいる。こうした情報提供の可能性を強化しながら、AHELOが更に取り組まなければならない課題は、大学のコンピテンス・プロフィールを国際的にベンチマーキングした情報に係る大学の要請にどう応えていくかという課題である。

大学が求める自校のコンピテンス・クラスター別の「強み」と「弱み」の診断、コンピテンス・プロフィールの作成は、AHELO-FSの調査結果の分析から実現することはできなかった。そうした情報提供をするために、どのような調査デザインを取り、いかなるテスト問題を開発する必要があるのかについて、専門的な検討を深める必要がある。一方、国際的なベンチマーキングの要請に応えるために、日本・豪州・カナダ（オンタリオ州）のナショナル・センターでは、データの共有・相互利用に関する覚書を交わし、3国の分析を行ってきた。この覚書では、3者がそれぞれのナショナル・データの機密性を相互に尊重し、教育研究以外の目的で使用することはないこと、互いのデータやその分析結果を事前に文書による承諾を得ることなく、第三者に開示することはないこと等について合意したことが明記されている。そして、この覚書に基づいて、3者では参加大学に国際ベンチマーク情報を提供することに努めてきた。こうした取組は、国際的なデータ共有のプロトコルづくりにむけて、重要な前例になると思われる。

5. 国際的な学習成果アセスメントの意義

AHELO-FSの取組を通して、国際的な学習成果アセスメントが政府や大学をはじめとするステークホルダーにどのようなメリットをもたらすのか、またどのような課題が残されているのかについて、具体的な議論を展開する材料が整ったことは、極めて重要な第一歩であったと思われる。

こうした意義に加えて、AHELO-FSに工学分野で参加し、テスト問題の作成・実施・採点といった一連の活動に参加したこと自体に大きな意義があったといえる。ここでは、そうした意義を3点に絞って整理して、本稿を締めくくりたい。

第1に、工学分野でどのような学習成果の習得が期待されているのかについて、異なる国や大学の専門家が共同作業として検討することを通して、国際的に共通認識が醸成されていることを、具体的な場面で確認することができた。またそれを具体的なテスト問題と採点ルーブリックに集約することができた。今後、こうした経験と成果に関する情報を広く普及させていくことで、日本の工学教育の国際通用性を高める方法の検討に対して、重要な示唆を提供することができると思われる。

第2に、AHELO-FSテスト問題（多肢選択式問題）では、日本側から提案した土木学会の認定土木技術者資格試験及び日本技術士会技術士第一次試験の問題が、国際的な工学専門家委員会によって、国際通用性のある問題として採択された。また、第1フェーズの小規模実査からその妥当性が質的に検証され、第2フェーズの大規模実査からもその妥当性と信頼性が量的に検証されたと、OECDは報告している。このことは、工学分野のコンピテンス枠組みを構成する「工学基礎・専門」に係る日本の標準的な問題が、国際標準に合致することが公認されたことを意味する。

第3に、工学分野のコンピテンス枠組みを構成する「工学分析」「工学デザイン」「工学実践」及び「工学一般的技能」について、新しい知見を得ることができた。これらのコンピテンスが重要であるという認識は日本でも既に共有されているが、具体的な取組として教育プログラムに広く反映されている状況とはいまだいえない。そうしたなかで、これらのコンピテンスを測定する試行的な取組を展開することで、その重要性について明確なメッセージを改めて発信することができた。さらに、これらのコンピテンスがどのような問題で測定され得るのかについて、大学教員や学生に具体的なイメージを提供することができた。

コラム1に示す記述式問題に対する学生のコメントが象徴するように、AHELO-FSに参加した経験は、関係者の多くに日々の教育について強烈な「気づき」をもたらしたと考えられる。そして、AHELO-FSは国際的な取組であることから、イニシアティブとして強い影響力を発揮することが期待される。

コラム1. 記述式問題に対する学生のコメント

- 学生A：「大学では理論的、抽象的な問題が多いなか、現実問題を扱っている点が面白かった。倫理的な問題も面白かった。」
- 学生B：「何が原因だったかを考え、どうするかという流れは、よい問題、面白い問題だったと思う。実務に役立つと思った。」
- 学生C：「記述式問題のような問題が重要なら、グループワークやケーススタディの授業が必要だが、今まではなかった。解くアプローチを学ぶ機会がなかった。」

【参考文献】

- 中央教育審議会『新たな未来を築くための大学教育の質的転換に向けて～生涯学び続け、主体的に考える力を育成する大学へ～（答申）』2012年。
- 東京工業大学「OECD高等教育における学習成果の評価（AHELO）フィージビリティ・スタディの実施のあり方に関する調査研究（最終報告書）」（文部科学省先導的の大学改革推進委託事業）2013年。
- 東京大学大学経営・政策研究センター『全国大学生調査』2008年。
- フリア・ゴンサレス/ローベルト・ワーヘナール編著（深堀聰子/竹中亨訳）『欧州教育制度のチューニング-ボローニャ・プロセスへの大学の貢献』明石書店，2012年。
- ENAEE (European Network for Accreditation of Engineering Education) (2008). *EUR-ACE (European Accreditation of Engineering Programmes) Framework Standards for the Accreditation of Engineering Programmes*.
(http://www.enaee.eu/wp-content/uploads/2012/01/EUR-ACE_Framework-Standards_2008-11-0511.pdf)
- IEA (International Engineering Alliance) (2013). Graduate Attributes and Professional Competencies
(<http://www.ieagrements.com/IEA-Grad-Attr-Prof-Competencies.pdf>)
- OECD (2011a), "A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering", *OECD Education Working Papers*, No. 60, OECD Publishing. (doi: 10.1787/5kghtchn8mbn-en)
(http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en)
- OECD (2011b), "Tuning-AHELO Conceptual Framework of Expected and Desired Learning Outcomes in Economics", *OECD Education Working Papers*, No. 59, OECD Publishing. (doi: 10.1787/5kghtchw3nn-en)
(http://www.oecd-ilibrary.org/education/tuning-ahelo-conceptual-framework-of-expected-and-desired-learning-outcomes-in-economics_5kghtchw3nn-en)
- OECD (2012). *Assessment of Higher Education Learning Outcomes Feasibility Study Report Volume 1 – Design and Implementation*.
(<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>)
- OECD (2013a). *Assessment of Higher Education Learning Outcomes Feasibility Study Report Volume 2 – Data Analysis and National Experiences*.
(<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf>)
- OECD (2013b). *Assessment of Higher Education Learning Outcomes Feasibility Study Report Volume 3 – Further Insights*.
(<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf>)


【ウェブサイト】

- 土木学会 技術推進機構「土木学会認定土木技術者資格制度-過去の筆記試験問題」
(http://www.jsce.or.jp/opcet/02_testQ.shtml) (2013年11月1日アクセス)
- 日本技術士会「試験・登録情報-試験関係資料」
(http://www.engineer.or.jp/c_categories/index02019.html) (2013年11月1日アクセス)
- 日本技術者教育認定機構
(<http://www.jabee.org/>) (2013年11月1日アクセス)
- Tuning Association. *Tuning Educational Structures in Europe*.
(<http://www.unideusto.org/tuningeu/>). (2013年11月1日アクセス)

国立教育政策研究所
平成25年教育改革国際シンポジウム
2013年12月10日(火) 13:00~17:30
パネル・ディスカッション報告

大学の教育改善に資する情報提供にむけて
5
日本・豪州・カナダにおける
AHELOフィージビリティ・スタディの取組

深堀聡子
国立教育政策研究所



1

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。
 - どのように採点するのか・・・採点ルーブリックをととして、学習成果の範囲(観点)と水準を規定し、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。

Satoko Fukahori 2013.12.10 NIER Symposium

2

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、
取組を進展させることはできない。

Satoko Fukahori 2013.12.10 NIER Symposium

3

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
 - 学生は在籍する大学(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」の得点を標準化して比較可能するには、問題の難易度や分布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。

Satoko Fukahori 2013.12.10 NIER Symposium

4

パネル・ディスカッションにむけて
- 問題提起 -

国際的な学習成果アセスメントは、 大学の教育改善に資する いかなる情報を提供しうるのか?

Satoko Fukahori 2013.12.10 NIER Symposium

5

大学が求めている教育情報

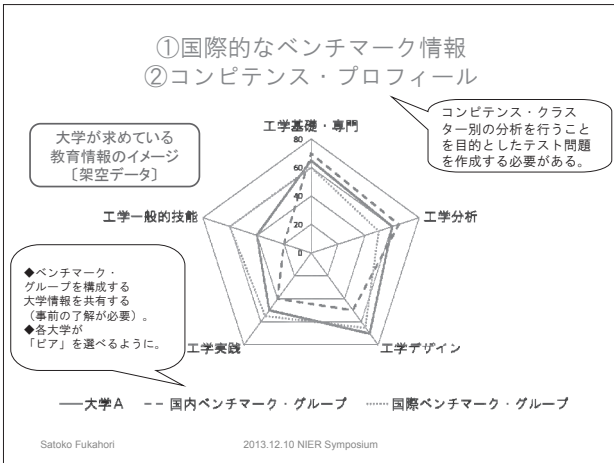
【検討の場】

- 文部科学省先導的大学改革推進委託事業「OECD高等教育における学習成果の評価(AHELO)フィージビリティ・スタディの実施の在り方に関する調査研究」(平成21~24年度)
- 国立教育政策研究所国際研究協力活動による研究「AHELO調査結果の分析に関する研究会」(平成25年度)

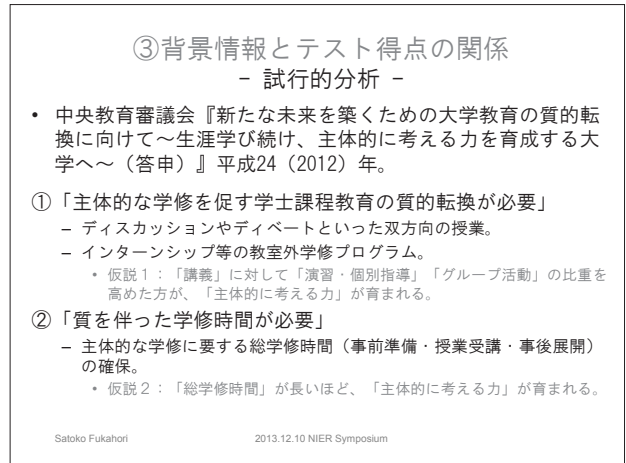
- ①国際的なベンチマーク情報
 - 「自校」の学生の学習成果は、国際的なグループと照合してどの程度か。
 - 日本の大学が「グループ」として、世界で「何位」なのかは、問題ではない。
 - ベンチマークする大学グループが、各国の大学システムを代表しているかどうかとも、問題ではない。
 - ベンチマークする大学グループが、具体的にどの大学から構成されているかの情報こそ重要。
- ②自校の学生グループの「強み」と「弱み」
 - コンピテンス枠組にもとづくコンピテンス・プロフィール。
- ③背景情報とテスト得点の関係
 - 教育改善の方向性に係る示唆を得るために。

Satoko Fukahori 2013.12.10 NIER Symposium

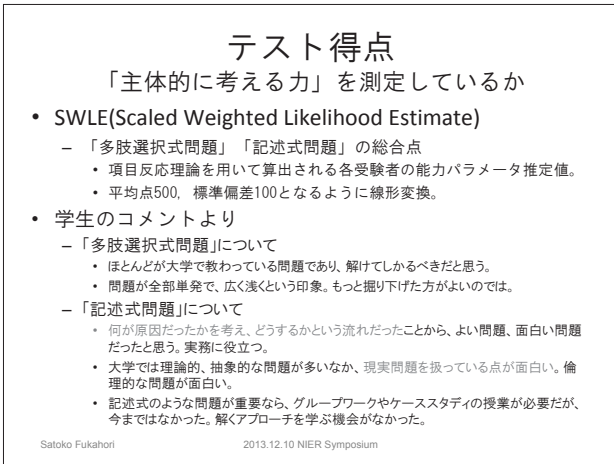
6



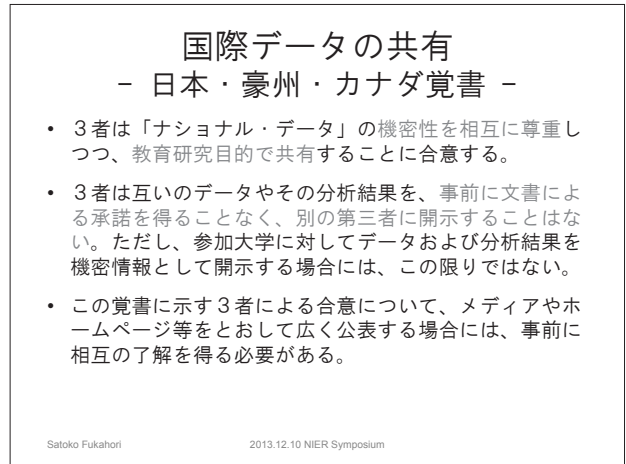
7



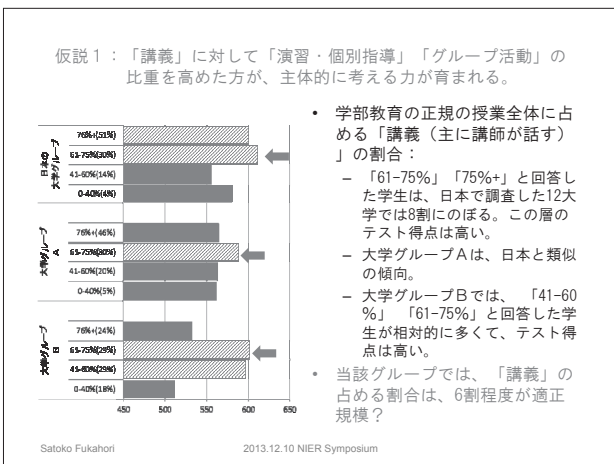
8



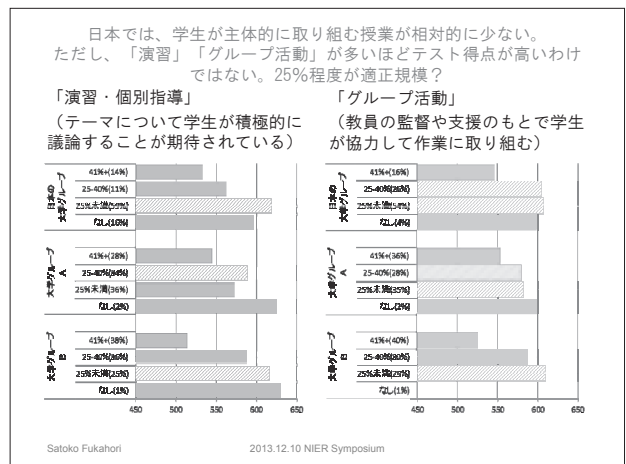
9



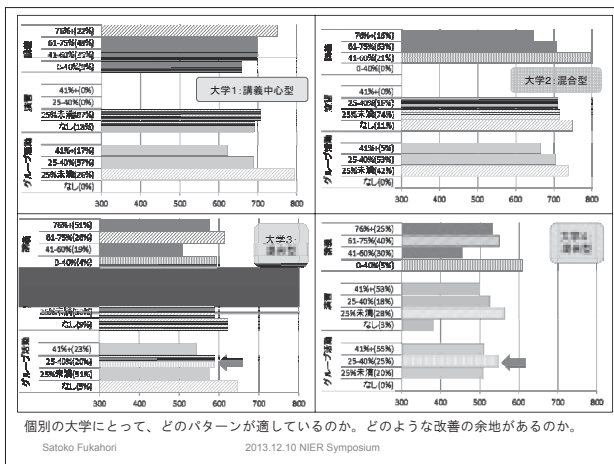
10



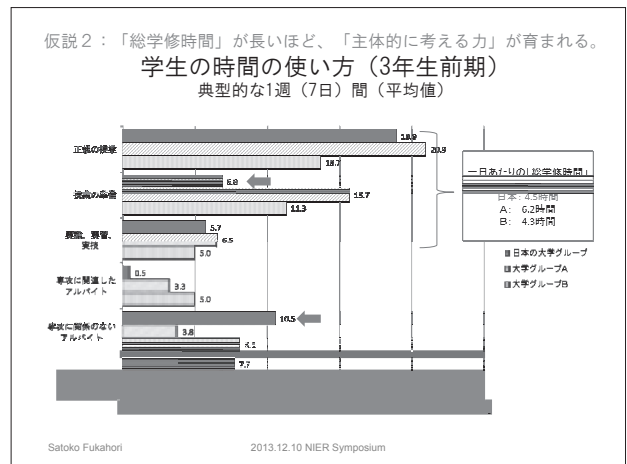
11



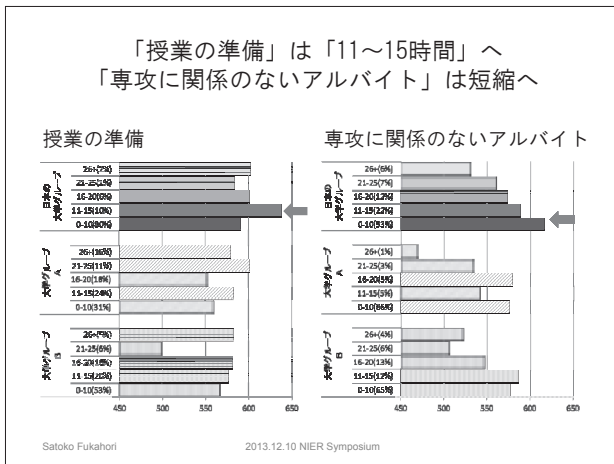
12



13



14



15

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
 - 自校の学生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
 - コンピテンス・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。
 - 大学の教育環境や学生の学習実態を、学生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スが、大学教育をとおして育成しようとしているコンピテンスと一致していること（整合性）。
 - コンピテンスにもとづく学位プログラムの体系化

Satoko Fukahori 2013.12.10 NIER Symposium

16

参考資料

- OECD. *Testing student and university performance globally: OECD's AHELO*. (<http://www.oecd.org/edu/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdcahelo.htm>)
- OECD (2011). *A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering*, OECD Education Working Papers, No.60, Paris: OECD Publishing. (2014年3月邦訳発行予定)
- OECD (2012). *AHELO Feasibility Study Report Volume 1 – Design and Implementation*, Paris: OECD Publishing. (2014年3月邦訳発行予定)
- OECD (2013a). *AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences*, Paris: OECD Publishing. (2014年3月邦訳発行予定)
- OECD (2013b). *AHELO Feasibility Study Report Volume 3 – Further Insights*, Paris: OECD Publishing. (2014年3月邦訳発行予定)
- Tuning Association. *Tuning Educational Structures in Europe*. (<http://www.unideusto.org/tuningeu/>)
- Gonzalez, J. and Wagenaar, R. (2008). *Tuning Educational Structures in Europe – Universities' Contribution to the Bologna Process*. Publicaciones de la Universidad de Deusto.
 - ゴンザレス・R, ワーゲナール(深堀暎子・竹中享訳)『欧州教育制度のチューニング-ボローニャプロセスへの大学の貢献』明石書店、2012年。

Satoko Fukahori 2013.12.10 NIER Symposium

17

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

ご清聴ありがとうございました。

Satoko Fukahori 2013.12.10 NIER Symposium

18

パネル ディスカッション

コーディネーター 金子 元久

パネリスト

ローベルト・ワーヘナール
ピーター・ユーウェル
岸本 喜久雄
メアリーキャサリン・レノン
ダニエル・エドワーズ
深堀 聡子

■ **金子氏** 大変な数の質問がきていますが、三つの基本的な問題をここからまとめさせていただきます。質問していただいたのに私のまとめの中に含まれてない方がいらっしゃると思いますが、時間の制限もありますのでお許しください。三つ基本的な問題の一つはこういうことです。テストをするために能力というものを何らかのかたちで概念的にまとめているわけです。これは特にチューニングの話で非常によく出てきたお話でありますけれども、メタプロフィールとかメタコンピテンシーとか、あるいはメタスキルとかですね。そういったものが能力を何らかのかたちで概念化するということが必要になってくるわけです。また、アメリカの場合には、ジェネリックスキルというような言葉がよく使われているわけですが、これらの概念自体がどういう根拠に立っているのか、また、例えばチューニングで言われているような能力の概念と、アメリカで言われているクリティカルシンキングのような概念というのが、同じものなのかどうかと。基本的にこの概念というものはどういうふうに出てきていて、特に実証的な道具に使うときに、その根拠はどこにあるのかというのが、第1に出てきている問題です。2番目は、このような大規模なテストを行ったときに、基本的には実施上どういう問題があるのか、発表では一応できたということでしたけれども、本当の問題はどこにあるのか、またユーウェル先生がおっしゃっていましたが、費用がどれぐらい掛って、特に効果はどこにあるのか、誰が負担すべきなのか、どのような効果があると言って社会に説得すべきなのか、という問題です。3番目の問題は、これからどうなるのかということ。一つは国際的にこのOECDの枠組みでさらに進められていくのかどうか、という点。もう一つは、これ以外の分野に広げていく可能性はあるのかどうかという点です。また、例えば工学の場合は今回土木工学だけでしたけれども、ほかにもいくつも工学だけでもいろいろな分野があるわけですから、そういったことができるのか、あるいは例えばリベラルアーツでこういったことができるのかどうかという問題もあるのではないかと思います。さらにそれを大学に広げるようなかたちで、大学、さらにほかの種類の高等教育機関にも広げていくことがあるのかといった問題もあるだろうと思います。

一応この三つの問題が提起されているわけですが、残念ながら1時間弱しかありませんので、今、申し上げた問題のいずれかにお答えいただければよろしいのですが、特に1番目の能力のフレームワークの問題については、ワーヘナール先生にお話ししたい。それからユーウェル先生にもこの点でお話しただけなら

と思います。もう一つはエドワーズ先生に先ほどお話に出ていましたが、ジェネリックスキルはCLAの問題と、それからACERで作った問題と、両方出ているわけですが、この二つを比べてみて何か考えるところはあるかどうかということをお話しただけならと思います。それから2番目の実施上の問題と、それから費用と効果、それから誰が負担するのかという問題については、レノン先生に特にカナダについてお話ししたい。また、深堀先生に日本の場合について、どのように考えるかということをお話ししていただけないかと思います。それから将来、これがどのように発展する可能性があるのかということについては、もちろんユーウェル先生にはお話しただかねばいけないと思いますし、それから岸本先生に特に日本の場合について、例えば工学部のあとで、これはもうちょっと拡大する可能性があるのかどうかということは、非常に重要になると思いますので、お話しただけなら



金子元久氏

ではワーヘナール先生の方からお一人ずつお願いします。

■ **ワーヘナール氏** わかりました。非常に難しい質問です。なるべく公平に答えさせていただきたいと思います。最初からCLAに関しては、私は懐疑的でした。率直に申し上げて懐疑的でした。なぜかという、ピーターが説明してくれたとおりでありますけれども、テスト問題が一つの文脈、つまりアメリカの文脈でしか開発されておりません。ですからアメリカの文脈であれば、有用性があることが立証されているのかもしれませんが、別の文脈に移した瞬間に、もしかしたら問題が発生するかもしれません。CLAには二つの問題があります。一つ目は、文化間の差異が世界中に存在する問題。二つ目は、冒頭の発表でも申し上げましたが、分野AかBで異なるという問題。一般的技能について、多くの同僚の研究者、学者は納得しているようでありまして、一般的コンピテンスは社会にとって重要だから育成すべきですけれども、それは専門分野の中で育成すべきだと。そして、専門分野において育成するのであれば、同じ文脈の中で測定するのが一番公平な方法だろうという考え方で一致しています。問題提起された「能力の概念化」に関してですが、どんな違いが地域間にあるかは、もっと調査しなくてはならないと思います。チューニングを地域別を実施しているのには理由があって、この仕組みで正しかったと思っています。現在、チューニングの成果を地域別に比較して、何らかの共通項が得られているか、文化的、あるいは教育発達の段階の違いの影響はあるのかを分析しております。その結果が出れば、果たして能力概念を比較できるかが明らかになってくると思います。私は可能だと思います。なぜならば、工学と経済学の両方におきまして、例えば工学の場合、地域間でかなりの共通項があることが既に立証済みです。ですから、展望はある程度あると思います。AHELOの今後について、私はよく分かりません。私自身の私見は確定しています。AHELOはある意味リスクが伴いました。なぜならばそれは共通項を17の場所で見いだそうとしたからです。国によっては、遅れて参加したものもあります。それ自体が問題だったのかもしれませんが。地域ごとで共通項を作り、そして地域間で比較するのがよいと思います。技術諮問グループでアプローチを議論しましたが、こういうステップ・バイ・ステップのアプローチの方が有効なのかもしれません。もう一つのモデルは、一つの地域で手がけてみて、終了すればその地域の自主的な取組に任せて、循環的に他の地域に展開するというアプローチも考えられます。時間はそちらの方がかかるけれども、得るものはむしろ大きいかもしれないと思います。チューニングの視点から、費用対効果について一言申し上げたいと思います。我々は努力して、何とか専門分野の枠組みを設定しようとしてきました。そしてお互い学び合うことができたと思います。枠組みを作っていくことによって、世界中からお互い学び合うことができると思います。たしかに、特定のコンピテンスに関しては、また一部の国・地域の法が他の国・地域よりも有効に開発できることも分かっていますけれども、やはり一層学ぶためには、比較はしなくてはなりません。最終的には、やはり学術界の責任であるべきだと思います。AHELOの課題は、トップダウン・アプローチだったということ。手がけたのが政府だったということ、そして大学が任意に招待されたということです。加えて学生に関しては、ユーウェル氏が示したとおり、実際に参加するかどうかかわからず、信頼性の高い結果が出るかどうか分からないという状況でした。次のステップでは、逆のアプローチはどうでしょうか。例えばより重要な役割、つまり主導権を大学に与えるということです。次のステップが実現するという点に関しては、私は確信しております。競争の存在する世界に我々は生きているわけですから、取り組んでいることの成果を比較できる仕組みを保証する必要があります。ただそのために枠組みはきちんと設定しなくてはなりません。既存の枠組みが十分に正しいものかどうかについては、私はまだ確信を持っていません。



ロバート・ワーヘナール氏

■ **ユーウェル氏** 一般的技能に関してですが、既存のテスト問題で始めましたので、枠組みはそもそもなかったわけです。基礎から少しずつ立ち上げるべきところ、出来上がったものから始めてしまったことほど、トップダウンなアプローチはないでしょう。この失敗から学ぶ、本調査に進む場合には、同じ轍を踏まないようにしなくてはなりません。チューニングやAHELOでは、トップダウンのアプローチと、教員間での議論に根ざしたボト



ピーター・ユウエル氏

ムアップのアプローチが中間地点で融合しているような状況を見て取ることができません。ワーヘナール氏からお話がありました部門別コンピテンスでは、こうした状況が一層顕著です。技術諮問グループにおいて、今後AHELOをどのような仕組みにすべきか、どうやって評価すべきかを議論しました、そして、二点修正するという事で意見が一致しました。一つは、一般的技能を測定するのであれば、専門分野の色調を濃くするという事。CLAのように日常的な文脈から仮説的にテスト問題を立てるのではなく、やはり専門分野や専攻の中で設定するアプローチを今後は尊重していかなければなりません。もう一つ、我々が提言したのは、学問分野にかかわらず、共通して実施する項目も設けることです。そうすることで、専門分野間で具体的にどのように違うのかが明らかになります。「能力の概念化」、あるいは概念の根拠という問題に戻りますけれども、きちんと概念が特定されていない領域をアセスメントの対象とするのは不可能に近いわけです。ですから、チューニングをまず行わなくてはならないわけです。何を測定するのかということに関して、ある程度合意を形成しなくてはなりません。そこで金子先生の最後の質問「AHELOの今後」に話が飛びますが、AHELOで一般的技能はできません。一般的技能は余りに文化的に偏りがありますし、文脈に依存するから不可能です。でも、工学分野ならばどの専攻でも、どの国でもできると思います。工学以外でもかなり多くの専門分野でうまくいくでしょう。ただし、法学、歴史学は難しいでしょう。確かに、実施が難しい専門分野はあると思います。

- 岸本氏 三つとも難しい問題ですけれども、答えられる限り答えていきたいと思います。一つめは工学部で広がるかという話ですが、その話はちょっとおいておきまして、まず学習成果の概念枠組みについて、どんな根拠とということですが、特にエンジニアリング教育の場合は、これまでも学生たちが国を越えて移動し、エンジニアが国を越えて移動するという状況から、この概念枠組みを教育プログラムの認証の中で、きちんとしていくという動きが、かなりありました。そういった意味で、今できているものが完全かどうかということはあるかとは思いますが、大学側の関係者だけではなくて、企業側の関係者、あるいは学んでいる人たち、育てていくエンジニアの人たちの意見も入れながら、教育の中で大切な学習成果は何かという議論を通じて、それがこういうかたちでできてきているということでもあります。そういった意味で、先ほど御紹介いたしましたように、ヨーロッパのグループと、ワシントンアコードのグループとで内容がほぼ同じものになっているというのは、エンジニアに求められている能力というのは何かというのが、ある程度概念として固まってきているということかと思えます。ただ、時代とともにエンジニアに求められている能力も、変わってきていますので、それに応じてフレームワークを進化していくということは大切なので、その意味での議論は続けていく必要があるかなと思います。それで今度はAHELOで、それに基づいて問題を作って試験をするということですが、正しく今回使用した問題で求めている学習成果が測れるようになっていくかどうかという観点からすると、まだまだ時間が足らず、やるべきことが多く残っているように思います。例えば、応用能力ということであっても、どのような問題ならば知識を応用できる能力を適切に計れるのかという課題もありますし、先ほどワーヘナール先生のお話の中にも出てきましたけれども、チームで働く能力というのが非常に大切に言われていますが、個人で試験を受けるようなタイプのもので、本当にそういうようなものが測れるのかということもあります。また、コミュニケーション能力といったときに、エンジニア、あるいはサイエンスをやっている人たちは文字だけではなくて、表を書いたり、式を書いたり、いろんなことで自分の考えを伝えるわけです。そういったもろもろのことを含めて全てを測れるようになれば、本当にそれで測れたということになりますが、まだ、その段階にはなっていないと思います。ただ完全な測り方はできないにしても、こういった取組を続けることによって、よりよい評価方法が生み出されていくようになるだろうなと思います。その一つとしては、学生たちは一つの国で学ぶのではなくて、様々な国で学べるようになってきて、大学を途中で移ったりしているということからすると、やはりどの大学でどのぐらいの教育が行われているのかということ、正



岸本喜久雄氏

しく評価していく必要があるのではないかと考えます。それでAHELOの取組が広がるかということですが、我が国の中でも国際的に競争力のある工学系の教育プログラムを提供していこうとするためには、本当に我々が教えていることがどのようなレベルにあるかということは知りたいし、現在いる学生たちがどのくらいのパフォーマンスを出しているかということを知りたいという状況があります。ただ、このようなことは一つの大学の中でできる作業ではないので、グループを組んで取り組んでいく必要があるだろうと思います。そういった意味で、一緒に今回取り組んできた大学の先生方とも話をしていますと、やりたいという意識はあるように思います。ただ、それが時間の問題だとか、誰がお金を負担するかとかの課題があります。さらに、誰が実際にお金も負担するかもそうですけども、これに関わる人の負荷は非常に大きいものがありますので、そういった意味で誰がこれを実際に動かしていくのかというのは、非常に大きな課題であると思います。また、国内にとどまらず、アジアの国の人たちも、このような動きには興味を持っていて、そういった意味で国際的にやっていこうとする機運はあるのではないかと思います。それを実際にどのように進めるかということですが、今回はOECDが主導というかたちでやってきたわけですが、それはOECDという名前があると参加しやすいということがありますが、やはり大学のグループで例えば地域的に、余りお金のかからないかたちで、実行できるような姿になっていくといいように思います。

■ **レノン氏** ありがとうございます。オンタリオ州での取組について、実践的、ロジスティクスな観点から、費用対効果があったのかについてお話をいただきました。もちろん、フィージビリティ・スタディは本調査よりも少し多くの費用がかかります。カナダは連邦政府ではなく、州が教育を管轄しています。そして多くの州が、参加することを選びませんでした。それは、費用がかかるからです。また、比較するほど多くの高等教育機関を持っていない州もあります。一つしか高等教育機関を持たない州もあるわけです。したがって、実際にかかるコスト、そして実際に何が州に戻ってくるのかというメリットが不明確の問題が、一つの課題でもありました。そうした中で、オンタリオ州は参加することを決めました。そして、全てのコストを負担することになりました。その背景には、オンタリオ州では、OECD-AHELOプロジェクトに留まらず、学習成果に関するいくつかのイニシアチブも走らせており、それらから包括的に学ぼうとしたことが挙げられます。オンタリオ州では、チューニング・プロジェクトにも参加しており、学習成果を大学のプログラムに埋め込み、実践に移す仕掛けとしてAHELOに注目する価値があると考えたのです。オンタリオ州は、この活動に投資することはよいことだと決断したわけですが、州が望んでいたのは、何らかの国際的な比較情報が得られることでした。自分たちを世界の中でベンチマーキングすることを望んだのです。そして、比較のための情報は、AHELOから得られることは明らかだったのですが、OECDはそういった形では提供してくれなかったのです。ただ、日本、オーストリア、カナダではそれは、国際的な比較情報を得ることは非常に重要だと考え、覚書を交わし、自分たちで手に入れようと試みてきたわけです。その結果として得られた比較情報は非常に興味深いものでした。一方、高等教育機関が期待したのは、少ないコストで自分たちの教育についての情報を得て学生にメリットをもたらすことです。オンタリオ州の工学プログラムには、様々な国から来た学生や教員が特に多くいます。ですから、国際調査に参加するのは非常によいことだと考えたのです。したがって、州も教高等教育機関も教員も、自分たちが払ったコストに対して、戻ってきたものに満足をしています。もちろん、高等教育機関はコンピテンス、あるいはコンピテンス・クラスターに関するより詳しい情報、そしてそれらを国際的に比較した情報を知りたいと考えています。もちろん、フィージビリティ・スタディには、改善しなければいけない点がいくつかあります。そしてそのためにはコストがかかることも分かっています。ただ、OECD報告書や機関別報告書を受けて、やる気がなくなったというわけではありませんけれども、望んでいたものがフルに戻ってこなかったという落胆はあります。そういったことも、教員や機関にとっては、メリットがあったと言えると思いますし、こうしたプロジェクトには継続的に参加したいと考えております。ここで問いかねなければいけないのは、州にとって得るものが多かったかということです。3国で自主的に導いた比較情報がなければ、各機関の教育改善活動を支援するという面以外



メアリー・キャサリン・レノン氏

に、州にとってどのようなメリットがあったのかは明らかではありません。もし、より国際比較情報を提供されるのであれば、価値があると考えられるかもしれません。以上です。

■ **エドワーズ氏** ありがとうございます。私からは3点についてお話ししたいと思います。まず「能力の概念化」と「一般的技能」についてですけれども、もうユウウェル先生もワーヘナール教授も、かなりおっしゃいましたけれども、私からは工学分野と一般的技能の枠組みを開発した経験について、逆の方向から説明したいと思います。工学分野の枠組みに関して、岸本先生がおっしゃったとおり、雇用主や学生も関与して、大学教員や研究員も参加して、チューニングの方法でコンピテンスを定義してから、AHELOに移りました。一旦コンピテンスを決めた後に、どういうテスト問題ならば、測定したいコンピテンスが測定できるのかを考えました。一般的技能の枠組みに関しては、設計が違うため必ずしも誰が悪かったということではないですけれども、概念として難しかったと言えます。土木工学の枠組みを見ると、工学分野の一般的技能が組み込まれています。今日は何度もワーヘナール先生がおっしゃっていますが、やはり一般的技能に関しても専門分野の枠組みの中で考えるべきです。例



ダニエル・エドワーズ氏

えば社会学と工学では、一般的技能の在り方は異なります。もっとも、一般的技能の中には、全ての専門分野に共通するものもあります。ただ、使い方が専門分野によって違うかもしれません。ですから、やはり一般的技能は難しい分野ではあるのです。たとえ土木工学のような相対的に明快な専攻であっても、あと10名違った研究者が加わっていたならば、我々が作った枠組みと違った枠組みになっていたと思います。もちろん重複はあったでしょうけれども、何が正しい、何が間違っているという問題はありません。今、医学でも同様のことをやっていて、同様の問題に遭遇しております。枠組みについては、違った意見を持つ人は必ずいるわけですから、本当に限定することはできません。次に「実施上の問題点」についてですが、私はAHELOコンソーシアムの一員として、テスト問題の作成に関わり、オーストラリアのナショナルプロジェクトマネージャーとして実施面にも関わった立場から、ロジ上の問題についてお

話します。学んだ教訓の一つは、ロジ的にAHELOは実施可能だということです。つまり、既にシステムは整備されており、ITも活用できますし、学生記録データも整備されています。教員、研究員も関心を持っています。だからロジ的には実施可能です。一番大きな問題は、いくらかかるかという費用面です。オーストラリアでは、かなり多額の費用がかかりました。政府からOECDにかなりの拠出金が支払われましたし、国内プログラムの実施にも政府補助金が提供されました。でも、お金だけではないのです。時間を初めとする、経済的な資源以外のコストも小さくはありませんでした。高等教育機関内の資源も活用しなくてはなりません。例えば、学生の参加を大学が奨励する働きかけもその一つで、実はそういったことがとても大切でした。ただ単に、学生に座ってテスト受けてもらって、データを集めるだけではなく、よりコミットしてくれる大学の方が、議論という意味でも有用な貢献をしてくれました。つまり、何のためにAHELOを行っているのかということへの大学の理解です。つまり、学生が将来の仕事で役に立つようなコンピデンスを身に付けられるよう保証すること、そこが見失われてしまうときが時々あります。オーストラリアでは、データに関してはかなり長文の記述式解答もあれば、欠落している部分もありました。しかしながら、AHELOは学習成果に関する会話を始める好機となりました。何を学習すべきか、ということについての会話です。コンピテンス枠組みを開発するには、チューニングが参考になりましたし、どのようにして測定するのかということについては、AHELOが参考になりました。「エンジニアのように考える」、そういう考え方ができるようになることが、一つのキャッチフレーズになりまして、教員や研究者、学部長だけではなく、学生の間でも会話を触発するのに役立ちました。特に深く関与した大学では、「授業で学んでいることと、テスト問題で問われることがかなり違う」といった会話が学生の中で頻繁に行われるようになりました。「こういうプロジェクトベースの問題解決型のテスト問題はとて面白い」というような声も聞こえてきました。「AHELOのテスト問題について考える過程で、社会人になったときに何に直面するかということへの理解が深まり、目的にかなった解答が書けるようになったと思う」という声も耳にしました。オーストラリアでは、こういう会話が触発されたことが、一番有用な結果ではなかったかと思います。最後に「AHELOの将来」について考えると、他の方たちもおっしゃっておられましたが、OECDに任せ続けるべきかということ

疑問視する人もいます。つまり国別ベースのモデルを採るOECDを情報発出源として活用するのか。OECDは適切な経路なのか。高等教育機関の調査を管理する主体としてふさわしいのか。そういった疑問が挙げられています。今、手を挙げている機関がお金を探して、イニシアチブを取るという方法も検討すべきです。手を挙げている主体があるからといって、必ずそのイニシアチブがうまく遂行されるわけではありませんが、いずれ確実に実現されると思います。日本は関心を持っているでしょう。オーストラリアは大学のランキングに取りつかれているようですが、既存の大学ランキングはほとんど全てが研究重視です。OECDはAHELOをランキングにしたくないという考えで、私もランキングであるべきでないとは個人的には思っていますけれども、今の比較情報やベンチマークの作られ方は、もっぱら研究ベースです。本来、教授・学習がコアであるべきなのに、ベンチマーク化されていません。大学が教授・学習に力を入れていたとしても、必ずしも既存の測定方法を使ってそれを立証することはできないのです。ですから、AHELOのような取組は続くと思います。対象とする専門分野について、今、オーストラリアは日本やカナダと機械工学で新しい取組を初めています。ドイツとも議論しています。トップダウンよりもボトムアップで協調して作り上げていけるのではないかと思います。オーストラリアにおきましては、かなり成功裏なコラボレーションが医学で実現しております。共通の枠組みを作って、医学部の学生の学習成果を測っております。結果はまだ全ては出ていませんけれども、似たようなイニシアチブがほかの国でも手掛けられていると聞いていますので、それらを比較したいと思います。ですから、医学もできる分野の一つではないでしょうか。もちろん文脈にもよりますが、共通項が見いだせると思います。

■ **深堀氏** はい、ありがとうございます。私からは2番目の「コストと効果」の点に注目して、お話しさせていただきたいと思います。まずコストに関しましては、文部科学省からOECDに支払った拠出金のほかに、先導的学改革推進委託事業の中で、日本の工学及び教育学の専門家の方々に、5年間にわたって研究していただくために必要だった経費、国研の中で問題の翻訳、適正化、テスト実施に取り組むための経費、参加してくださった学生の皆さんに対する参加記念品のための経費等が必要でした。そうした金銭的なコストのほかに、やはり多大だったのは、5年間にわたって専門家の先生方に関わり続けていただき、大変な労力を使っていた点を強調しなければなりません。この場を借りて、厚く御礼申し上げます。そういった多大なコストの中で、どんなメリットがあったのかという点ですけれども、まず私は国内コーディネーターとしてこの活動に5年間関わってきましたが、その中で、岸本先生からも詳しくお話がありましており、コンピテンス枠組みがどのようなものなのか、それについて日本は国際社会と合意できるのか、日本の既存の問題は国際的な枠組みとどのぐらい整合性があるのか、それぞれのコンピテンスを具体的な学習成果に落とし込んで、テスト問題を作り上げていく作業がどのようなものなのか、採点にあたっては解答のそれぞれについて学生が具体的にどこまで記述できていれば正解とみなすのか、どのような観点と水準まで書き込めれば学士としてふさわしいといえるのか。これらの点について、専門家の方々と検討を深めてきました。こうした経験そのものが、非常に有益なベネフィットではなかったかと思えます。関わってくださった先生方も、願わくば、AHELOの経験から御自分の教育実践を振り返られ、非可逆的な、もう元には戻れない変容を経験されたのではないかと期待しています。残念なのは、その先生方がそれぞれの大学に戻って、AHELOの経験がとても役に立ったということ、同僚の先生方にお示ししていただくための資料を十分に提供できていない点です。例えば、テスト問題がどのような内容だったのかということに関しては、一部の問題や採点ルーブリックしか公開されていません。守秘義務のもとで多くの部分が隠されているのです。本来、教育改善のための取組なので、テスト問題や採点ルーブリックを全て公開して、議論の対象にできるように提供したいわけです。どういう観点と水準で採点を行ったのか、具体的に各大学で話し合うための資料を提供して、FDの材料にしていきたいわけです。さらに、国際ベンチマークを伴うデータ分析の結果を、それぞれの大学で活用していただけるよう、報告書を出したいわけです。関わってくださった先生方が、それぞれの大学で胸を張って、面白い取組だったからうちの大学でも真剣に考えてみようと思掛けをしていただけるための材料を、まだ提供できていない点は非常に大きな課題だと思います。逆に、そうした



資料が出せた場合には、AHELOの取組は大学にとってかなりメリットになるのではないかと思います。AHELOにしてもチューニングにしても、大学の教育を改善できるのはやはり大学教員の先生方一人一人の力なくしてはあり得ないわけですから、それを支援していく体制を整えることが重要なのではないのでしょうか。文部科学省ができることは、そのためのアクションスキームを提供すること、こういった枠組みの中でやってみませんかというリーダーシップを取ることかもしれません。それぞれの大学が、競争的環境の中で孤立して取り組むのではなく、共通の枠組みの中で知恵を出し合ってみましょうと提案すること、そして、そうした共同の取組に補助金をつけていくことではないかと思います。そして、国立教育政策研究所のような公共の機関ができることは、コーディネーター役として専門的な知見、経験を提供していくことではないかと思います。アクションスキームとリーダーシップ、そしてコーディネートがあれば、それぞれの大学が共同の枠組みの中で、それぞれにとって意味のある教育改善に取り組んでいける状況が作れるのではないかと考えます。AHELOの将来が、その方向を採ることを願っております。以上です。

■ **金子氏** つけ加えることがあればどうぞ。

■ **ユーウェル氏** 私どもの今の応答に共通するテーマの一つは、このようなデータからは、改善のための会話を始めるべきだということです。教育改善のための議論は、AHELOフィージビリティ・スタディの議題として意図的に設定したものではありませんでした。ほとんど偶然に、うまく始まったわけですが、やはり続けていかなければなりません。その差異、議論を始めるためのデータは、必ずしも技術的に堅牢なものでもなくてもいいのです。もちろんできるだけ健全で、適切なものであるに越したことはありませんが、その一方で、PISAのようなデータ、すなわち国別の推定値を導くようなデータを収集するコストは莫大なわけです。恐らくワーヘナール氏も同じ意見だと思いますが、まずよい議論を始める、専門分野の文脈の中でよい議論を始める、そして地域から始める、というのが正しい方法ではないかと思います。皮肉なことに、OECDはそのようなプロセスを今まで一度もやったことがないのです。OECDの経験とは、大規模なテストを実施し、国別に大まかな推定値を出すというものなのです。今回、高等教育機関に対して同じようなことをしようとしたのですが、それでは決してうまくいかなかったわけです。

■ **金子氏** 先ほど深堀先生の発言にもありましたが、今、ユーウェル先生もおっしゃったように、これ自体が成功だったかどうかは別として、かなりここからいろいろと議論は出発するところはあるわけです。ただ、この枠組み自体がかなり国際機関を通じたもので、実はデータ自体も相当使いにくいといえますか、そこの制約も今のところ多くて、既にあるデータを十分に使うことも、なかなか今は難しいことでもあると思います。それを含めて、これからどのようなことをやられるおつもりなのか、そういった計画があるのであれば、是非お話しいただければ。先ほどエドワーズ先生からもオーストラリアはやっておるといってお話もありましたし、深堀先生からも岸本先生からも日本の活動を御紹介いただきましたけれども、どういったことがこれから、少なくともこれまでのデータを使ってやろうとされているのかを、短く御紹介いただければと思います。どなたでも結構です。

■ **深堀氏** はい。今回、レノン先生とエドワーズ先生に来日していただいております、明日AHELOに参加していた大学のクローズドなミーティングを開きます。その中で、それぞれの国でAHELOがどのような経験だったのかということについて議論することを計画しております。本日司会を務めてくださっている、濱中研究官がデータ分析を進めてくださっていて、クローズドな私たちで各大学にできる限りのデータを提供しようと考えております。また、今年3月末には、国研の中で行っている「調査結果の分析に関する研究会」の研究成果報告書をまとめて公開いたします。その中でデータ分析の結果をどこまで書き込めるかは、これから検討しなければなりません。全国の各大学において、教育改善に取り組んでいく際に参考になる資料としてまとめていく予定でございます。

■ **金子氏** エドワーズ先生、先ほど少しお話しいただきましたけど、特にコンテクスチュアル・バリエーションについ

ては、まだほとんど分析はされてないんですけど、そこらはどのように進められるのでしょうか。

■ **エドワーズ氏** ありがとうございます、金子先生。我々は現在、日本とカナダと協力して、合体データセットの分析を進めており、背景情報調査について作業を始めたところです。この作業と関連して、いくつかの計画がありますが、一つは現在オーストラリア教育研究所ACERにおいてまとめようとしている論文があります。深堀氏は既に面白いデータを発表してくれておりまして、それは時間数に関するものです。「専門分野に関係のないアルバイト」の時間数について、我々は更に分析して、どんな要因が差異をもたらすのかを特定しようとしています。これまで分析してきて興味深いと感じているのは、多変量解析で多くの要因を統制すると、大学での成績は、AHELOのテスト得点と強く関連していることが分かります。つまりAHELOのテスト問題には一定の併存的妥当性があると言えそうです。また、明日の会議でも話すつもりですが、オーストラリアのデータを見る限りにおいては、専門分野のアルバイトをやっている、既にエンジニアとしてアルバイトをしている学生とAHELOのテスト得点の間には正の相関があるようです。それは時間数の上限を超えると余り有効ではなくなりますが、その上限までは、アルバイトの時間数が長いほど、AHELOテスト得点も高くなっています。もう一つの知見は、お手元のペーパーの中でも触れていますけれども、学生たちが卒業後どんな職に就くつもりかについて、エンジニアになろうとして勉強している学生の方が、この分野で職を求めていると回答した学生よりも、AHELOテスト得点は高かったことです。これはまだ出発点ですが、今後、こうした分析を積極的に手掛けていきたいと思っています。

■ **金子氏** ユーウェル先生どうぞ。

■ **ユーウェル氏** エドワーズ氏と私は、それぞれの国の結果を比べてみたのですけれども、つまりアメリカの一般的技能とオーストラリアの工学分野と比較してみました。すると、異なる分野であるにも構わず、結果のパターンにいくつかの共通性を見ることが出てきました。確かに、データから何か面白いことを明らかにできそうなのです。例えば、学生による学業成績の自己報告、テストに要した努力についての自己報告は、テスト得点と強く関連していることが分かりました。その一方で、卒業後に何をやるつもりか、そのやろうとしていることについてAHELOテスト問題の内容はどれほど妥当かという問いに対する学生の回答も、テスト得点と強く関連していました。正課の中での経験にも、テスト得点と有意な相関を持つものがありました。したがって、AHELOの成果として、教授・学習プロセスの影響について探究する確かな手ごたえを見つけることができたと言えそうです。

■ **岸本氏** 今回は問題を作って、ただ何点取れたかだけではなくて、今日お話があったように背景情報を丹念に取っていて、その背景情報と得点との関係がどのようになるかという分析が進められています。大学関係者も、そのような分析結果を知りたいと思います。また、こういう教育をしたらこういう成果が出ているということが、正に客観的データとして出てくるのが非常に大切だと思います。それと、より強く期待していることは、大学側が教育のやり方を変えたときに、この結果がどう変化するのかということを実は本来知りたいわけで、それに対する貢献です。それには1回やっただけでは本当のデータがまだまだ取れていなくて、このあと数年後に同じようなことを繰り返し実施してみてどうなるかとかといった継続活動が求められます。そのようにすることで、こういったデータが生きていくのではないかと考えます。このような活動を活かすためには、きちんとデータを取り続けることが必要であると思います。

■ **金子氏** そろそろ終わりにしたいのですが、もう一つ、カナダあるいはアメリカあるいはEUで、このAHELOの経験が何か別の動きにつながってくるというようなことはあるのでしょうか。あるいは今、行われている動きに何か影響、どのような影響を与えたいと思いますか。

■ **ワーヘナール氏** それにお答えしてみようと思います。実際のところ、少なくとも世界的な規模では、余り大きな動きはありません。しかしながら大学間で、これは重要な取組だとの認識が高まってきたと思います。ヨーロッパ

パにおいても、そのように受け止められてきています。お気づきになったかと思いますが、ヨーロッパの高等教育機関はAHELOに余り参加しませんでした。OECDはもともと欧州の機構であるにも関わらず、参加していないことは驚きでもあるわけです。いくつかの高等教育機関に話を聞くと、AHELOに関心を持っているのは確かです。しかしながら、現在の取り組み方は、持続可能ではないようです。ほかの多くの関係者も論じているとおり、持続可能性の基盤は、ボトムアップのアプローチであって、トップダウンではいけないというのです。大学も教員・研究者間も、AHELOのメリットを最初から見いだしているはずですから、ユーウェル氏が言ったように、取組の最初から取り込まなければならなかったのです。それが実現すれば、人々は同僚との対話から学ぶことに気付くことでしょう。その枠組みが地域なのか、世界なのか。学生も関心を示すでしょう。テスト問題が学んでいることに対応していれば、もちろん学生たちも受けたいと思うはずですから。結果がよくなければ捨ててしまえばいいわけで、よければ将来の雇用主に見せたいと思うかもしれません。したがって、人々の関心を集める方法はいくつかあります。その中で、比較の要素も重要です。高等教育機関は、研究の観点からだけではなく、教授・学習の観点からも比較したいわけです。そしてこのような調査は、継続することが論理的に正しいと思います。したがって、新しい枠組みを作るかもしれません。この枠組みは、この大学のコンソーシアムになるかもしれません。OECDが補助金を出してくれるような枠組みにできるかもしれませんが、OECDには資金がないので、それはちょっとおもしろいかもしれませんが、補助金は、各国の政府や欧州委員会から得られるかもしれません。あるいはルミナ財団などの民間組織から得られるかもしれません。そうした民間組織は、ドイツにもアメリカにもありますし、他の国にも存在します。その方が、これまでよりもよいアプローチかもしれません。ただ、これはOECDがイニシアチブとして手掛けてくれたおかげなのです。OECDが初めてくれなければ、この取組の価値は分からなかったでしょう。それゆえ、OECDに対して感謝の意を表する必要があると思います。

■ **ユーウェル氏** 同感です。OECDに我々は借りがあります。これを始めてくれて感謝すべきです。ただ、このような形成的評価の取組を今後も維持していく上で、OECDが望ましい組織かどうかということに関しては、まだ解答は出ていないと考えます。取組の目的は、厳密な方法論を採ることではありません。そうではなく、十分によいデータを収集し、それに基づいて、改善について議論を促進することです。地域的なコンソーシアムは、よいアイデアだと考えます。例えば、コロンビア、エジプト、メキシコといった国々は、AHELOに積極的に参加して、ほぼ100%の学生参加率を達成しました。これらの国々は今、データの分析を初めています。その中でメキシコとコロンビアは、中南米のイニシアチブとして、データを共有して分析する共同研究を手掛けています。もともとOECDの計画の中では、それは想定されていませんでした。でも実現したのです。なぜかという、需要あるからです。AHELOの将来を予測するならば、少数の国がまとまって、独自のイニシアチブを始めていくでしょう。そして、第三者の独立財団からの援助を得るとというのが、望ましいアプローチだと思います。同一ではないけども、似たような枠組みの下で、です。ありがとうございました。

■ **レノン氏** AHELOは巨大なる潮流の一部だと思います。学習成果を確立して、実践して、測定する取組は、高等教育機関、政府、専門分野、学位プログラムの枠組みでなされています。それだけ様々なレベルで関心があるということですから、学習成果調査を異なるレベルで実施するという正に好機だと思います。カナダの工学プログラムは、近年、卒業生に期待されるコンピテンスと学習成果を定義し、それを実践するだけでなく、2016年までに測定して挙証することを義務付けられています。そのような中で各工学プログラムでは、アセスメント・ツールの独自開発に件名に取り組んでいます。AHELOへの参加意欲、会話に参加したいという希望も、その一環として高まりを見せています。日本やオーストラリアとの機械工学のアセスメント開発にも参加の意欲を示しております。そういった様々なイノベーションが一緒になりつつある中で、国際協調によって、概念的にも発展するチャンスがあると思います。OECDは声高なる会話をする場にはなっていますが、唯一無二でもないと思います。もっと小規模の革新的な活動を推進してもいいかと思っています。

■ **ダニエル氏** 手短かにコメントをさせていただきたいと思います。トップダウンに対するボトムアップアプローチという話がありましたけれども、オーストラリアにおきましては、このAHELOに工学分野で参加してきました。

そして医学の分野においても、同じようなプログラムがあります。AHELOでは、初めに政府がOECDの枠組みに参加することを決め、オーストラリア教育研究所がテスト実施の委託を受け、大学に対して参加を募りました。各大学からは、かなりよい参加率が得られたものの、アセスメントの在り方を決定する過程に、大学が大きく関与することはありませんでした。率直に申し上げて、オーストラリアでそういうアプローチを採った場合、余り多くの参加者を得ることができないと思います。やはり、大学側からのコミットメントが、非常に重要だと思います。その一方で、医学プログラムの取組みでは、大学の側からオーストラリア教育研究所に対して、こういう試みをしてみたいのだと相談があったのです。「そのための専門性がないので、あなたの組織と組ませてもらえないか」と持ちかけられました。オーストラリア教育研究所は、二つの医学部と2か月ほど取組ましたが、その過程で大学側から、研究者、臨床医、自主的に関心を持った関係者等、非常に多くの人が集まりました。少し経った後に、実際にアセスメントを実施してみようということになりましたら、結果的に16の大学が参加することになり、大学間の共同イニシアチブとして進められることになったのです。そうした状況が、大学間で共有する一般性の高いアセスメントを可能にしたのです。これは小さな事例ではありますが、大変重要な成功事例だと思うのです。こういうボトムアップのアプローチこそ、うまくいくのだと教えてくれます。今後は、こうした取組を地域的に広げていきたいと考えております。まずは小さく始めて、それがうまくいったら広げていくというやり方を進めたいと思います。医学プログラムの取組では、テストを受けた全ての学生に対して、結果の報告書を提供しています。AHELOの工学の取組では、それを行うことができませんでした。工学の分野におきましては、テストを受けさせること自体が難しかったのです。トップダウンの工学の場合と、私たちが取り組んだ小規模な医学の場合では、違うやり方を採ったことで結果が大きく異なったわけです。

■ **金子氏** 個人的な感想を言わせていただければ、先ほど言われたように、日本の大学教育改革も、大学間の教育の質をめぐる競争が重要になってくる段階に来ていると思います。そうだとすれば、その大学が何をできているのか、それから個々の先生がどういった違いを学習のアウトカムにもたらしているのかということが問題になる。したがって、何らかのかたちで学生が何を学習しているのかというのを測ると言いますか、余り抽象的なかたちだけではなくて、具体的にどう学習しているのかを測ることが非常に重要にならざるを得ない。このAHEROのプロジェクトはそういう課題に国際的な共同作業を通じて応えるという、非常に野心的な試みとして始まったことです。ただ実際にやってみますと、やはりかなり今日お話にありましたように、テストの概念とか、あるいは方法とか様々な問題があるということが改めてわかってきたわけです。ただこれで全てが終わりというわけではありません。OECD全部が参加しなくても、数箇国が自主的に参加すれば、さらに作業を進められることで意味のある成果を得ることができるかもしれません。あるいは国内でも数機関が参加して互いに比べるということも、非常に大きな意味があるのだということも、まただんだんこの試行を通じてわかってきたと思います。そうした意味で、今回のお話は日本のこれからの高等教育改革が、さらに具体的な段階になっていく上で、非常に重要なお話だったのではないかと思います。最後に全ての方の質問に代表できなくて大変申し訳ありませんでした。ただ、また質問へのお答えは、大変面白かったと思いますし、刺激的だったと思います。改めまして参加された先生方に拍手をしたいと思います。どうもありがとうございました。サンキューベリーマッチ。

閉会挨拶

杉野 剛
国立教育政策研究所次長

- 報告書作成に当たり、当日の発言内容に修正を加えていることがあります。
- 所属団体、職名は2013年12月10日現在のものであります。

閉会挨拶

国立教育政策研究所次長
杉野 剛



皆様おつかれさまでございました。本日は御多用の中シンポジウムに御参加いただきまして誠にありがとうございました。各大学、民間団体から多くの皆様方に御参加いただき、更には在日の外国大使館の関係者の方々にも御参加をいただきました。ゲストの先生方からは大変興味深いお話を伺うことができました。

最初に木村孟先生からは過去30年間にわたります大学改革の流れと、大学評価の課題を俯瞰するお話をいただきました。その見取り図の上にこのシンポジウムを位置付けて進めることができたと思っております。

ローベルト・ワーヘナール先生からはコンピテンス枠組みに基づく学位プログラムを構築するチューニングの具体的な方法につきまして、最新の情報を御紹介いただきました。

ピーター・ユーウェル先生からはAHELOの全体像を御紹介いただきますとともに、その結果と課題、更にはこの調査の今後の育成について貴重なお話をお伺いすることができました。

岸本喜久雄先生からはAHELOの工学分野におけます日本での取組についてお話をいただきました。こういった国際的な調査につきましては、単に調査に参加するだけではなく、その調査問題の開発の段階から参加してこそ大きな成果が得られるということを痛感した次第でございます。

深堀聡子先生からは、今日も御参加いただきましたメアリーキャサリン・レノン先生、ダニエル・エドワーズ先生とともに取り組んでおりますカナダ、オーストラリア、日本の3か国の調査結果の分析に関する報告がございました。いろいろと制約がある中ではございますけど、この3か国の取組には大きな意義があるのではないかなと考えております。

こういった御講演、報告を踏まえましてのパネルディスカッションでは、更に率直と言いましようか、金子先生のお言葉では刺激的な御意見、あるいは貴重なお話をお伺いすることができました。論点を整理いただき、実り多い議論を深めていただきました金子元久先生に御礼を申し上げます。TUNING-AHELOは我が国の大学政策、あるいは各大学での教育実践に重要な示唆を提供していただくものだと考えております。木村先生のお話にもありましたけれども、日本の大学教育はどうやら日本のウィークポイントだということでございますけれども、このシンポジウムを契機に日本の大学教育の改善が進むことを祈念いたしまして、また、本シンポジウムに御参加いただきました全ての皆様方に厚く御礼を申し上げまして、閉会の挨拶とさせていただきます。ありがとうございました。

資料

AHELOの経験：

オーストラリアにおける実施，成果及び教訓*

ダニエル・エドワーズ

オーストラリア教育研究所主任研究員

AHELO：

オンタリオ州高等教育機関の経験*

メアリーキャサリン・レノン

オンタリオ州高等教育質保証カウンシル上級研究アナリスト

登壇者プロフィール

- *印が付いている原稿は，英語の原稿を仮訳したものです。
- 所属団体，職名は2013年12月10日現在のものです。

AHELOの経験：オーストラリアにおける実施、成果及び教訓

ダニエル・エドワーズ博士

オーストラリア教育研究所主任研究員，及びオーストラリアAHELOナショナル・プロジェクト・マネジャー

はじめに

オーストラリアは経済協力開発機構（OECD）による高等教育における学習成果調査（AHELO）のフィージビリティ・スタディに参加した。本稿はそれに関する背景、実施プロセス及び成果を提供するものである。オーストラリア技術革新産業科学研究高等教育省（The Federal Department of Innovation, Industry, Science, Research and Tertiary Education, DIISRTE）は、OECDへの直接的拠出金を通じて、そしてオーストラリア国内大学での調査実施を担当するナショナル・センターへの財政措置を通じて、オーストラリアの本プロジェクト参加に資金を供給した。

AHELOフィージビリティ・スタディは、OECDによる大型プロジェクトである。その目的は、複数の文化・言語を越え、また各機関の立場やミッションの違いを越えて、有効な方法で学習成果を測定するための確固たるアプローチを見いだすことであった。本プロジェクトに関するねらいや最終報告等の詳細については、<http://www.oecd.org/edu/ahelo>を参照されたい。

AHELOフィージビリティ・スタディでは、一般的技能、経済学及び土木工学の3分野において学生の試験を行った。本フィージビリティ・スタディのため、オーストラリアは土木工学分野に参加した。この分野に参加したことが本章の内容の中心となる。こうしたテスト問題に加え、参加した全教育機関の学生、教員、スタッフに対して背景情報の側面に関するアンケートが実施された。

本フィージビリティ・スタディの実施は二つの段階に分かれる。第1フェーズは2011年に実施された。このフェーズでは、教育機関の参加促進、参加大学におけるフォーカス・グループの調整、テスト問題の草案の学術的検討、学生の回答の採点、そして教育機関及び学生からのフィードバックの照合及び連絡調整が行われた。第2フェーズは2012年に実施され、大学の再募集、学生及びスタッフのサンプリング、AHELO調査の準備に係る技術支援、テスト問題及びアンケート調査の実施運営、テスト問題の解答の採点及び結果報告が行われた。

本稿は、特に学生関連のデータに重点を置きつつ、オーストラリアにおけるAHELOフィージビリティ・スタディのプロセス、活動、成果の周知に関する洞察を提供する。最初に、AHELOの概念及び本プロジェクトに関する実用性についての背景説明を簡潔に行う。次に、第1フェーズ及びテスト問題草案に関するフィードバックの回収から始まる、オーストラリアにおける本研究の実施プロセスを検証する。その後、参加者数に関する洞察と、注目には値するものの依然大まかな調査結果の報告とともに、AHELOのテスト間の実施（第2フェーズ）が詳述される。本稿の最終節では、オーストラリアの視点から本フィージビリティ・スタディの全体的な成果について議論し、普及及び報告の代替的アプローチの可能性を考察する。

AHELOの背景

本節ではAHELOプロジェクトの背景について述べる。本節はAHELOコンソーシアムのメンバーと共同して執筆され、本研究の目的に対する洞察を提供する。また、本稿で詳述される取組の枠組みについての全体的な視点についても触れる。フィージビリティ・スタディの全体的な目的及び意欲は重要であるが、時としてこの取組におけるプロセスの詳細な論議の中で失われてしまうため、本節は達成しようとしていたことの「全体像」を振り返る機会を提供する。

最終学年の学生が自身の学位プログラムから習得した知識及び論理的思考を使用し、応用し、それに基づき行動する能力に対して、国際的アセスメントを実施することは可能なのか。そしてその結果を効率的かつ国際的に比較可能な方法で評価することは可能なのか。高等教育の学習成果のアセスメントが教育過程における重要なチェックポイントであることを、政策立案者、教育機関の長、教職員及び学生に納得させることはできるのか。これらの問いがOECDのAHELOフィージビリティ・スタディの根幹にある。

AHELOでは一般的技能、経済学及び土木工学という三つの中核領域におけるテスト問題の開発及び検証と同時に、テスト結果データの解釈を補助するための背景情報調査ツールの開発が行われる。テスト問題は学士号の最終学年の学生を対象としており、彼らのスキル及び知識を現実世界の課題に応用する能力を測ることを目標とする。AHELOは、高等教育における学習成果のアセスメントのための新しい方法及び技術標準を開発することを目指すという点で野心的なプロジェクトである。これは現在世界規模で実施されており、17か国がテスト問題の開発及び検証に参加し、世界中の専門家、教育機関、政府及び主要な高等教育団体が関与している。プロジェクトは、オーストラリア教育研究所が主導する国際機関のコンソーシアムによって運営されている。

AHELOは重大な情報格差に対応している。指導の質を改良したり学生の学習成果を高めたりしようとする取組は、様々な教育機関や国の学生の能力、あるいは指導の質についての比較判断を可能にするような信頼できる情報の欠如により妨げられる。学習及び指導といった中核的な高等教育活動に関してこのようなデータがない場合、高等教育機関が置かれる状況はおおむね評判や研究実績に基づくものとなる。AHELOの目的は、指導の実践及び学習成果をある程度重視するテスト問題の設計や試行を通じて、情報源を充実させることである。

AHELOのアセスメントは世界中の専門家の共同作業によって開発され、その後全ての参加国で使用するために翻訳され適応されている。参加国の学生を対象としたアセスメントの検証は2011年で完了した。AHELO向けに開発されたテスト問題は、2012年に17か国で実施された。

本研究は高等教育及びアセスメントについて行われている議論の在り方に影響を及ぼしており、AHELOには重大な点で高等教育の展望を再構築する可能性があることを示唆している。第一に、一部利害関係者間の議論は、「学習成果を測定すべきかどうか」から「どのように測定すればよいのか」に移ったように見える。同時に、(OECD内外の)国家システム及び教育機関の関与がフィージビリティ・スタディの期間を通じて顕著に高まった。これらを総合して考えると、改善を知らせたり、質を明示したりするためにも、指導や学習の質に関するデータに対して要望があることを示唆していると言える。

オーストラリアにおけるAHELOの推進

オーストラリアのAHELOへの参加に当たっては、連邦政府による資金提供と支援を受けた。政府は、ナショナル・プロジェクト・マネジャー (NPM)、出張に係る若干の基本的支援及び手当を含め、ナショナル・センターの組織化を促進した。オーストラリアのNPMは、AHELOコンソーシアム及び本研究のOECD事務局の指示の下、基本的にはAHELOフィージビリティ・スタディの全ての側面の実施に責任を負った。NPM及びナショナル・センターの重要な活動の多くは、第1フェーズ及び第2フェーズの活動を扱った本報告の各節に詳述されている。この取組のうち、以下の節で余り取り上げていない点の一つは、プロジェクトの期間を通じて必要とされたコミュニケーションの要素である。この点についてここで簡単に触れておく。

NPMは、フィージビリティ・スタディの全期間にわたり技術革新産業科学研究高等教育省 (DIISRTE) を通じて連邦政府と緊密かつ頻繁なコミュニケーションを維持した。プロジェクトの繁忙期中には月次経過報告を提供し、定期的に電話会議を行った。経過報告及び議論はDIISRTEに対しNPMの日々の活動について継続的に知らせるとともに、本プロジェクトが大学、大学主導者、研究者及び学生にどの程度受け入れられているかについての洞察の提供を続けた。

ナショナル・センターも、参加教育機関との関係及びコミュニケーションの構築に重要な役割を果たした。本研究中の各教育機関との接触及び協力は、電話会議、シンポジウム、現地訪問及び電子メールによって維持された。NPMによって維持された教育機関との関係のうち重要なものは、大学内における活動を担当した各教育機関コーディネーターとのものであった。こうした人々との人脈を築くことは、プロジェクトの重要かつ非常に価値ある側面であった。

本稿の他の節で余り扱っていないNPMのもう一つの重要な役割は、AHELOフィージビリティ・スタディの実施において、オーストラリアの国益が反映されるためのルートを提供した点である。オーストラリア政府は、テスト問題の開発及び適正化、並びにデータ収集の際に採用されるプロセスや手続において、オーストラリアが確実に重要な貢献ができることを重視していた。

AHELOの実施

本節では、本研究の国内実施の観点からAHELOフィージビリティ・スタディの全体概要を示す。フィージビリティ・スタディに参加した国及び教育機関内の実施には二つの局面があった。第1フェーズでは、アンケート調査及びフォーカス・グループ（focus groups, cognitive labs）と呼ばれるグループ・インタビューに基づくテスト問題項目に関する質的調査が実施され、本研究向けに構築されたテスト問題に関連するフィードバックがOECD及びAHELOコンソーシアムに提供された。第2フェーズでは、大学内で学生を試験状況下に置いてAHELOのテストが実施され、教員及び高等教育機関を対象としたアンケート調査が行われた。

フィージビリティ・スタディの第1フェーズ中に、多くの国でテスト問題草案の質的検討が実施された。学生は、フォーカス・グループの形で、テスト問題に関するフィードバックを提供した。一方、土木工学分野の多くの上級研究者はAHELOテスト問題草案についてコメントし、検討内容を提供するよう依頼された。まとめると、第1フェーズの重要な側面は以下のとおりであった。

- 大学への説明及び募集：大学の関心及び参加の確保。
- 参加大学に対する支援：教育機関コーディネーターのためのマニュアル、ディスカッション及びトレーニングといった参加に関する詳細情報、並びに継続的な連絡及び支援の提供。
- フォーカス・グループの実施：教育機関コーディネーターの指揮の下、各大学で実施された。テスト問題草案の試験実施、アンケート調査及びフォーカス・グループによるフィードバック提供には学生も関与した。
- AHELOテスト問題の学術的検討：土木工学教育の専門家によるもの。
- 学生及び大学関係者からの解答及びフィードバックの照合：NPMによって計画された。テスト問題解答の採点、調査の照合及びコーディング、全国データベースの構築等。
- テスト問題に関するフィードバックの提供：AHELOコンソーシアムへのフィードバックのための解答の分析及び要点の照合等。

オーストラリアにおける第1フェーズの活動は2011年に実施され、10大学が参加した（詳細は本稿の「成果」の節を参照）。

第2フェーズでは、学生に対するAHELOテスト問題のオンラインでの大規模実施を行った。教員及び高等教育機関コーディネーターへのアンケート調査も行われた。

まとめると、第2フェーズにはオーストラリアの参加に当たって、以下の活動が行われた。

- 教育機関の募集：大学の関心及び参加の再確保。
- 教育機関向けマニュアルの開発：オンライン試験実施と学生募集のための支援。
- 教育機関コーディネーターの研修及び支援：確実に全ての参加教育機関がAHELO試験システム及びAHELO試験実施手順を熟知しているようにした。
- 教育機関のサンプル抽出：参加教育機関における適格な学生及び職員に関する母集団データの収集。
- 試験用コンピュータの設置：試験会場の安全なオンライン環境確立に関する助言及び情報の提供。
- 試験の運営：参加大学における学生へのAHELO試験の実施。ルールに基づいて試験が実施されていることを確認する実施体制のモニタリング。
- アンケート調査の運営：特定されたサンプルに対する教員用・高等教育機関用アンケート調査の送付。回答数を最大化するための追跡（再依頼）実施。
- テスト問題解答の採点：採点リーダーの採用、採点トレーニング、専門家集団による採点の支援等。
- データ処理及び国際分析の支援：AHELO調査ルールの固守を確実にし、データの国際比較可能性を最大にすること。
- 国際研修及び普及会議への出席：他国における洞察を役立てるとともに、オーストラリアの背景状況から得られた教訓の価値を最大化する。

オーストラリアにおける第2フェーズは2012年中に実施され、3月から5月の間に8大学で試験が行われた（「成果」の節に詳述）。

教育機関の募集

OECD及びAHELOコンソーシアムからの要請に基づいて、オーストラリアは大学10校のフィージビリティ・スタディ参加を目標とした。こうして2011年の初めに、オーストラリアの10大学の大学副総長及び工学部学部長に書面で参加を要請した。招待された全10校の教育機関が本研究の少なくとも第1フェーズに参加することに同意した。

大学の選択に当たっては、NPM, DIISRTE, オーストラリア工学部長評議会 (ACED) 間での連絡調整が行われた。各大学は以下を基準として選択された (選択肢の優先度順)。

1. 以前から (政府との非公式討論を通じて) AHELOに対する関心を積極的に示していること。
2. 教育雇用関係省 (DEEWR) に教育機関コーディネーターの連絡先の詳細をあらかじめ提供していること。
3. アクティブに活動している土木工学プログラムが存在していること。
4. 「サンプル」に幅広い地理的位置の大学が含まれていること。
5. 「サンプル」に様々な教育機関の「グループ」(すなわち、研究インセンティブ、「新しい」大学、地方の大学等)からの大学が含まれていること。
6. 「サンプル」に様々なサイズの大学が含まれていること。

各大学がフィージビリティ・スタディへの参加を約束すると、NPMから連絡があり、プロジェクトの概要を説明された。教育機関コーディネーターが指名され、NPMから直接連絡を受けてプロジェクトへの参加についての詳細の説明を受けた。

2年のプロジェクトの間、11の異なるオーストラリアの大学から学生が参加した。これらの教育機関は以下のとおり。星一つ (*) の教育機関は第1フェーズのみ参加した。全体としてこれらの教育機関はその後の参加を妨げる主要な障害として資源の不足を挙げた。星二つ (**) の教育機関は、2012年の初めにNPMから招待された後で、第2フェーズのみ参加した。その他全ての教育機関は、フィージビリティ・スタディの両方のフェーズに参加した。

- チャールズ・ダーウィン大学
- カーティン大学*
- ジェームズ・クック大学
- スウィンバン大学
- アデレード大学
- メルボルン大学
- ニューサウスウェールズ大学*
- ニューカッスル大学*
- シドニー工科大学
- ウェスタンシドニー大学
- RMIT大学**

成果

本節では、AHELOに参加しているオーストラリアの教育機関の学生の成果について幅広く議論する。議論の内容は比較的幅広く、AHELOアセスメントの実際の結果よりも参加レベルの話により重点を置いている。

質的検討：第1フェーズ

第1フェーズは、2011年4月及び8月にオーストラリアの10校の参加教育機関において実行された。全体で78名の学生がフォーカス・グループに参加し、テスト問題草案のいくつかのセクションを選択解答後、アンケート調査の回答を完了し、フォーカス・グループ設定でフィードバックを提供するよう求められた。

この試験に使用されたテスト問題草案は以下のとおりであった。

- 記述式問題：学生はシナリオ又は問題を示されて、これに関して多数の問題に答える。ほとんどの場合、これらの問題は「自由記述」(つまり選択式でない) による回答を要求する。

- 加えて多肢選択式問題一式。

全ての試験は、参加的大学の中で教育機関コーディネーター又は学部の指定された教員によって運用された。試験は約1時間継続し、土木工学の学士号の最終学年の学生を対象として行われた。表1は、第1フェーズのフォーカス・グループに参加した学生の人口統計上の特徴を示す。

表1: フォーカス・グループの統計データ

統計データ		(n)	%
性別	男性	60	76.9
	女性	18	23.1
年齢	20 歳	4	5.1
	21 歳	18	23.1
	22 歳	26	33.3
	23 歳	12	15.4
	24 歳	9	11.5
	25 歳	3	3.8
	26-30 歳	3	3.8
国籍	オーストラリア	54	69.2
	外国籍	23	29.5
母国語	英語	52	66.7
	その他の言語	25	32.1
学生総数		78	100

試験に参加した後のモデレーターが進行する議論では、学生は以下の論点を重点的に扱うよう求められた。

- 課題は考えることへの挑戦を求めるものだった (挑戦)
- 資料は課題への関心を刺激するものだった (関心)
- 課題は知識やスキルを現実社会で使うように応用させるものだった (応用)
- 課題は適切な範囲の知識やスキルを測定するものだった (範囲)
- 課題は私の研究プログラムと関連性のあるものだった (プログラム)
- 課題は将来の専門実務と関連性のあるものだった (将来)

学生の回答は採点され、AHELOコンソーシアムに提供された第1フェーズの国のデータセットに入力された。国レベルでは、ナショナル・センターが学生の回答を一部分析した。

オーストラリアの第1フェーズから得られた主な成果は、「挑戦」「関心」「応用」「範囲」「プログラムとの関連性」「将来の職業」の各基準において学生が試験をほぼ肯定的に評価した点であった。記述式問題は、多肢選択式問題との比較では、これらの基準に関して僅かに高く評価された。

AHELOテスト問題に対するオーストラリアの学生の解答の概要が以下の図1及び図2に示されている。これらの図は、学生にとって記述式問題が非常に難しかったこと、そして調査対象の他の特定領域についても全体的に肯定的であったことを示している。特に、プログラムとの関連性、将来の職業との関連性に関心が高かった (図1)。難易度及びこれらの学生の学位課程との関連性は選択式の問題でも評価が高かったが、図2は、選択式問題の関心、応用及び将来との関連性が参加学生によってより低く評価されたことを示している。

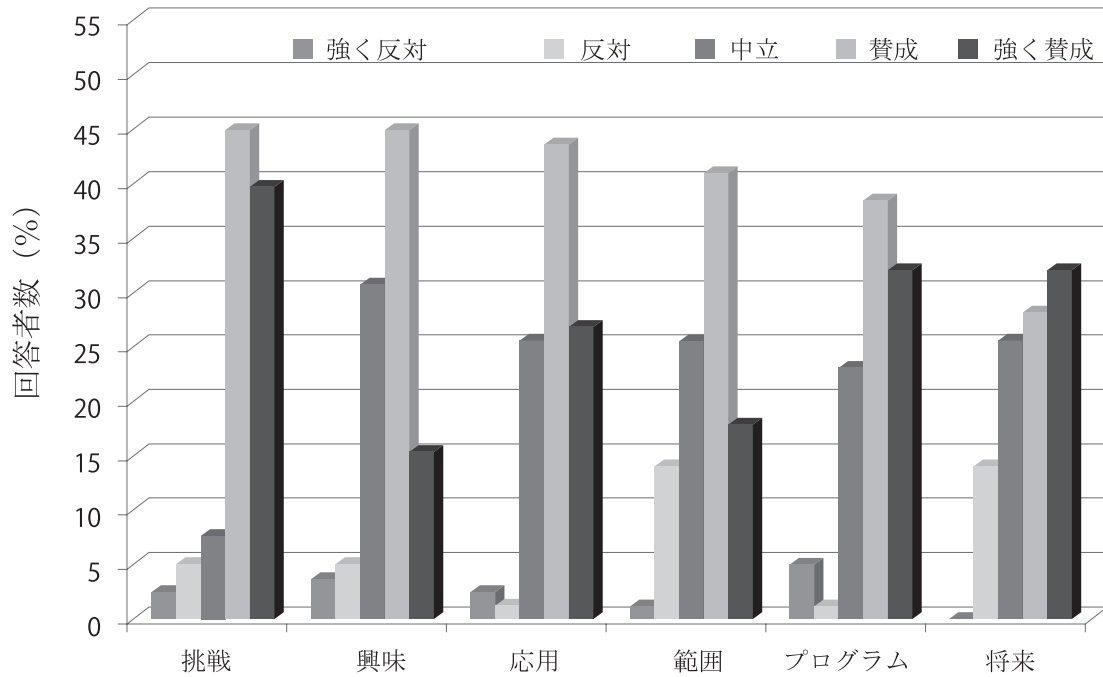


図1 記述式問題に関するフィードバック

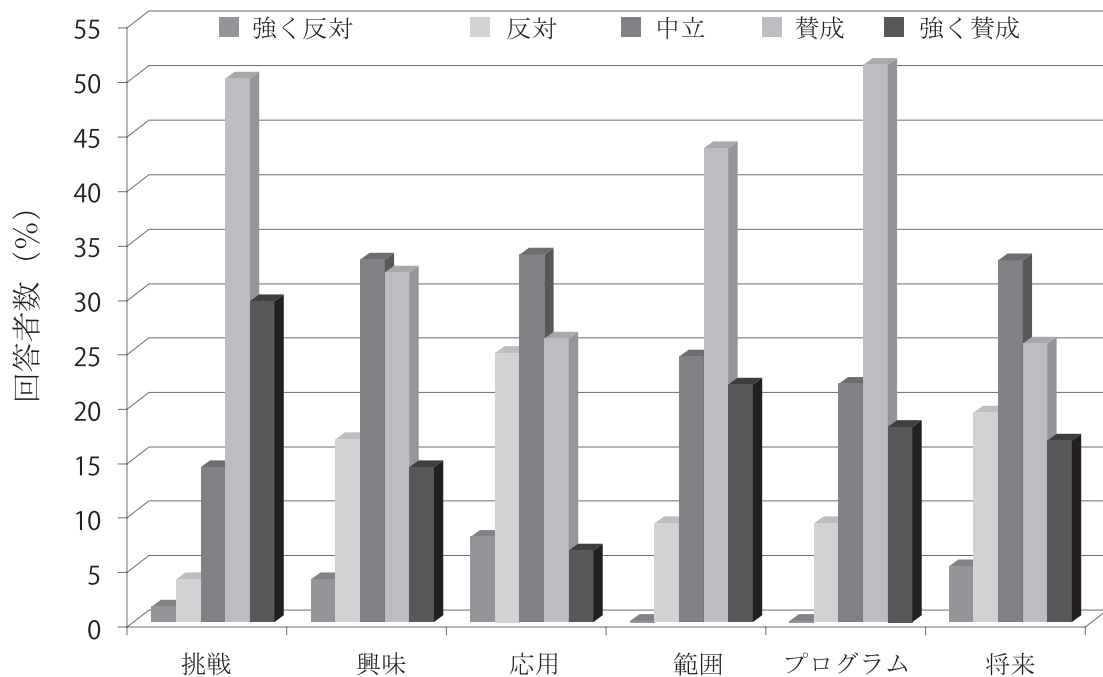


図2 選択式課題のフィードバック

アンケートによる試験の評価に加えて、学生のフォーカス・グループにもモデレーターが進行するディスカッション調査が実施された。要約すると、このディスカッションでの主な回答は以下のとおりであった。

- 記述式問題について尋ねられると、ほとんどの学生が、内容は理解でき、指示は適切だったと答えた。圧倒的に多くの学生が課題は面白かったと答えており、その理由は図や現実の問題を扱っていたことと、試験が通常のテスト内容とは違っていたためであった。
- 選択式の課題についての学生の一般的な意見は、これは彼らが大学で良く見ている課題により近いというものであった。この課題はより記憶力指向であり、学位との関連性はあるが、長期的には将来のキャリアとの関連性は低いと思われると答えている。

テスト問題草案の検討は、参加したオーストラリアの大学所属の土木工学の3名の研究者によって行われた。学術的検討の成果は以下のとおりに要約できる。

- 全般に、フィードバックは学生のものと同様であった。
- 内容が学士課程の最終学年よりもむしろ初期のものに集中していたという指摘が数件見受けられた。
- 技術的な説明や使用文言に関連する更に具体的な項目ベースのフィードバックもあり、若干の項目について質問内容がより明確になるように変更することが示唆された。

オーストラリアとその他の参加国における学生フォーカス・グループ及び学術的検討の結果に基づき、AHELOコンソーシアムによって、土木工学分野で使用されたテストには、数多くの変更が加えられた。本研究における現時点での主な成果は、当初の計画より長い時間が試験に配分されたという点であった。記述式問題の一つは難易度が高すぎたため試験から完全に除外された。他の記述式問題は改正された。

量的検討：第2フェーズ

AHELOの第2フェーズは、2012年にオーストラリアの8大学で実施された。フィージビリティ・スタディに参加した学生は、記述式問題1問と多肢選択式問題のモジュールを含む90分の試験で構成されたAHELOの土木工学アセスメントを受けた。学生は試験セッション終了後、学生の人口統計及びその他背景の特徴に関してデータを収集する背景アンケートにも記入した。

サンプル抽出及び学生募集

試験運営準備のための主な作業は、AHELO実施ルールに従い、各参加大学からサンプリング枠を収集することであった。AHELOデータの国際的な比較可能性を確保し、収集されたデータの処理を促進するため、全ての参加教育機関は全ての適格な学生及び職員のリストを所定の表書式で提供することを要求された。この書式によって、国際的に対象母集団を検証し、必要な場合にはサンプルの選択をすることが可能となった。オーストラリアでは土木工学部の規模がサンプリングというよりもむしろ学部内の全数調査が可能にほど小さかった。試験の母集団に入るための適性は、土木工学プログラムに所属する学士号取得前の最終学年の学生と定義された。

NPMに対するサンプリング枠の提供は、教育機関コーディネーター側に多大な尽力を要する取組だった。なぜならこれは、学生の就学率状況及び人口統計上の様々な特徴に関する最新情報を必要とするからである。教育機関はサンプリング枠にデータを強化すると思われる背景指標を追加するよう要請され、教育機関によっては学生が属するプログラム、学生の出生国、又は成績評価点の平均等の学生の資質・適性の指標といった情報を提供した。その他の教育機関はデータ共有の規則による拘束がより厳しく、性別等の基本的事項の提供しかできなかった。

試験の運営は2012年の4月と5月に全ての教育機関で実施された。学生のアセスメント及びアンケートの実施のため、何度もテストのセッションを設定した教育機関が多かったが、その他の小規模の教育機関ではセッションを一度計画しただけであった。一部の教育機関では、教育機関コーディネーターは全ての試験運営を監督したが、NPM予算からの財政支援により試験運営のアシスタントを雇ったところもあった。試験期間中、AHELO試験システム及びその他のAHELO実施ルールに関する問合せに関してNPMによる支援が可能であった。

オーストラリアのAHELO実施に対する主な課題は、学生に本研究に参加する動機付けを与えることであった。これは、2011年10月の国際的なNPMミーティング中にも他の一部の国によって強調された問題でもあった。オーストラリアで参加している教育機関の多様性を考慮して、NPMは参加奨励のために弾力的アプローチを取った。教育機関コーディネーターに地域に合ったインセンティブについてアイデアを提案してもらい、NPMがそれに賛同して資金を供給した。表2は教育機関によって参加学生に提供されたインセンティブの種類の詳細である。様々な選択肢及びアプローチが提供されたことが示されている。一部の教育機関では各参加者にバウチャーを提供できるほど全学生の母集団の数が小さかったが、参加者向けに賞品の抽選をしなければならない大学もあった。金銭的インセンティブのほかに、将来の職業経験に対する試験の関連性及び本研究の成功にとっての参加の価値を説明することにより、学生達は参加に意欲を持った。ある大学は学生向けに金銭又は「賞品」のインセンティブを提供しないことを選択し、代わりに唯一アセスメントの経験だけを参加者にとってのインセンティブとした。

全ての教育機関は、試験を受けた学生に与える参加証明書を提供された。これらの証明書はAHELOコンソーシア

表2：学生の参加インセンティブ

教育機関	インセンティブ
A 大学	参加者一人当たり 100 オーストラリアドルのパウチャー
B 大学	Ipad の抽選。参加者のための結果評価ランチディスカッション
C 大学	参加学生に 50 オーストラリアドルのパウチャー
D 大学	Ipad の抽選。会場別でパウチャー抽選 (8 会場) , 会場ごとに 200 オーストラリアドル又は 100 オーストラリアドルのパウチャーが当たるチャンスあり。
E 大学	Ipad2 個の抽選。5 名に 50 オーストラリアドルのパウチャー抽選, 昼食の提供
F 大学	1000 オーストラリアドルの賞金の抽選。加えて全体参加率の高い学生ソサエティに 1000 オーストラリアドル (75% : 未達成)
G 大学	なし
H 大学	中核科目の必修テストとした。

表3：学生の参加人数及び参加率，オーストラリアのAHLEO参加校

教育機関	参加人数	参加率* (%)
A 大学	3	30.0
B 大学	27	12.3
C 大学	4	12.5
D 大学	29	19.3
E 大学	27	29.0
F 大学	25	14.2
G 大学	19	12.8
H 大学	53	98.1
合計	187	21.1

*参加率は、本研究の対象教育機関によって構築されたサンプル抽出枠で特定された全学生数に対する参加学生数の比率に基づく。

ム及びOECDによって作成され、OECDのロゴが付いていた。

表3は、各大学の学生及び職員の参加人数並びに（サンプリング枠の選択によって特定された適格参加者の数に基づく）参加率を示している。オーストラリアの教育機関からは全部で187名の学生がAHELOアセスメントに参加した。関係した教育機関全体では、最終学年の土木工学部の学生の5名に1名がアセスメント調査に参加している。ただし、この平均値は、教育機関の間の学生参加率の明らかな差異を適切に表していない。八つの教育機関のうちの五つで参加率は25%を下回ったが、その一方で一つの教育機関（H大学）では学生がほぼ完全参加した。

数多くの問題がオーストラリアにおける学生の参加率に影響を及ぼした。その中でも、最終学年のインターンシップにより、学生の課程の中の「学内」での時間が限られていること、相当数の学生がパートタイム労働に就いており、したがってほとんど自由な時間がないこと、そして試験期間の通知期間が短かったことが挙げられる。

しかし、この表が示すようにH大学が高い参加率を確保できたことは、この種の事業でオーストラリアの高等教育における学生群の参加拡大の達成が可能であることを示している。同校が学生からこれだけの参加を得ることに成功したことを受け、以下の欄で同校のアプローチの情報を更に詳しく説明している。

成功した参加モデル

H大学が取ったアプローチは、理想的な実施モデルに対する洞察を与えるものである。このアプローチには教育機関内における強固な及び分散したリーダーシップ、洞察力のある計画、そして学生と職員間の学習成果についての対話をアセスメントに統合したことが関係していた。

同教育機関におけるAHELOの実施は前もって計画されており、この計画はナショナル・プロジェクト・マネージャーによって推進された説明会に続き、試験の4か月前にあたる前年の11月に開始された。プロジェクトのリーダーシップは、大学副総長代理のオフィスの職員が直接関与する大学総長事務局、工学部の長、そして実施に最終的な責任を持つ上級教職員という、同教育機関内の三つのレベルに存在した。試験の計画で、教育機関内のチームは土木工学プログラムの最終学年の中核科目（プロジェクト/論文科目）のうち学習成果及び卒業能力の題材が適切なものを特定した。この科目はAHELOが学期中の学業にとって不可欠な部分となるよう計画された。学生は研究科目の一部としてアセスメントに取り組むよう依頼され、アセスメントの後にはディスカッションに参加し、自分たちのアセスメント経験並びに、授業内容、卒業後の職業での応用が予想されるスキル、教育及び実践水準の保証に関連した職業上の責任の間の関係について考察を行った。

参加者の特徴

AHELOの学生参加者の特徴が表4に示されている。学生の人口統計で顕著なのは、工学部では男子学生が多数派であり、主要な言語が英語以外である学生の比率が高い点である。参加学生の平均年齢は23.2歳であった。大部分の学生はフルタイムかつ通学（遠隔教育ではない）の学生であった。サンプル枠と確保された参加者集団との基本的な比較は、性別ではサンプルが母集団を表していることを示唆しているが、参加者の年齢は完全母集団の平均よりも僅かに高いと思われる。

表4：学生参加者の特徴

特徴	項目	参加者人口 (全体に占める 割合 (%))
性別	男性	80.1
	女性	19.9
年齢	20歳以下	5.3
	21歳	20.5
	22歳	30.4
	23歳	19.3
	24歳	9.4
	25-30歳	11
	31歳以上	4
就学タイプ	主にパートタイム	4.1
	主にフルタイム	95.9
就学方法	通学	91.2
	通信と通学の混合	8.8
出生国	オーストラリア	53.8
	中国	14.6
	インド	2.9
	マレーシア	5.8
	その他	22.8

テスト結果

オーストラリアによるAHELO参加から得られたデータ結果に関する本稿での議論は、制約されたものである。なぜなら、このデータが比較的取扱いに注意が必要なものであることと、大多数のオーストラリアの関係大学の学生のサンプル数と代表性において、信頼性が高いデータの収集に必要とされるレベルを下回ったという事実によるものである。その代わりに、オーストラリアの完全なデータセットを使用して比較的幅広いレベルの所見をいくつかここで示す。学生の成績とAHELOの結果の関係、そして職業の「実践」とAHELOの成果の関連性という、二つの問題が議論される。以下の得点を検証する際のおおまかな参考値として、AHELOの国際平均は500、標準偏差は100である。

図3は関係を学生のAHELOアセスメントの結果と大学での研究中の彼らの成績（学生による自己申告）の間の関係性を示している。グラフではオーストラリアの学生のうち自分がクラスの上位にいると特定した者が他の学生よりAHELOの平均点を大きく上回っていることがはっきりと示されている。これは、AHELOテスト問題に若干の併存的妥当性が存在することを示す助けとなる。

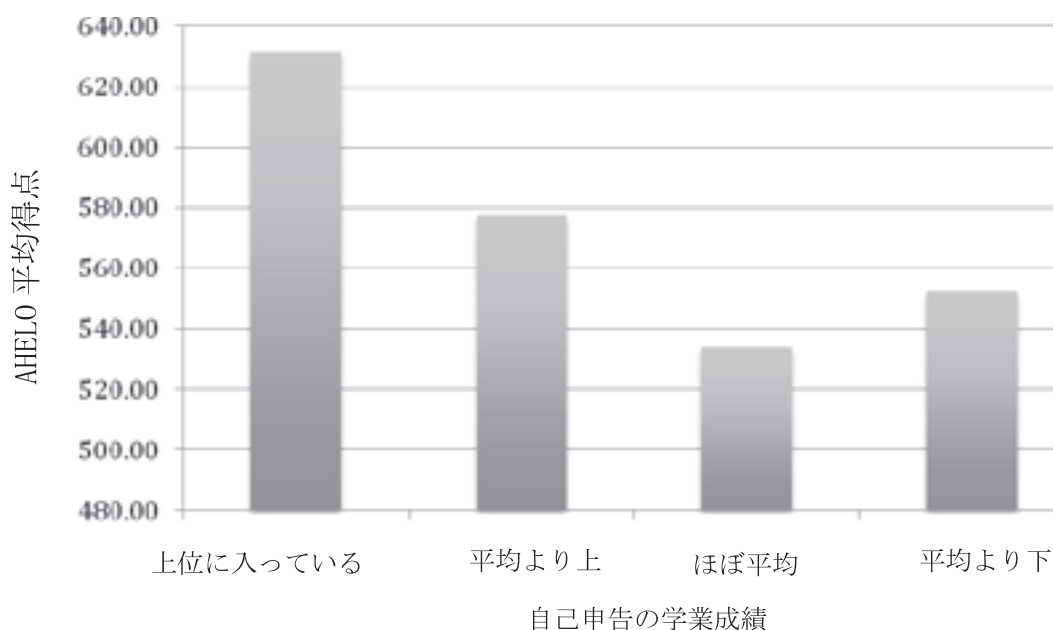


図3 オーストラリア人参加者：自己申告された大学での成績別のAHELOの平均得点

オーストラリアの高等教育における重要な問題は、研究と労働の関連性であり、どの程度研究が若者に将来のキャリアのための準備をさせているかである。上記に示される第1フェーズからの成果は、少なくとも試験の内容に関しては、学生が記述式問題に将来の職業との相当な関連性を見だしていたことを示していた。興味深いことに、オーストラリアのAHELOのデータは、その学位に関連した有給の仕事に就いている大学の最終学年の学生及び卒業後に研究に関連した業種で働くことを計画している学生は、現在専門分野で働いておらず、将来もその計画がない学生よりも得点が高いことを示している。

図4には、1週間に最低6時間、31時間未満、自身の研究に関連した有給の仕事に就いている学生は、関連分野で有給の仕事に就いていない学生、又はそうした仕事に30時間超就いている学生よりも平均してAHELOアセスメントの結果が良かったことが示されている。これは、専門分野における労働から得られたどの程度のスキルが、学生がAHELOアセスメントの質問にうまく答えることができる度合いへの影響を助長するのかについて考える上で、興味深い発見である。加えて、労働の負担が非常に重い最終学年の学生（30時間超の有給労働）にとって、こうした有給労働が大学でのパフォーマンス能力に関して負の影響を与えている可能性もあることが示されている。

その他のオーストラリアのAHELOの参加者の若干の興味深いデータを示す労働と研究に関連した問題は、学生の卒業後に対する意欲に関連したものである。図5のデータは、学位取得後すぐに土木工学のキャリアを追求しようと計画している学生は、AHELOアセスメントでより高い得点を得ている可能性が明らかに高かった。こうした学生は学士課程のこの時期には最もモチベーションの高い学生であり、恐らく、専門分野で仕事をする計画のない学生

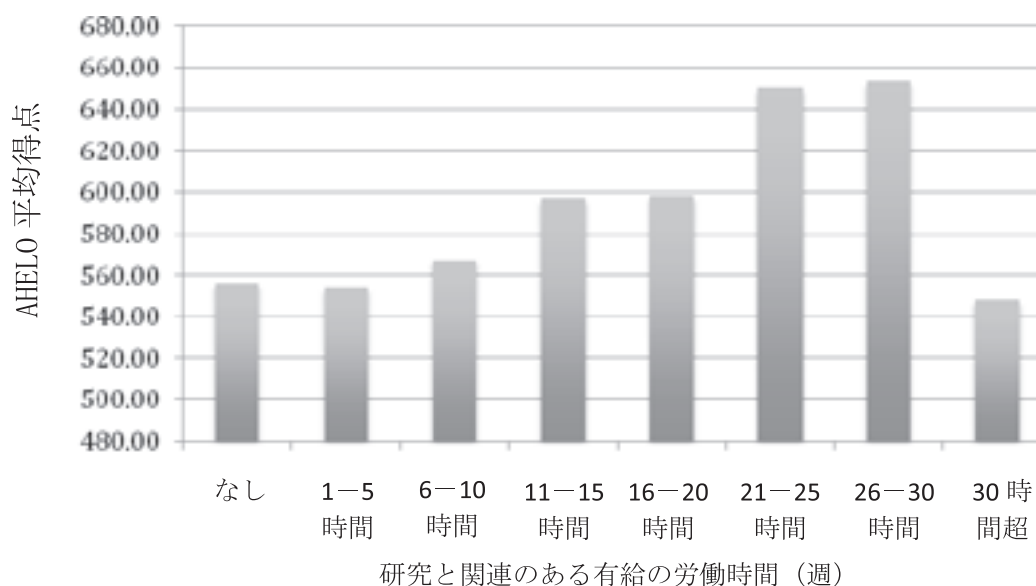


図4 オーストラリア人参加者：研究に関連した仕事における1週間の有給労働時間別のAHELOの平均得点

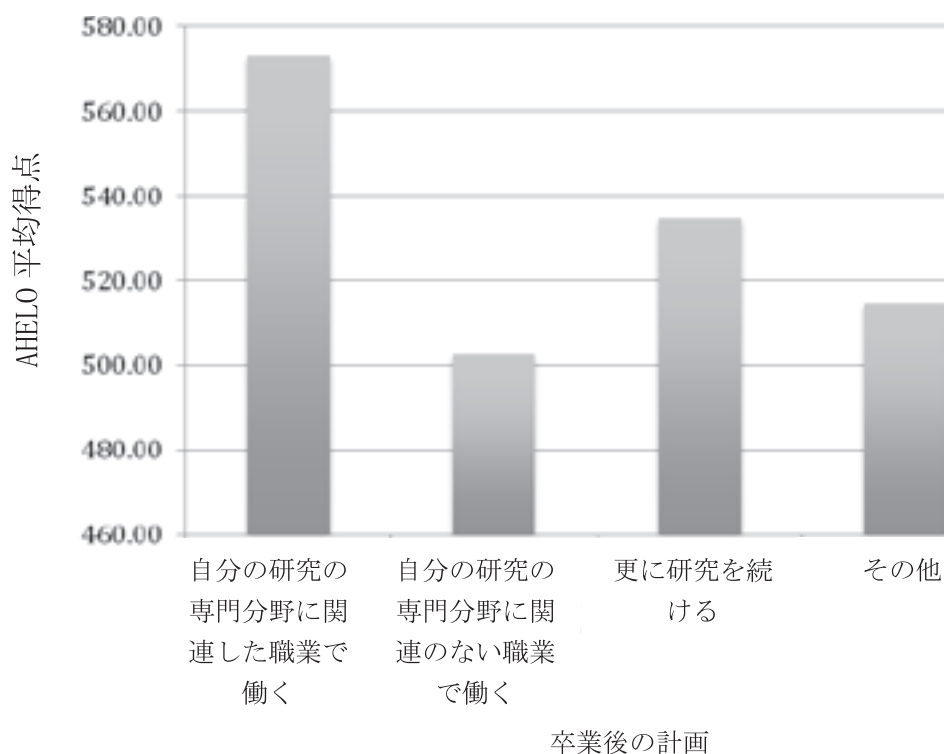


図5 オーストラリア人参加者：学生が計画する卒業後の進路別のAHELOの平均得点

よりも集中力が高まっている可能性が高い。興味深いことに、専門分野で更に研究を続けることを意図した学生の平均得点は専門分野で働く計画のない学生よりも高かったが、土木工学分野の労働に参加しようとしている学生より低かった。

結論及び将来に向けたアイディア

AHELOフィージビリティ・スタディへのオーストラリアの参加によって結果的に多くの価値ある教訓が生まれた。第1に、オーストラリアがこの種の国際研究に参加する能力があることを示した。利害関係者も参加に関心を持っており、教育機関内の運営システムにより比較的直接的に学生及び職員の必要なデータの生成が可能となった。教育機

関コーディネーターのネットワークを監督したナショナル・センターのプロセスはよく機能し、そして、コーディネーターはこの種の協力に価値を見出したと報告をしている。

第2に、国際的に適用可能な工学分野の試験の開発と試行に参加することは、参加教育機関及び学生の双方にとって、新たな理解と発見があった。AHELOは、学生たちが自分たちの将来の職業を反映した応用・統合タイプの問題にはそれまで出会ったことがなかったと報告しているという点で、真に草分け的な試験であると判明した。会議及びオンラインコミュニケーションによる複数の国からの工学の専門家間の協力は、国際的なきずなを強める優れた機会を提供した。加えて、参加教育機関同士の協力は、工学部間の将来の協力に向けた機会をもたらした。

現実的な面をみると、AHELOはオーストラリアの学生が自主的な試験又はアンケートへの参加を容易には希望しないことを示した。H大学によって用いられたような実施モデルを使用し、将来におけるAHELO又は類似した研究を繰り返していけば、この種のアセスメントを最大限カリキュラムに統合するのに役立つであろう。オーストラリア向けフィジビリティ・スタディの参加率が期待外れではあったが、その実施プロセスは、将来の研究における学生及び教育機関の間のこの種類の関与のために必要なプロセス及びシステムに関して相当な知識を築き上げた。

AHELOは、選ばれたオーストラリアの大学における工学学士号教育の現状に対して、いくつかの小さな洞察を提供した。他国の学生と比較しても、このデータからはオーストラリアの学生のスキルの水準についていかなる強固な結論も出しえない（そのため本稿では成果の分析量を削減した）が、将来再開される際に何が可能であるかを垣間見ることはできる。重要なのは、AHELOフィジビリティ・スタディへの関与がオーストラリアに価値ある教訓及び当該アセスメントの将来の実施に向けたモデルを提供した。

オーストラリアにとっては、フィジビリティ・スタディによって明らかになった、将来の国際参加に向けて価値ある反省材料がいくつか存在する。その一つは、このフェーズで使用された試験期間では、オーストラリアの学生が技術的に他国（日本を除いて）の教育機関から1学期分遅れるという点である。試験期間はオーストラリアの学生の実績が最終学年の1学期に当たったが、他の大部分の参加国では最終学年に行われた。正確な国際比較のためには、本研究が将来再開される際には全ての国で学年度の比較可能な時期に行われるべきである。2番目は、オーストラリアの参加に関与した多数の教育機関が最終学年に非常に大規模な最終学年インターンシップ又は研究プロジェクトを行っている点である。これは学生がこの年にキャンパス外で重要な時間を費やすことを意味する。このように、教育機関にとって信頼できるアセスメントに大きな集団が参加できる時期を見つけるのは困難である。こうしたアセスメントの運用をより長期的に計画すれば、最終学年のこれらの重要行事の影響の最小化に役立つ可能性がある。しかし、インターンシップ及び研究プロジェクトには、柔軟性が重要であるため、この問題は達成するのが難しいであろう。将来考慮するよう推奨される最後の問題は、参加した個人に向けた学生レベルの成績報告書の作成である。これはフィジビリティ・スタディの範囲を超えると考えられるが、将来こうした機能を構築することがこの種の研究への学生の関与を刺激するのに役立つとオーストラリアは考える。

フィジビリティ・スタディの終了以来、オーストラリアのNPMは他国のNPMとともに本研究の制約（特にアセスメント項目がまだ公開されず機密であるという事実）の中で教育機関向けの報告の開発に役立つ方法の特定に取り組んだ。この点に関してはいくつかの革新的な取組が行われてきており、教育機関レベル及び国家レベルの概要報告の見本が本稿の最後に示されている。これらの報告は個々の項目レベルの成果、難易度別指示項目を示しており、各項目によって試験されている能力タイプに関する参照を提供している。この見本は、将来このような報告を作成することを考慮した有益なものである。学生及び教育機関レベルにより明確かつ詳細な報告を行っていれば、フィジビリティ・スタディの実施後には更に大きな機運が得られていたであろうというのが本稿の著者の意見である。

謝辞

ACERコンソーシアムのメンバーであり、オーストラリア教育研究所の同僚である、Eva van der Brugge, Sarah Richardson, Hamish Coates, Jacob Pearce, Yan Bibby and Xiaoxun Sunに対して、本稿の内容への助言、及びフィジビリティ・スタディを通じての助力に感謝の意を表します。

改訂教育機関報告用書式例
：学生の成績パターン

土木工学学習成果
項目記号の説明

工学分野の一般的技能	
A	工学関係者や一般社会と効果的にコミュニケーションを図るために、多様な方法を駆使する能力
B	工学の学際性に関する理解
基礎科学及びエンジニアリング科学	
C	科学的・数学的原理の理解・知識の表明
D	工学の各分野の要点・概念の全体的理解の表明
E	素材・建設の包括的知識の表明
F	構造工学の包括的知識の表明
G	環境地盤工学 の包括的知識の表明
H	水力工学の包括的知識の表明
I	都市・農村計画の包括的知識の表明
工学解析	
J	既存の方法を用いて工学課題を見極め、解決法を考案、解決する能力
K	工学の成果、過程、方法を分析したりするために、知識と理解を応用する能力
L	適切な分析方法やモデルを選択・適用する能力
M	文献を検索し、データベース等の多様な資料を活用する能力
N	適切な実験の計画及び実施、データ解釈、結論付け
O	ワークショップ及び研究室におけるスキルの表明 ワークショップと実験を行う能力
工学設計	
P	特定の定義された要求に応える設計を開発して実行するために、知識と理解を応用する能力
Q	設計の方法を理解し、活用する能力
工学的実践	
R	適切な装置・道具・方法を選択・使用する能力
S	工学課題を解決するために、理論と実践を統合する能力
T	適用できる技法・方法とその限界を理解する能力
U	工学実践の非技術的な意味合いに関する理解。技術者倫理・工学実践の責任と規範に従う能力
V	健康・安全・法律の問題、工学実践が伴う責任、工学による解決策がグローバル・経済的・社会的・環境的文脈に及ぼすインパクトに関する理解
W	リスク・変動マネジメントを初めとするプロジェクト・マネジメントやビジネス慣行に関する理解。その制約についての認識

2012年 AHELO フィーダリティ・スタディ
オーストラリアの大学用見本
土木工学の学習成果
記述式問題

課題*	項目	項目説明 記号	平均素点	
			AHELO 母体のサバール ナースクリプトシステム	卒業見込み 課題
課題1 ダム	ENGCRIM17	B, U, V	0.9	0.6
	ENGCRIM16	U, V, W	0.7	0.9
	ENGCRIM14	V	0.7	0.9
	ENGCRIM11	D, J	0.7	1.1
	ENGCRIM15	K, L	0.7	0.9
	ENGCRIM12	Q	0.5	0.5
課題2 防樹堤	ENGCRIM28	M, W	1.2	2.1
	ENGCRIM23	P, R	0.6	0.7
	ENGCRIM24	S	0.6	1.3
	ENGCRIM25	O, Q	0.4	0.7
	ENGCRIM27	M, T	0.4	0.5
	ENGCRIM21	K	0.4	0.5
課題3 コンクリート 橋	ENGCRIM26	D, T, V	0.3	0.4
	ENGCRIM22	J, Q	0.2	0.3
	ENGCRIM36	A, B, U	1.3	2.2
	ENGCRIM34	T	0.7	0.8
	ENGCRIM33	K	0.7	0.7
	ENGCRIM32	D, T	0.6	0.4
ENGCRIM35	L, O	0.5	0.6	
ENGCRIM31	N	0.1	0.1	

取得単位項目における部分点
 実施なし

* 生徒一人につき1課題。

全国概要報告用書式例（改訂版）

2012年 AHELO フィージビリティ・スタディ
 オーストラリアの教育機関
 土木工学の学習成果
 記述式問題

取得単位項目における部分点
 実施なし

課題	項目	項目 説明 記号	平均素点									
			AHELO 全体のサンプル	オーストラリア全体のサンプル	教育機関 3	教育機関 4	教育機関 9	教育機関 6	教育機関 2	教育機関 5	教育機関 10	教育機関 1
課題 1 ダム	ENGCRMT17	B, U, V	0.9	0.9	1.0	0.9	1.4	0.8	1.2	0.9	0.6	
	ENGCRMT16	U, V, W	0.7	1.0	1.0	1.0	1.0	0.9	1.0	1.0	0.9	
	ENGCRMT14	V	0.7	0.9	1.0	1.0	1.2	0.7	1.0	0.9	0.9	
	ENGCRMT11	D, J	0.7	0.9	1.0	0.8	0.8	0.8	0.7	1.0	1.1	
	ENGCRMT15	K, L	0.7	0.9	1.0	1.0	1.0	0.9	1.0	0.9	0.9	
	ENGCRMT12	Q	0.5	0.5	1.0	0.6	1.2	0.2	0.5	0.6	0.5	
課題 2 防潮堤	ENGCRMT28	M, W	1.2	2.1	3.0	2.7	1.5	2.7	2.0	1.8	2.1	1.0
	ENGCRMT23	P, R	0.6	0.7	0.0	0.6	0.9	1.1	0.8	0.4	0.6	1.0
	ENGCRMT24	S	0.6	1.3	2.0	1.7	1.4	1.1	1.6	1.5	1.1	1.0
	ENGCRMT25	O, Q	0.4	0.7	1.0	0.7	0.9	1.0	0.4	0.8	0.5	0.3
	ENGCRMT27	M, T	0.4	0.5	1.0	0.3	0.9	0.8	0.0	0.5	0.4	0.3
	ENGCRMT21	K	0.4	0.5	1.0	0.4	0.4	0.4	0.4	0.6	0.4	0.3
	ENGCRMT26	D, T, V	0.3	0.4	1.0	0.7	0.3	0.3	0.4	0.3	0.4	0.7
	ENGCRMT22	J, Q	0.2	0.3	1.0	0.3	0.3	0.6	0.4	0.4	0.2	0.7
課題 3 コンクリート・橋	ENGCRMT36	A, B, U	1.3	2.2	2.5	2.5	2.1	2.0	1.8	1.7	2.4	
	ENGCRMT34	T	0.7	0.8	0.5	0.9	0.9	0.7	0.8	0.9	0.6	
	ENGCRMT33	K	0.7	0.7	1.0	0.7	0.7	1.0	0.8	0.7	0.6	
	ENGCRMT32	D, T	0.6	0.4	0.0	0.8	0.3	0.3	0.4	0.1	0.4	
	ENGCRMT35	L, O	0.5	0.6	1.0	0.6	0.6	0.8	0.7	0.6	0.4	
	ENGCRMT31	N	0.1	0.1	0.0	0.0	0.0	0.2	0.4	0.1	0.1	

2012年 AHELO フィージビリティ・スタディ
 オーストラリアの教育機関
 土木工学の学習成果
 多肢選択式問題

□ 項目別正答率

実施なし

組	項目	項目説明記号	正答率(%)									
			AHELO 全体のサンプル	オーストラリア全体のサンプル	教育機関 5	教育機関 4	教育機関 6	教育機関 9	教育機関 3	教育機関 2	教育機関 10	教育機関 1
第1組	ENGMCQ3	G	54	81	95	82	95	94	100	43	72	100
	ENGMCQ1	C	51	47	73	41	40	35	0	57	49	0
	ENGMCQ2	E	35	36	18	45	60	59	50	36	16	100
	ENGMCQ5	H	20	18	23	23	20	12	50	14	14	0
第2組	ENGMCQ7	C	61	57	65	65	38	50	50	65	57	50
	ENGMCQ8	F	54	50	39	52	62	53	75	71	36	50
	ENGMCQ9	F	54	64	61	78	62	69	50	71	59	0
	ENGMCQ10	G	50	51	78	61	52	63	100	41	27	0
	ENGMCQ6	C	47	55	65	70	67	63	25	41	41	50
第3組	ENGMCQ13	C	55	66	81	67	64	77	100	73	53	0
	ENGMCQ14	E	33	30	42	26	36	46	33	33	19	0
	ENGMCQ11	H	28	34	62	37	55	31	33	20	14	0
	ENGMCQ12	H	26	24	35	26	27	15	33	27	16	0
	ENGMCQ15	E	22	25	42	30	23	23	0	20	16	33
第4組	ENGMCQ16	F	59	57	38	76	62	76	67	63	40	100
	ENGMCQ20	I	45	57	67	52	71	65	67	50	47	67
	ENGMCQ18	F	40	41	52	28	38	53	33	25	44	67
	ENGMCQ17	G	25	31	43	60	38	18	0	6	20	67
	ENGMCQ19	H	18	14	19	16	14	6	0	13	16	0
第5組	ENGMCQ22	F	48	49	63	75	57	53	33	53	26	0
	ENGMCQ23	G	38	42	54	33	52	33	33	27	39	100
	ENGMCQ25	I	32	51	46	58	43	47	0	67	52	67
	ENGMCQ21	C	27	30	42	42	33	33	33	33	13	33
	ENGMCQ24	H	24	32	58	29	33	40	100	13	20	0
第6組	ENGMCQ29	F	63	70	95	71	50	82	100	77	61	50
	ENGMCQ27	F	31	31	32	42	25	35	33	31	27	0
	ENGMCQ30	H	16	22	32	25	35	18	67	23	9	0
	ENGMCQ28	I	6	6	5	0	0	6	0	0	14	0

AHELO: オンタリオ州高等教育機関の経験

Mary Catharine Lennon

オンタリオ州高等教育質保証カウンシル, 上級研究アナリスト

はじめに

2011年にオンタリオ州は、経済協力開発機構（OECD: Organisation for Economic Co-operation and Development）による「高等教育の学習成果調査（AHELO: Assessment of Higher Education Learning Outcomes）」のフィージビリティ・スタディに加わった。オンタリオ州高等教育質保証カウンシル（HEQCO: Higher Education Quality Council of Ontario）は、職業訓練・カレッジ及び大学省（MTCU: Ministry of Training, Colleges and Universities）に代わり、カナダ教育閣僚会議（CMEC: Council of Ministers of Education）と協同でプロジェクトを主導した。

2006年に始動したAHELOは、大学生の知識と能力を測定する手段として、標準化された一般的試験並びに専門分野固有の試験を様々な国で利用できるかどうかについて判断するためのフィージビリティ・スタディである。学習の質を表す優れた指標の確立に向けた国際対話に資することを意図した本調査は、共通の学習成果を開発し、様々な教育文化・言語・教育機関において、標準化された試験を通じて学士課程（第1サイクル）終了時に学生のパフォーマンスを測定することを目的とした。このフィージビリティ・スタディでは三つの分野でテスト問題開発がなされた。一つは一般的技能、その他二つは経済学及び土木工学という専門分野固有の技能に関するものであった。

この世界規模のプロジェクトには17か国が参加しており、カナダは工学分野に参加した九つの国・地域の一つである。土木工学プログラムを持つオンタリオ州内10大学のうち9校が本調査に参加し、土木工学部で卒業を迎える最終学年の全学生のうち、その約61%にあたる450名近くの学生が参加した。

本報告書では、フィージビリティ・スタディに参加したオンタリオ州の経験を振り返る。ここでは主に、本調査の実施・運営活動並びに教育機関にとってのその価値に焦点を当てる。参加教育機関は自校プログラム作りに関して具体的な洞察を得ることはなかったが、AHELOは国際的アセスメントと比較理解の観点から相当な関心呼び、大規模アセスメントの運営において大きな経験を提供するものとなった。

AHELOフィージビリティ・スタディの背景と実施根拠

20世紀後半には、高等教育というものに世界規模で大きな変化が起こった¹。高等教育のマス化は、教育機関とそのプログラム及び学生の多様化を招き、小規模なエリート集団育成というそれまでのシステムとは異なるものであった（OECD, 2012a）。こうしたマス化の背景には多くの要因がある。高等教育はもはや従来の教育施設や国・地域内に留まるものではない。技術的に強化された学習の拡大により、学生が望む時間と場所でプログラムを利用できるようになり、各国の学生と教員にもたらされた可動性により教育機関が世界に向かって開放された。さらに、学生からの教育に対する需要に後押しされて、私立の教育機関や専門プログラムを備えた教育機関といった新たな代替的教育プロバイダーが出現した。

中等後教育を対象とした現在の複雑かつグローバルな「市場」は、新たな形態によるガバナンス・説明責任・信号伝達の仕組みを必要としている。従来の説明責任及びガバナンスとともに、インプット（資金調達、図書館所蔵等）・アウトプット（在籍率、卒業率、出版物）の評価結果が国内外でのパフォーマンス指標となった。これらのパフォーマンス指標に基づく国際的ランキングは、教育機関の研究能力を評価する際に重要な役割を果たす。

一方、こうした従来の方法では捉えられないものが教育の質であり、多くの教育機関の中核となる教授と学習である。こうしたギャップが認識されたことにより「代理変数（proxies）」と呼ばれる新たな指標が導入された。教育の質を示すこの「代理変数」には、学生の評価・学生の満足度と授業参加に関する調査・労働市場における成果等が含まれる。これらの指標から、学生が中等後教育以降も満足いく順調な生活を送っていた場合、その教育機関による教育

¹本セクションの情報はOECDの文書に基づいている。詳細はTremblay, Lalancette and Roseveare (2012) 及びOECD (2013a) を参照されたい。

の質が高かったことが伺える。

ここ最近では、教育の質を把握・立証・評価する手段として学習成果に注目が集まってきており、これは一部ではパラダイム・シフトとも考えられている（OECD, 2012a, 35）。学習者の知識と能力を明確化した学習成果は、適切な評価方法を伴えば、学生の学習を測定する透明性の高い手段となる。このため、教育機関の教授と学習の質、最終的には提供されている教育の質について客観的に把握することが可能となる。

教育機関の教授と学習の質を明確化することは多くの意味で有益である。学生・雇用者・一般市民に対する教育の価値への理解を深め、教育機関とそのプログラムレベルでの連携・カリキュラム開発・教育実習の改善を実現する。さらには、これまで品質保証と説明責任の分野において明確に把握できないものとされてきた教育の質について評価が可能となる。加えて、透明性のあるプログラム作りを実現し、これにより教育機関とプログラムに対する国際的理解並びに比較理解を促進する。このように学習成果の明確化によって、教育というものが教育機関・プログラム・多くの一般市民と国際社会の利益となるよう、より身近な分かりやすい存在となるが、究極的には学生をその中心に据えたものである。このことは学生中心の学習へと大きくシフトすることを意味する。

本フィージビリティ・スタディの目的は、「概念実証（proof of concept）」（OECD, 2009a, 15）であり、これによって学生の知識と能力が国際水準にあるかどうかを判断するとともに、高等教育機関・政府・学生や雇用者を始めとするその他利害関係者に対して関連情報を提供するものである（OECD, 2009b, 2）。

こうした取組から生まれた主な疑問事項は以下のようなものである。

- 予想される学習成果について国際的な合意を得ることは可能か。
- 異なる文化・言語間で同じ試験を実施することは可能か。
- そうしたアセスメントは有効かつ信頼性が高いものか。

本フィージビリティ・スタディには三つの主要作業領域が存在し、学生の成績評価のため三つのアセスメント（テスト問題開発）分野が設定された。その一つは、批判的思考力・分析的思考力・問題解決力等の一般的技能を評価するためのものである²。その他二つの分野は、経済学³及び工学⁴における専門分野固有の技能に特化したものである。二つの専門分野のテスト問題では、知識内容の評価よりはむしろ、知識の応用力（つまり学生が「技術者のような思考ができるか」）に注目した。追加的分野として設けた「付加価値」は、学生の学習に寄与する高等教育について調査することを目的とし、中等後教育開始から終了までの間に、学生のスキルと能力に向上が見られることが予想される（OECD（2013b）を参照）。

これら二つの作業領域に加えて、背景情報の側面（Contextual Dimension）⁵の設定を行った。これは、各国・地域における教育環境への理解を高め、観察した学生間に生じた成績差の原因を特定するため、学生・教員・高等教育機関並びに国・地域を調査することを目的としている。

アセスメント枠組みの開発と同時に、OECDはその加盟国⁵並びに興味を示したその他の国に対し、フィージビリティ・スタディへの参加とテスト問題の共同開発、さらに自国の教育機関をサンプルとした本調査の試験的实施を要請した。その結果、6大陸から12言語にわたる17の国・地域が本調査への参加に同意した。オンタリオ州は土木工学分野に参加した。

オンタリオ州におけるAHELO

2011年7月、オンタリオ州の職業訓練・カレッジ及び大学省は、カナダの他州から独立して本調査に加わることを決定した⁶。オンタリオ州政府からは独立した機関であるHEQCO⁷がプロジェクトを主導するよう要請された。

²詳細は、Tremblay et al. (2012) ; p 112-114を参照のこと。

³詳細はFor more information, see OECD 2011a; Tremblay et al. (2012), p 116-118.を参照のこと。

⁴詳細は、OECD 2011b; Tremblay et al. (2012), p 121-123, OECD, 2012a, OECD 2012bを参照のこと。

⁵参加国の一覧については<http://www.oecd.org/general/listofocdmemberscountries-ratificationoftheconventionontheoecd.htm>を参照のこと。

⁶OECDでの審議において高等教育分野を代表するカナダ教育関係協議会（CMEC）は、全州の合意の下、様々なOECDプロジェクトに携わっている。しかし、一般から関心が高いにも関わらず、CMECは、本フィージビリティ・スタディへの参加については各州からのコンセンサスを得られなかった。

⁷詳細については www.heqco.ca を参照のこと。

中等後教育の価値を把握・立証するためには広範な研究課題が存在するが、その一環として、オンタリオ州はHEQCOを通じて、学習成果を調査する多くの研究プロジェクトに携わった。この分野でHEQCOは多数の活動に関与してきており、各種専門分野にわたる学習成果の確立（チューニング・プロジェクト, Lennon, Frank, Humphreys, Lenton, Madsen, Omri & Turner（近刊予定A）を参照）、大学学習アセスメントの試験的实施による全般的学習成果の測定（Lennon（近刊予定B）を参照）、並びに教育機関の様々な活動を通じた学習成果の収集と立証支援を行った。こうした国際的調査への参加により、システムとプログラム作りに関する比較情報が得られ、本調査の基盤が形成された。

ナショナル・センターとして、HEQCOはOECD並びにAHELOコンソーシアムと連携してアセスメント調査の実施と運営にあたった。さらにOECD-AHELO間の対話において国側の専門家としての役割も担い、ナショナル・プロジェクト・マネジャー（NPM）1名を派遣して、参加教育機関とともに本調査を実施・運営し、その後の分析を行った。オンタリオ州で本調査を準備・運営するに当たり、多くの活動が実施された（図1を参照）。コンソーシアムによって操作可能となったフィージビリティ・スタディでは、全ての国・地域において実施する活動が確実に同等なものとなるよう、調査の実施ガイダンスとプロトコルが設けられた。次のセクションでは、主なフィールドワークと実施活動領域を振り返り、OECDプロトコルとオンタリオ州の対応双方について述べる。

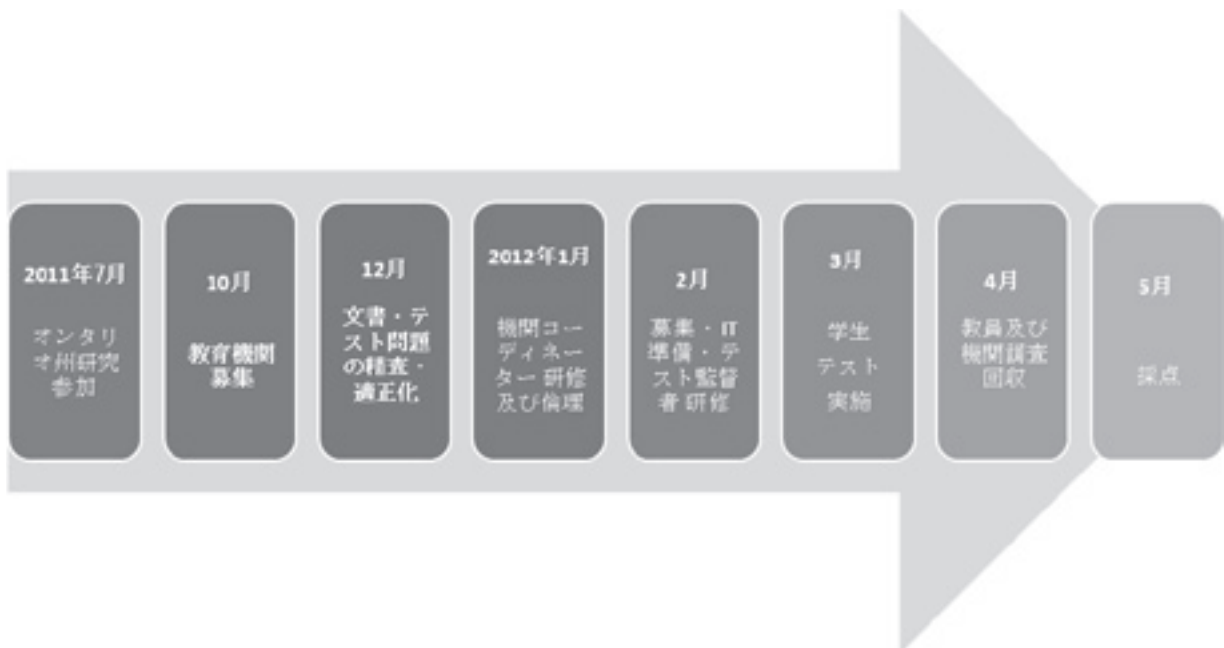


図1. オンタリオ州における実施活動の基本スケジュール

参加教育機関の募集と採用

OECDは各国・地域に対し、学士号レベル（又は第1サイクル）の土木工学学位を授与する教育機関で様々なタイプを代表する10校を便宜的サンプルとして採用するよう推奨し、その中には、規模とマンドレート（教授又は研究といった任務領域）が異なり、地方と都市に所在する公立及び私立の大学・工科大学校・カレッジを含めることが望ましいとされた。各システムについて有効かつ信頼できるデータを収集するため、あらゆるタイプの教育機関と学生を適切に代表するサンプリングの実施を目標とした。

オンタリオ州には学士号レベルの土木工学プログラムを提供する教育機関が10校ある。これら教育機関の副学長並びに工学部学部長に対し、書面によりプロジェクトへの参加を呼びかけた。参加を促進するため、各教育機関には少額ながらプロジェクト運営費用が支給された。参加を呼びかけた教育機関からは直ちに極めて良好な回答が得られ、10校中9校が参加に同意し、自校のみならず海外にある教育機関の学生とプログラムを理解する手段としてこうした国際的アセスメントへ参加することに興味を示した。

オンタリオ州の教育機関は、自校の工学プログラムを評価するため学習成果を測定しプログラム戦略を立てるという概念を既に持っていた。ワシントン協定の合意事項の一部として、またプログラム認定を目的として、工学プログ

ラムを提供するカナダ国内の教育機関が2016年を目途に学習成果に基づく学位プログラム作り・アセスメント・認定取得に向けて取り組んでいる（EGADS, 2013年）。こうしたことを背景に、今回の大規模なアセスメント調査にはツールとして相当の関心が集まった。オンタリオ州の全土木工学生の約90%が、これらのプログラムに参加した。

参加教育機関は全て公立校であり、文系から理系いたる幅広いプログラムを博士号レベルまで提供している。主に都市部に位置し、その規模はフルタイム換算（FTE）した学生数にして14,595名から75,941名に上る。オンタリオ州から参加した9大学は以下のとおりである。

- カールトン大学
- マクマスター大学
- オタワ大学
- クイーンズ大学
- ラリアソン大学
- トロント大学
- ワーテルロー大学
- ウェスタン大学
- ウィンザー大学

土木工学プログラムは工学部に所属するが、環境工学と連携したのももあった。プログラムの教員数は17名から40名であった。フルタイム換算（FTE）した学生数は231名から573名であった。1年間に授与する学士号の数は、2010年度で20名から101名に上った。

文書の適正化・精査・検証

テスト問題開発において専門知識を提供する採点リーダーを各国で1名ずつ選任した。この採点リーダーは、自身の専門分野において専門家と認められた個人であり、その責務は、テスト問題に関する全ての文書を確認し、国・地域にとって適切なものとなるようテスト問題と翻訳を適正化し、さらに学生と教員を対象とした試行試験を実施して検証することであった。各国の採点リーダーは、試験項目を適切に採点する作業を共同で行い、国内の採点チームの訓練を担当した。

オンタリオ州の採点リーダーはナショナル・プロジェクト・マネジャーと協力の下、学生に対する試行試験の実施⁸や教員と上級生である大学院生に試験のフィードバックを求めるなど、様々な活動を実施し、オンタリオ州の背景状況と照らし合わせて試験内容を綿密に調査した。収集したフィードバックは、試験問題内容と難易度の適切性並びに（言語ニュアンスや技術的専門用語などの）オンタリオ州の背景状況に配慮した妥当性の確保についてコメントするものであった。

こうしたフィードバック情報はOECDに報告され、テスト問題に盛り込まれた。オンタリオ州では大きな変更が求められることはなかった。実際のところ、国際採点チーム内でも採点リーダー間の摩擦はほとんどなく、各国・地域の採点リーダー間で早急に解決・合意に達した。

高等教育機関の活動

今回の調査に参加した各教育機関は、教育機関コーディネーター（IC: Institutional Coordinator）を選任した。機関コーディネーターの責務は、自身の教育機関内の調査を管理しナショナル・センターであるHEQCOと連携することであった。また、AHELO調査において必要不可欠な役割を担い、自身の教育機関内においてAHELOを実施する責任を負った。

現地活動の標準化は、活動の結果得られる情報の信頼性を確保するために重要であった。アセスメント設計がいかにも優れていても、その運営と実施の問題がデータの有効性と信頼性に影響を及ぼす可能性がある。そのため、フィージビリティ・スタディにおける重要事項の一つとして、現地活動の共通化確保が求められた。

教育機関におけるAHELOの実施

各教育機関は、AHELO実施国・地域間での整合性と必要な支援を確保するため、ナショナル・センターと緊密に連絡を取り合った。オンタリオ州では、各教育機関内で機関コーディネーター（多くは土木工学部の学部長）が選任された。AHELOに関する状況・背景並びに活動ロードマップが提供される一日トレーニング会がHEQCOで開催され、

⁸ 5名の3年生が試験を受けフィードバックを提出した。4年生の母集団が損なわれないよう3年生が選ばれた。

9名の機関コーディネーターが出席した。この会議の後、機関コーディネーターとナショナル・プロジェクト・マネージャーは、毎週電話会議を行って活動に関する支援と助言を提供し合った。

フィールドワークの実施には数多くの作業を要し、各機関コーディネーターは作業補助チームを結成した。その推奨メンバーにはITの専門家と試験管理者が含まれた。オンタリオ州にある教育機関の多くはこうした提言に従ったため、チームの規模が他よりも大きくなることもあった。

機関コーディネーターが最初に行う作業の一つは、試験実施に向けて教育機関による倫理的承認を申請することであった。その他の国・地域とは異なり、オンタリオ州の教育機関では、学生といったヒトを対象とした調査の実施には許可を必要とする。実施スケジュールが短期間であるため迅速に試験を遂行する必要があることから、オンタリオ州の教育機関は、倫理的承認取得のため、調査パラメータを修正する必要があった。

例えば、(サンプルが確実に母集団を代表し、GPAや高校の平均値等に基づくテスト結果の理解を高めるため) AHELOは機関コーディネーターに対し、学生の管理情報にアクセスすることを要請したが、オンタリオ州の教育機関では、AHELOの試験結果を学生ファイルとリンクさせる許可を取得できなかった¹⁰。その結果、試験を受けた学生サンプルに関する情報の信頼性と、教育機関のデータ検証能力の低下を招いた。例えば、教育機関に対し学生の管理情報へのアクセスが許されていれば、AHELO試験の個人別得点をGPA等の他の学生能力指標とともに検証することができたであろう。教育機関でこのような性質の情報が確保できていれば、オンタリオ州でのAHELO試験の得点について、その有効性に対する理解の向上につながった可能性もある⁹。

フィールドワークの実施

学生募集

学生募集はAHELOフィージビリティ・スタディの重要な要素であった。各教育機関は、試験を受ける学生サンプルとして200名を選出し、募集率75%を目指すよう求められた。指定専門領域(工学・経済学、一般的技能の場合は学校全体)に200名以上の学生が在籍する教育機関の機関コーディネーターは、どの学生を潜在的参加者として選出するかを定めたサンプリング枠を設ける必要があった¹⁰。こうした合目的サンプリング枠は、母集団を代表する学生サンプルが確実に試験を受けるようにするためのものである。潜在的受験者が200名未満の教育機関では、学内調査として全学生参加を目標とした。オンタリオ州の全土木工学教育プログラムでは、最終学年の学生数¹¹が200名未満だったため、全ての適格な学生¹²が各教育機関でのアセスメントの潜在的候補者となった。

各国・地域では、それぞれ独立した学生募集戦略を展開した。学生参加を義務化した地域もあるが、任意参加制とした地域もあった。任意参加制の地域全てではないが、その大多数で、学生にインセンティブが与えられた(OECD, 2013a, 169ページ)。オンタリオ州の学生参加は完全に任意であった。機関コーディネーターは地域の倫理プロトコルが定める範囲内で適切と判断した学生を採用した。参加勧誘のため、ポスター・説明会・ギフト券・賞品抽選会・土木工学部学生会への寄附といった様々な活動やインセンティブによる勧誘が行われた。あらゆるタイプの分析を行うに足る十分な学生数を確保することが必須であるとの認識を持っていた機関コーディネーターにとって、学生募集が最も時間がかかり懸念を要する活動であった。学生参加が任意だったため、オンタリオ州の教育機関では全て、非無作為抽出による任意参加となった。

教員に対する短期背景情報調査も同様に実施した。40名以上の教員を有する学部(一般的技能の場合は教育機関)に対しては、代表的サンプルを確実に採用するためのサンプリング枠を設けた。オンタリオ州の全参加教育機関と同様、教員が40名未満の場合には、全学生参加による学内調査が試みられた。全ての教員に対し参加を義務化する又は勧誘活動を行うことができないため、オンタリオ州の教育機関では、志願者による任意参加となった。

⁹学生の管理データにアクセスできたのは機関コーディネーターのみであった。ナショナル・センター、コンソーシアム、OECDはアクセスできなかった。

¹⁰本フィージビリティ・スタディでは、倫理的承認の取得が困難であったが、今後こうしたアセスメントは、倫理的承認申請と研究目標達成の時間が確保できれば、大規模な修正なしに受け入れられる可能性が高い。

¹¹このサンプリング枠組はコンソーシアムの指導のもと実施された。

¹²最終学年の学生数は34名から137名であった。

¹³国外にいる学生又は特別措置を必要とする障害を持った学生は除外された。

試験の運営

AHELO試験システムは、2012年2月から6月まで稼働し、教育機関はこの期間中であればいつでも自校の学生をテストすることができた。各教育機関は、技術的問題点がないことを確認するため、コンピュータプラットフォームのテストを実施するよう求められた。試験の時間と日程はあらかじめナショナル・プロジェクト・マネジャーに報告された（技術的又は緊急の問題が生じた場合）。ナショナル・プロジェクト・マネジャーはこうした情報をさらにコンソーシアムに報告し、試験のオンラインシステムが一連の活動に対応できるよう準備態勢を整えた。機関コーディネーターはAHELO実施期間を設定し、試験運営者（試験監督官）に（AHELO訓練ガイドラインに基づいた）適切な訓練を受けさせる責任を負った。

オンタリオ州では試験の実施時期が課題であった。大学の学期カレンダーでは、中間テスト後の2月中旬に1週間の休みがあり、期末プロジェクトと期末テスト後の4月に学年末を迎える。こうしたことから、試験を実施できる期間が非常に限られていたため、多くの教育機関では3月上旬から中旬にかけて試験を実施した。教育機関は学生の授業スケジュールを考慮し、様々な時間帯と日程で試験を行った。

システムのテスト稼働を実施したにもかかわらず、オンタリオ州の教育機関のうち1校が技術的問題に直面し、学生が試験を提出できなかった。コンソーシアムの技術的問題は是正されたが、当該教育機関では学生の回答のほぼ8%が損失した。

採点

学生テスト問題の選択式問題では、コンピュータによる自動採点方法が採用されたが、各国・地域の教育機関では、自校の学生による記述式問題の回答を人力で採点した。採点リーダーは他の採点リーダーとともに2回の研修会に出席し、テスト問題の質問を確定の上、適切な採点ルーブリックを決定した。これにより各国・地域の採点方法に整合性が確保された。採点リーダーはナショナル・プロジェクト・マネジャーと協力し、採点官による小チームを編成して試験システムと採点ルーブリックについて訓練を行った。

オンタリオ州では技術者6名による採点チームを編成した。採点官の多くは、オンタリオ州の学生のアセスメントと取組の検証に関心を持つ機関コーディネーターであった。採点は2012年6月に2日間にわたってナショナル・センターで行われた。オーストラリアとの了解覚書（MOU）¹³の一部に基づき、オーストラリアが管轄する試験の回答をオンタリオ州が一部採点し、その逆の採点も行われた。こうした採点方法は、採点を行う国・地域の採点官同士の信頼関係を評価する目的で行われたが、このようなことがなければ海外の学生の回答を見る機会がなく、採点チームにとって関心の対象でもあった。

試験の成功と課題

AHELOフィージビリティ・スタディの主たる目的は、世界中の学生を対象とした共通テストを共通の方法によりオンラインで実施することが運営上可能であるかを把握することであったが、実際にその実行可能性が証明された。専門家と教員が共通の学習成果並びにテスト問題での質問について合意し、プロジェクト管理と各種作業が世界共通プロトコルに従って実施された。運営面での軽微な問題が各国で起こり、一部の教育機関による離脱・回答率の低さ・技術的問題の発生もあったが、今回の国際的アセスメントがおおむね成功裏に実施されたことは明らかである。AHELO報告書の第1巻には、本調査の運営面で各国が学んだ多くの教訓が掲載されている（OECD, 2012a, 6章）。運営面の成功と課題に関しては、三つの分野で見られたのと同様、参加国・地域間で大きな相違があった。例えば、ある国の教育機関では、テストを実施する学生数に対して十分なコンピュータ台数が確保されていないことが分かり、当該教育機関にコンピュータを届けるための移動用のバンが必要となった。信じられないことに、エジプトでは、同国で「アラブの春」が繰り返されている最中での実施が予定されていたため、それが大きな困難となったが、なんとか乗り越えることができた。一般的技能分野では学生募集が問題となった。

こうした実施面の問題は、本調査の主たる目標である試験運用の実用性を把握するために極めて重要である。そのため、各国の経験を収集すべく多くの取組がなされてきた（OECD（2013a）を参照のこと。各国の経験については

¹³詳細は以下MOUのセクションを参照のこと。

8章)。

オンタリオ州内での成功と課題は、大規模に実施する様々なテストに見られるものと、今回のAHELO特有のものがあつた。受験者にとって重要性が低いテストでは、学生募集が困難を極めるのが通例である。広報活動(ポスター、電子メール、授業参観等)やテストセッション実施の双方において時間がかかり、学生に金銭その他の大きなインセンティブが与えられる場合には、多大な費用が発生する可能性がある。今回の学生募集は困難を極めたが、教育機関の募集戦略は極めて創造的であり、最も時間のかかる活動であつたにもかかわらず、自校の学生の参加を見ることは機関コーディネーターにとって非常に喜ばしいことであつた。

フィージビリティ・スタディというものを実施する際に予想されるとおり、オンタリオ州にも小さな問題が数件発生した。倫理的承認を迅速に取得する必要があつたため、アセスメント枠組みを修正してAHELO結果と個人レベルの学生データを関連づける教育機関の能力を縮小した。このため、教育機関における調査結果の潜在的価値が低下した。本報告書執筆時点では、どの教育機関もまだ学内データを分析していなかったが、国・地域別の比較報告を入手した際にそうしたデータを検証する意欲を示していた。

試験後の面接と追跡調査において、参加学生と教員は国際的プロジェクトに参加できたことに感謝の意を表し、調査に参加するという自身の決断を評価していた。このことから、ベンチマーキングやプログラムを比較されることに対して恐怖感を抱いているのではなく、各国・地域の教育機関との比較を通じて自校のプログラムの特性・長所・短所を理解することに参加者全員が素直に関心を抱いていたことが伺える。

議論及び結論

オンタリオ州の結果分析

九つの国・地域から70校を超える教育機関がフィージビリティ・スタディの工学分野に参加した。オンタリオ州では、現在土木工学プログラムを提供している10の教育機関のうち9校がAHELOの工学分野に参加した。環境情報が教育機関から収集された。これら9校の教育機関からは、教員計155名、最終学年の学生計443名がAHELOの工学分野に参加した。オンタリオ州の回答率は教員72%、学生61%¹⁴であつた。AHELOに参加した教員の属性データ、職業上の身分・資格は、オンタリオ州の教育機関全体のそれと類似していたが、教授と研究に費やしていると報告された時間に顕著な相違があつた。AHELOに参加した土木工学プログラムの学生に注目すると、ここでも学生がどのくらいの時間を授業の準備と実際の授業に費やしているか、自身の研究分野との関連の有無にかかわらずどのくらいの時間を有給の仕事に費やしているかといった学業への取組の度合いに相違が見られた。オンタリオ州の参加教育機関に関する詳細(学校・教員・学生の特徴)、オーストラリアとその他九つの参加国・地域との比較結果については、Lennon(近刊予定C)を参照されたい。

AHELOの貢献

ここで思い出していただきたいのは、本フィージビリティ・スタディの目的が比較ランキングデータを収集することではなく、この種の取組の実施可能性をフィールドワークとテスト問題の試行双方において検証することである。このため、得られたデータの徹底分析は、次のいずれかの理由により、実施していない。a) 枠組みにおいて同意されていないため。b) 後日得られたデータの制約を受けるため。

収集したデータは細心の注意を払って解釈する必要があるとの認識から、データ結果とそれが利害関係者(すなわち国・地域、教育機関、教員及び学生)にもたらす価値についてコメントすることは不可能である。これは、現在そうした関係者がその情報を使用することができないためである。

このように、今回のフィージビリティ・スタディから、収集した情報をどのようなことに役立てることができるかを明確にするためには、テスト問題の改良を要することが明らかとなつた。さらに、情報を用途に合わせて適切に手直しするために各種の枠組みやテスト問題を作り直せば、その情報が誰にとって有用であるのかという問題への答え

¹⁴第9教育機関ではオンラインシステムの技術的問題の結果、36名のアセスメント参加学生のうち33名の回答が記録されなかった。オンタリオ州の回答率は、アセスメント参加学生数に基づいて算出されており、実際に記録された回答数に基づくものではない。

が曖昧となってしまふことも判明した。

ここでフィージビリティ・スタディの結果について、その価値を2方向から考察する。

1. 様々な利害関係者集団がAHELOへの参加から期待したものは何か。
2. こうした集団のニーズに応えるため、アセスメントにはどのような調整が可能か。

国・地域

AHELOを開始した当時、その目的が国内外の教育機関のランキングではないことが明確にされた。しかし、国・地域間で比較・対照できるということは大きな魅力であった。また各国政府にとって、自国の教育機関とそのプログラムがどのように組織・構成されているかを把握し、それらが学生の学習成果へ及ぼす影響について検証することができれば、価値ある情報を得る機会となる。こうしたことが、政策の比較理解向上と、システムと教育機関の間における相互作用の改善につながる可能性がある。

上述のようなデータ比較は、様々な理由から、今回のAHELOフィージビリティ・スタディでは実現できなかったが、システムレベルの学習においてはその可能性が存在する。例えば、単純なデータ・ポイントに基づいてランキングできないような形で情報が提示されても、相手国に比較情報又は国際平均が提供された場合には、国際比較が可能となる。

教育機関・プログラムレベルのデータ

AHELOフィージビリティ・スタディの主たる目的は、他校と比較して自校の学生が自身の様々な能力をどのように発揮するのか、それについて教育機関に情報を提供することであった。さらに、背景情報調査で収集された情報により、学生の教育環境が持つ特性について認識を深めることも目的としていた。

AHELOで実施したテスト問題には、コンピテンス領域¹⁵ごとに学生の能力に関する教育機関レベルの情報を収集するに足る高い精度はなかった。このため、学生の総得点以外のものを基準に教育機関を比較することが不可能であった。これはフィージビリティ・スタディに参加した教育機関の期待を裏切るものであったが、ツールの改良により、教育機関における学生のコンピテンス領域を比較する試験の開発が可能となろう。また、こうした情報は、教育機関の長所・短所を把握する際に非常に役立つといえる。例えば、設計と実践において非常に優れた教育機関が、基本的工学スキルでは劣っているということが明らかになることもある。こうした詳しい情報が、プログラムと教員にとってカリキュラム設計の際に有用な情報となる可能性がある¹⁶。国内外にある他校のプログラムをどのような形で自校と比較できるかを把握することが、オンタリオ州の教育機関が今回の調査への参加を決定した主たる理由であった。こうした情報の入手はできないが、そのために試験を再構築することは可能である。

教育機関及びプログラムレベルで収集する背景データも有用情報を含んでいる場合がある。オンタリオ州の教育機関でも教員・学生に関する管理情報を収集しているため、こうした情報のほとんどは他のソースでも容易に利用可能である。したがって、自校の学習環境に関する情報も教育機関やプログラムにとっては特別に新しい情報ではなく、国・地域内の比較情報にも同様のことがいえる。

しかし教育機関とそのプログラムを他校や国際平均と比較して見れば興味深い点が明らかとなる。世界中の教育機関とそのプログラムを比較することにより、学習環境の共通点・相違点を解明し、それらが学生の能力に及ぼす影響について把握することが可能となろう。さらに、プログラム設計と環境特性に見られる国際的傾向をプール解析すれば、成功プログラムの傾向に関する情報の収集、さらに共通の構造的特徴によるベンチマーキングに大いに役立つと考えられる。

このように、現在提供されている教育機関情報とその比較にも価値があるが、世界的傾向に関して情報をプールするだけでなく、その国際的な対比からもより大きな洞察が得られるかもしれない。

¹⁶工学設計、工学実践、工学分析、並びに一般的・基本的工学スキル。

¹⁷コンピテンスレベルの適正評価には、アセスメントの再設計が必要となるため、学生にとっては試験期間の長期化につながる事が指摘されている。

学生レベルのデータ

学生層の内訳・特性・テスト得点に見られる注目すべき傾向に関しては、AHELOで収集した学生レベルのデータから有用な情報が得られる可能性があった。

しかしフィージビリティ・スタディの枠組みの中で期待されていたのは、自校の学生の状況を同地域内にある他校のそれと比較した情報を教育機関に提供し、その情報を活用して当該教育機関が学生の学業の成功をサポートできるようにすることである。例えば、恐らくは学外で就業していることが原因で、年齢が高い学生ほど良い成績を修める可能性が低いことが判明すれば、その教育機関でそうした学生への支援策を改革するという選択ができる。さらに、こうした情報を他校と共有すれば、同様の問題に取り組んでいるプログラムとの連携促進へとつながる。現行のAHELOの枠組みでは実施されていないが、学生の成功例に見られる国内外の傾向を特定し、国・地域に関係なく教育機関の間で協力と連携を促進していくことは興味深い取組といえる。

今回の試験は、教育機関レベルで学生データとの連結を行うことを意図していたため、学生レベルのフィードバックを行うようには設計されていなかった。しかし、その後の議論の中で、学生レベルのフィードバックを行うことは、学生にとっては自身の能力を同輩との比較により明らかにする手段となるため、そうしたフィードバックに意義があり、またそれに対する関心も高いことが明らかとなった。学生レベルのフィードバックを提供するには、まず試験を大幅に変更して、学生全員が同等か又は同一の試験を確実に受けるようにする必要がある。

AHELOに参加した国・地域、コンソーシアム並びにOECDが時間を費やして議論した問題の一つは、学生に自身の得点についてフィードバックするか否かであった。今回のアセスメント枠組みは、教育機関レベルでフィードバックを行うことを意図しており、信頼性が高い得点情報を個人レベルで提供するようには設計されていない。しかし他校の学生（教育機関内あるいは国内・海外）との相対評価による得点は、自身の能力について世界規模で客観的に評価された記録でもあるため、学生からの関心が高まり、さらには、学生募集活動と試験を受ける学生の取組が改善される可能性がある（情報の信頼性の向上につながる）。こうしたことはAHELOフィージビリティ・スタディが本来意図したものではないが、多くの参加国・地域にとって非常に重要な問題である。以上のことは、今回のようなアセスメントの結果得られる数多くのアウトプットの一つといえるが、こうしたアウトプットを現実化するためには改めてアセスメントの枠組を設ける必要がある。

結論

このたびのAHELOフィージビリティ・スタディは、国際関係の構築・システムレベルでの比較理解の補助・教育機関とそのプログラムの理解支援・学生レベルでの国際的アセスメントにおける実施可能性の探究という点で成功を収めた事業である。参加国・地域だけでなくOECD加盟国レベルにおいても、今回の調査から多くの教訓を得ることができた。事実、世界中の学生を対象に標準化された試験を実施できることが確認された。また、工学プログラムを始めとする様々な分野の利害関係者がそうした試験の実施に関心を示しているようである。

オンタリオ州は、学習成果の確立と測定における幅広い有用性について国際的な対話に参加することができ、これは州にとって有意義な経験となった。同州は、学習成果を取り入れたシステムの開発に取り組んでいるため、州外のシステムがどのように学習成果を活用しているかを把握できることを評価している。中でも興味深かったのは、様々な国・地域が国際的な学習成果アセスメントの活用を望む中、システムレベルのベンチマークとして活用すべきか、教育機関とそのプログラムレベルの改善を目的として活用すべきかで意見の相違が見られることだった。

AHELOの取組はフィージビリティ・スタディとして多くの教訓をもたらした一方、課題も多く残した。こうした国際的アセスメントに対し、政府・教育機関・プログラム・教員並びに学生が関心を寄せていることが判明し、期待される学習成果と適切なアセスメントについては世界レベルで合意を得られる可能性があることも確認できた。さらに、世界中の学生を一貫した方法でテストすることが可能であることも判明した。以上のことが今回のAHELO調査で確認できた主要事項である。

References

- EGAD (2013) *Engineering Graduate Attribute Development Project*. Retrieved from: <http://egad.engineering.queensu.ca/>
- Lennon, M. C. (2010). *Signalling Abilities and Achievement: Measuring and Reporting on Skill and Competency Development*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming A). *Tuning: Identifying and Measuring Sector-Based Learning Outcomes in Postsecondary Education*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming B). *Piloting the Collegiate Learning Assessment in Ontario: Lessons Learned*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming C). *The AHELO Feasibility Study in Ontario*. Toronto: Higher Education Quality Council of Ontario.
- OECD (2009a). *AHELO Terms of Reference*. Paris, OECD
- OECD (2009b). *Roadmap for the OECD Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study*. Paris, OECD
- OECD (2011a). *Tuning-AHELO Conceptual Framework of Expected and Desired Learning Outcomes in Economics*. Paris, OECD Publishing.
- OECD (2011b). *A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering*. Paris, OECD Publishing
- OECD (2012a). *Engineering Assessment Framework*. Paris, OECD.
- OECD (2012b) *Final Engineering Assessment Development Report*. Paris, OECD.
- OECD (2012c) *Contextual Dimension Framework*. Paris, OECD.
- OECD (2013a) *AHELO: Feasibility Study Report: Volume 2: Data Analysis and National Experiences*. Paris, OECD.
- OECD (2013b) *AHELO: Feasibility Study Report: Volume 3: Further Insights*. Paris, OECD.
- Tremblay, K. Lalancette, D., and D. Roseveare (2012). *AHELO: Feasibility Study Report: Volume 1: Design and Implementation*, Paris, OECD
- Tuning Association (2009). *A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering*. Netherlands, Tuning Association.

登壇者プロフィール

ローベルト・ワーヘナール

ワーヘナール氏は歴史学者で、現在、フローニンゲン大学（オランダ）人文学部長・同大学院研究科長である。同時に、ビルバオ（スペイン）及びフローニンゲンに置かれているチューニングアカデミーの共同ディレクターでもある。また、高等教育チューニングジャーナル編集委員会のメンバーであり、エラスムス・ムンドゥス修士プログラムのコーディネーターも兼任している。

さらに、ワーヘナール氏は、欧州委員会の高等教育外部専門家であり、1988年以降の欧州単位互換・累積制度（ECTS）や欧州高等教育圏資格枠組み、欧州生涯学習資格枠組みといった欧州の高等教育を調和させるための主要なイニシアティブに携わってきた。そして、オランダの高等教育機関における「ボローニャプロセス」実施に向けた国内の専門家チームの議長も務めている。

ワーヘナール氏は、「世界の教育構造チューニング」という大規模な革新的プロジェクトを、デウスト大学（スペイン、ビルバオ）のフリア・ゴンサレス博士と共同で、構想、デザイン、コーディネートしてきた。このプロジェクトにおいて、2009年には、OECDからの依頼を受けて、経済学・工学の分野で期待される/望まれる学習成果について、Tuning-AHELOの概念的枠組みが作成された。

ピーター・ユーウェル

ピーター・ユーウェル氏は、大学の効果向上を目的として設立された研究開発所である全米高等教育経営システム研究所副所長である。1981年にスタッフに加わり、それ以降、主に教育機関の有効性や大学の成果の測定を専門とし、測定・収集した情報を計画・評価・予算編成に活用することに関して、調査を行ったり教育機関や州政府等に対して直接コンサルティングを行ったりしている。この分野において、数多くのプロジェクトを指揮した経験を有し、プロジェクトの中には、ケロッグ財団、ルミナ財団、フォード財団、私立高等教育振興コンソーシアム、ピュー慈善信託が資金援助を行ったイニシアティブがある。また、375以上の大学、24の州政府の高等教育制度、12の国に対して、学習成果、プログラム評価、エンrollment・マネジメント（入学者管理）、戦略的計画といったトピックについてコンサルティングを行った経験を有する。

ユーウェル氏は、学習成果に基づくアプローチによる学部教育の改善について、これまでに著書6冊と多数の記事を執筆している。また、各州の教育審議会、全米知事協会、州議会議員連盟、高等教育政策研究所、及び経済協力開発機構（OECD）等の外部機関からの委託を受けて論文も執筆している。講演を求められることも多く、1985年に、米国の高等教育における学習成果をテーマとする最初の全国規模の会議で基調講演を行って以来、国内外で広く講演を行っている。1996年には、ウェスタン・ガバナーズ大学（WGU）の学習成果に基づく教育課程の編成、1998年には、現在米国の1200以上の大学が活用する全米大学生学習状況調査（NSSE）の作成チームを指揮した。

ユーウェル氏は、ハヴァフォード大学の卒業生で、1976年にイェール大学より博士号（政治学）を取得している。また、シカゴ大学での教員経験もある。

メアリーキャサリン・レノン

メアリーキャサリン・レノン氏は、オンタリオ州政府の独立研究機関であるオンタリオ州高等教育質保証カウンシル（HEQCO）の上級研究アナリストで、現在、学習成果の確立・測定プロジェクトを指揮している。

これまでに、カナダ教育閣僚協議会や英連邦大学協会などを含む、個別大学、州及び州間、そして国際的な教育関係機関における高等教育政策の開発、助言及び研究に携わってきた。

レノン氏は、修士課程及び博士課程にて比較・国際高等教育の課題や政策を研究し、現在トロント大学の博士候補生である。

このような経歴から、国際比較に基づきつつ、システムデザイン、アカウントビリティ、質保証、ガバナンスといったシステムレベルの政策課題に主に取り組んでいる。

現在、HEQCOが行う一連の学生の学習成果プロジェクトを担当しており、Collegiate Learning Assessment（大学学習評価）の試験的实施や学問領域3分野におけるチューニング・プロジェクトの促進、またOECDによるAHELO フィージビリティ・スタディのカナダ国内プロジェクトマネージャーを務めるなどしている。

ダニエル・エドワーズ

ダニエル・エドワーズ氏は、オーストラリア教育研究所（ACER）における高等教育研究プログラムの主任研究員である。ACERの高等教育研究の更なる発展のため、2008年1月に任命された。

エドワーズ氏の研究は、教育問題の多岐にわたり、特に高等教育を重点分野としている。これまでに、高等教育の需要（学生及び雇用者間両方）、学生の学習達成度、学生の目標・希望と進路、大学入学選抜方針、教育の「選択」理論といったテーマに関する調査を行ってきた。また、社会階層や人口変動といったより幅広い社会問題についてリサーチを行った経験も有する。

エドワーズ氏は、政府、大学、関係団体等に対し、高等教育領域における政策や実践に関する助言や見識を提示する様々なプロジェクトを主導している。具体的には、オーストラリアのNational Research Student Survey (NRSS)、AHELO国内プロジェクトマネジメント、AHELO背景情報調査開発等である。また、ACERの「Joining the Dots（点をつなぐ）」と呼ばれるオーストラリアの高等教育に関する政策情報リソースの編集及び管理を担当している。

エドワーズ氏は、モナシュ大学人口・都市研究センターの非常勤リサーチフェローでもある。

木村 孟 (きむら つとむ)

文部科学省顧問

- 略歴** **大学・大学院** 東京大学工学部土木工学科卒業（1961）
 東京大学大学院数理系研究科土木工学専攻修士課程修了（1964）
 東京工業大学工学博士課程修了（1968）
- 職業** 日本鋪道株式会社（1961）
 東京工業大学理工学部助手（1965）
 東京工業大学工学部助教授（1968）
 英国ストラスカイド大学研究員（1973）
 英国ケンブリッジ大学研究員（1979）
 東京工業大学工学部教授（1982）
 東京工業大学工学部長（併）（1992－1993.10）
 東京工業大学学長（1993）
 ケンブリッジ大学チャーチルカレッジフェロー（1993）
 東京工業大学工学部教授（1997）
 東京工業大学名誉教授（1998）
 学位授与機構長（1998）
 大学評価・学位授与機構長（2000）
 独立行政法人大学評価・学位授与機構長（2004）
 独立行政法人大学評価・学位授与機構特任教授（兼）（2009）
 文部科学省顧問（2009）
 日本技術者教育認定機構会長（2009）
 独立行政法人大学評価・学位授与機構名誉教授、特任教授（2009）
 現在に至る
- 公的委員** 中央教育審議会副会長（2001－2007）
 経済産業省独立行政法人評価委員会委員長（2001－）
 国土交通省独立行政法人評価委員会委員長（2001－2010）
 国立大学法人評価委員会委員（2003－2004）
 東京都教育委員会委員長（2004－）
 日本学術会議会員（2005－）
 中央教育審議会臨時委員（2007－）

主要著書・論文

- 『土の応力伝播』（鹿島出版会、1978）
 『土質力学（土木工学大系第8巻）』（彰国社、1980）

その他

- 土木学会賞論文奨励賞（土木学会）（1966）
 手島記念賞著述賞（手島工業教育資金団）（1969）
 土質工学会賞論文賞（土質工学会）（1982）
 土木学会賞論文賞（土木学会）（1988）
 名誉大英勲章 C B E（英国）（2004）
 日本学士院賞（2012）
 瑞宝重光章（2013）

岸本 喜久雄 (きしもと きくお)

東京工業大学大学院理工学研究科教授, 工学系長, 工学部長

略歴 **大学・大学院** 1975年3月 東京工業大学工学部機械物理工学科卒業
 1977年3月 東京工業大学大学院理工学研究科機械物理工学専攻修士課程修了
 1982年1月 工学博士 (東京工業大学)

職業 1977年4月 東京工業大学助手
 (1987年3月～1988年1月ケンブリッジ大学客員研究員)
 1989年1月 同助教授
 1995年9月 同教授
 2012年4月 同副学長 (2012年9月30日まで)
 2012年10月 同大学院理工学研究科工学系長, 同工学部長

公的委員 日本学術会議機械工学委員会委員長, 日本技術者教育認定機構副会長など

主要著書・論文

【研究分野】 材料力学, 破壊力学, 計算力学, 実験力学

【著書】 著書・訳書11編

主な著書

- Kikuo Kishimoto, Case Studies of "Electronic Packaging Failure, Interfacial and Nanococale Failure, (分担), Elsevier, 2003.
- S. Aoki, K. Kishimoto, Application of BEM to Galvanic Corrosion and Cathodic Protection; Topics in Boundary Element Research Vol.7 (Electrical Engineering Applications), Springer-Verlag, 1990.

【論文】 原著論文255編, 総説・解説17編, 招待講演
 (APCFS2004, JSMAMS2010, SPAWDA2010など多数)

主な論文

- Farid Triawan, Kikuo Kishimoto, Tadaharu Adachi, Kazuaki Inaba, Toshio Nakamura, Tohru Hashimura, The Elastic Behavior of Aluminum Alloy Foam under Uniaxial Loading and Bending Conditions, Acta Materialia, Vol. 60, pp. 3084-3093, 2012.
- 岸本喜久雄, 界面き裂の支配力学パラメータと界面強度評価試験 (総説), 日本接着学会誌, 43-11, pp.31-39, 2007
- Hiroyuki Mae, Kikuo Kishimoto, Modeling and Simulation of Impact Failure Characteristic of Polypropylene by Elastoviscoplastic Constitutive Law, Journal of Solid Mechanics and Materials Engineering, 1-1, pp.35-46, 2007.

その他

1980年 日本機械学会論文賞
 1993年 防食腐食協会論文賞
 2000年 日本機械学会論文賞
 2006年 日本材料学会学術貢献賞
 2007年 日本機械学会材料力学部門業績賞 など

金子 元久 (かねこ もとひさ)

筑波大学・大学研究センター 教授, 東京大学 名誉教授

略歴 **大学・大学院** 東京大学教育学部 卒業 (1972)
同大学院修士課程 修了 (1974)
シカゴ大学 Ph. D. (1985)

職業 アジア経済研究所研究員
世界銀行コンサルタント
米国ニューヨーク州立大学アルバーニ校客員助教授
広島大学助教授 (大学教育研究センター)
東京大学助教授
東京大学教授 (大学院教育学研究科)
東京大学大学院教育学研究科長・教育学部長
国立大学財務・経営センター 教授、研究部長
筑波大学・大学研究センター 教授 (現職に至る)

公的委員 中央教育審議会委員
日本学術会議会員
日本高等教育学会長
中国北京大学、華東師範大学、復旦大学、中国科学技術大学 各顧問教授

主要著書・論文

『大学教育の再構築－学生を成長させる大学へー』(玉川大学出版部、2013)
『大学の教育力－何を教え、学ぶか』(筑摩書房、2007)
P. Altbach & T. Umakoshi eds. Past and Future of Asian Higher Education: Baltimore: Johns Hopkins University Press, 2004 (共著)
『グローバル化時代における工学系大学院教育』(日本学術会議工学研究連絡委員会報告書、2003)
『教育・経済・社会』(放送大学教材)(1996) など

深堀 聰子 (ふかほり さとこ)

国立教育政策研究所高等教育研究部・総括研究官

- 略歴** **大学・大学院** 京都大学教育学部卒業（1991）
 京都大学大学院教育学研究科教育学専攻博士前期課程修了（1993）
 京都大学大学院教育学研究科教育学専攻博士後期課程退学（1997）
 コロンビア大学大学院教育学科博士課程修了 Ph.D.（2000）
- 職業** 東京大学社会科学研究所・助手
 京都女子大学短期大学部・専任講師、准教授
 国立教育政策研究所高等教育研究部・総括研究官（現在に至る）
- 公的委員** 人事院国家公務員採用総合職試験(人間科学)試験専門委員
 東京学芸大学附属高等学校 S S H 運営指導委員
 日本教育社会学会研究委員会委員・紀要編集委員会委員
 Tuning Journal for Higher Education, Editorial Board.
 大学教育学会国際化委員会委員
 広島大学高等教育研究開発センター客員研究員

主要著書・論文

- 『学習成果アセスメントのインパクトに関する総合的研究（研究成果報告書）』（平成21－23年度国立教育政策研究所プロジェクト研究事業）国立教育政策研究所、2012年、1-200頁。
- 深堀聰子・竹中亨（共訳）（Gonzales, J. and Wagenaar, R.編著）『欧州教育制度のチューニング－ボローニャ・プロセスへの大学の貢献－』明石書店、2012年、1-198頁。
- 「学習成果の評価－工学分野の取り組みを例に考える－」『比治山高等教育研究』第4号、比治山大学高等教育研究所、2011年、89 - 101頁。
- 岸本喜久雄・深堀聰子「学部教育の現状と今後」『機械の研究』第62巻第1号、養賢堂、2010年、117～126頁。
- 『高等教育における学習成果アセスメント - 特筆すべき事例の比較研究－OECD教育関連ワーキングペーパー（Nusche, D.著）』国立教育政策研究所、2008年、1-57頁。（<http://www.oecd.org/dataoecd/41/8/41771582.pdf>）

Opening Remarks

Haruki Ozaki*
Director General, NIER

N.B.

- The * mark indicates that the original language of the speech was Japanese and that the transcript is a tentative translation based on the simultaneous interpretation provided during the symposium.
- The transcripts include changes made after the symposium for purpose of publication.
- The affiliations and professional titles of the speakers are as of December 10, 2013.

Opening Remarks

Director General, NIER
Haruki Ozaki



Hello, ladies and gentlemen. On behalf of the National Institute for Educational Policy Research (NIER), the organizer of the International Symposium on Educational Reform 2013, I would like to say a few words. This symposium has been held yearly since 2001, with the aim of learning from the experiences of countries by inviting the experts who are in the forefront of educational reform, and utilizing the knowledge gained for Japan's educational reforms. This year's theme, as you can see above the stage, is Tuning-AHELO Global Quality Assurance through Sharing Competence Frameworks and Degree-Level Specifications. Twenty embassies here in Japan and around 400 participants expressed interest in this theme. Thank you very much for your participation. AHELO (Assessment of Higher Education Learning Outcomes) is a study conducted by the OECD (Organisation for Economic Co-operation and Development) for university students in the final year of a bachelor's degree program. Its aim is to measure, through a common global examination, the kind of knowledge and skills they have acquired and the level of learning outcomes they have achieved through university education.

A feasibility study to review the possibility of an international learning outcome survey has been conducted in three domains: generic skills, economics, and engineering, with the participation of 17 countries. During the informal OECD meeting of education ministers held in Tokyo in January 2008, Mr. Tokai, Japan's Minister of Education, Culture, Sports, Science and Technology at that time, expressed his intention to participate in the feasibility study. He decided to participate in the domain of engineering after discussions were held in the working group regarding the OECD assessment of learning outcomes in higher education in the university subcommittee of the Central Council for Education.

As for the National Institute for Educational and Policy Research, we put our efforts into developing examination questions as a member of the AHELO's international consortium, as commissioned by the OECD. At the same time, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) designated the Institute as the national center for conducting the examination, and we acted as the secretariat for that purpose. "Tuning" represents the method individual universities use to define the competence of a student—i.e., his or her knowledge and skills, in the context of the subject areas—and build a degree program by sharing the competence framework.

Since its birth in Europe in 2000, Tuning has expanded to South America, Russia, North America, Africa, Australia, Canada, Central Asia, Thailand, and China. In the economics and engineering disciplines of AHELO, the definition of the competence framework has been set through the Tuning Method, and test questions have been created. That is why the theme of this symposium is Tuning-AHELO. The efforts of Tuning and AHELO include an extremely crucial insight into Japan's higher education policies. As you know, in Japan, in order to cultivate human resources that will drive our globalizing knowledge-based society, high expectations are placed on university education. On the other hand, the issues related to the Quality Assurance of University Education are gaining prominence, as Japan's university population is expanding. At the same time, the 18-year-old population is declining. The Japanese

educational policy has pursued the path of diversity and individuality since the end of the 1980s. However, in recent years, the weight has been shifting to the creation of a common framework that will become a platform of quality assurance.

The shift was evident in the Central Council for Education's so-called "report on quality transformation" that came out in 2012. The focus is on the necessity of establishing a degree program and setting up education management properly. The environment of Tuning in AHELO is the specific method for sharing competence frameworks and degree level specifications—important themes that are very timely.

Today's symposium has invited seven experts from overseas and Japan. I would like to extend my heartfelt welcome to all experts who are participating in this symposium despite their busy schedules. The first keynote speech will be given by Mr. Tsutomu Kimura, present advisor to the Minister of Education, Culture, Sports, Science and Technology, and former president of the National Institute for Academic Degrees and University Evaluation, a position he held for many years. He will be speaking of current issues regarding university assessment systems. The first presentation of this section will be made by Prof. Robert Wagenaar, Director of Undergraduate and Graduate Studies of the University of Groningen, who has been involved in important initiatives to harmonize European higher education, such as the ECTS (European Credit Transfer and Accumulation System), framework for qualifications of the European higher education area, and European qualification framework for lifelong learning. He will share with us the role that Tuning and competence framework should play in the global measuring and comparison of educational reform, and learning.

After Prof. Wagenaar, Dr. Peter Ewell, vice president of the National Centre for Higher Education Management Systems and chair of the AHELO Technical Advisory Group (TAG), will speak about the results of the AHELO feasibility study and the conclusions made by the technical advisory group. Next, Prof. Kikuo Kishimoto from the Tokyo Institute of Technology will talk of AHELO in Japan and its prospects. An expert in AHELO's engineering discipline, he has played a central role in the development of globally acceptable test questions and has also pushed for the implementation of the study in 12 universities in Japan.

In the panel discussions in Section 2, the facilitator will be Prof. Motohisa Kaneko, a technical advisory group member of AHELO like Dr. Ewell. Prof. Kaneko has reviewed the feasibility study. We also welcome two other panelists: Ms. Mary Catharine Lennon, a Senior Research Analyst at the Higher Education Quality Council of Ontario, and Dr. Daniel Edwards, a Principal Research Fellow at the Australian Council for Education Research.

Before we start the discussions, Dr. Satoko Fukahori, our Institution's senior researcher, will report on the AHELO feasibility study in Japan, Australia, and Canada, from the viewpoint of utilizing study results; she will also join the panel discussions. (Dr. Fukahori was the emcee at the start of the symposium.) The AHELO feasibility study was concluded last December and is currently being reviewed by OECD as to whether or not it will evolve into a main study. Regardless of its direction -- for Japan's higher education policies that aim for a systematic degree program and for the universities working on specific majors -- the insights from this Tuning-AHELO symposium will be extremely important. I expect the symposium to contribute tremendously toward the improvement of Japan's university education. With this, I would like to conclude my speech on behalf of the organizer. Thank you for your kind attention.

Section 1: Presentations

Keynote Speech

Tsutomu Kimura

Advisor to the Ministry of Education, Culture, Sports, Science, and Technology

How has the current assessment scheme been introduced and how has it been received by HEIs?

Presentation 1

Robert Wagenaar

Director of Undergraduate and Graduate Studies, University of Groningen

TUNING Competences Frameworks: Key Tools for Educational Reform and the Assessment and Comparison of Learning in Global Perspective

Presentation 2

Peter Ewell

Vice President, National Center for Higher Education Management Systems

The AHELO Feasibility Study: Study Results and the Conclusions of the Technical Advisory Group (TAG)

Presentation 3

Kikuo Kishimoto*

Dean, Graduate School of Engineering, School of Engineering, Tokyo Institute of Technology

AHELO in Japan and Its Prospects

N.B.

- The * mark indicates that the original language of the speech was Japanese and that the transcript is a tentative translation based on the simultaneous interpretation provided during the symposium.
- The transcripts include changes made after the symposium for purpose of publication.
- The affiliations and professional titles of the speakers are as of December 10, 2013.

Keynote Speech

How has the current assessment scheme been introduced and how has it been received by HEIs?

Advisor to the Ministry of Education, Culture, Sports,
Science, and Technology

Tsutomu Kimura



Good afternoon, distinguished guests, ladies and gentlemen. I'm extremely honoured to be here as the keynote speaker of this very ambitious international symposium organised by National Institute for Educational Policy Research. Before beginning my presentation, I would like to express my sincere appreciation towards the participants who have flown a long way from different parts of the world to Tokyo. I have to say this because I have been working as the chairperson of advisory board for National Institution for Educational Policy Research. In fact, I was in Brussels last week. I just got back to Tokyo Saturday last week, which is only 3 days ago. So I'm still suffering from very bad jet lag, which is a usual sort of a routine for me in these years, probably because of old age. I hope none of our foreign colleagues are in that condition.

My presentation is centred around the current assessment scheme going on in this country. It is slightly away from the topic of Tuning and AHELO, but I believe that this type of assessment is one step before tuning. As many other countries in the world, Japan has been trying very hard to achieve higher education or university reform for the past, say, 20 to 30 years. The current assessment scheme is certainly one of the products of that effort. My presentation consists of two parts. In the first part, I would like to describe how the idea of implementing new scheme has been developed in the course of higher education reform in this country. In the second part, I would like to show you how the current assessment scheme has been received by higher education institutions, by showing the results of post-event surveys carried out by National Institution for Academic Degree and University Evaluation to which I used to belong.

The majority of the proposals in the line of higher education reform in this country came from the University Council. This University Council was established on the basis of the request put forward by Emergency Council for Education. The Emergency Council for Education was very powerful. I would dare to say it is the most important inquiry committee in the history of discussions at all levels of education in this country. This Council for education was established by then Prime Minister Nakasone who was a very far sighted Prime Minister. The Council identified the different types of problems at all levels of educations. Particularly with respect to higher education, the Council pointed out that higher education is one of the weak points of Japan. This message came out around late 1970's to 1980's. At that time, Japan was boasting its prospective economy and unheard-of prosperity in many sectors. All the Japanese were very confident about everything. "Japan as number one" is the phrase which described Japan at that time. Therefore this type of critical message sent out by the Emergency Council for Education has met fierce resistance from various sectors. But I still think that the indication by the Council was exactly right. Receiving the criticism from the Council, Ministry of Education set up the University Council immediately. It was in 1987. The Council continued to exist until 2000. This Council was very active and worked very hard. The Council has produced more than 20 reports, I believe. The number of proposals they hammered out has exceeded one hundred. Amazing

activities!

The first report came out in 1988. The Council said that this country has to diversify the missions of PhD courses. Traditionally, PhD courses in this country were considered to be organizations or institutions to produce high quality researchers, or high quality academicians. But the Council bravely said that this mission should be diversified so that PhD courses can also foster highly professional human resources who can work at non-research sectors of the government, and even at industry. On receiving this recommendation, this country has carried out some improvement, but still we have a big problem in this line. Majority of the graduates of PhD courses still tend to work in research-related areas.

In 1991, a very important comprehensive report came out. I myself call this the 1st highlight report because of its comprehensiveness. This report includes many important proposals. For example, for the first time in the history of discussions related to higher education, the Council pointed out the importance of teaching at university levels. Before this time, Japanese university used to think that only mission they had was to carry out first class research. But the Council clearly said that the universities have to recognise the importance of teaching. With respect to teaching reform, the Council put forward a proposal to deregulate university curricula. This has given a big impact on the Japanese university. Immediately after the last war, this country carried out a drastic reform in curricula construction. As a result Japanese universities came to have two groups of disciplines. One is the general or liberal arts education group and the other is specialised education group. At most universities, general education was provided in the first or second year after the students' entering a university. Then specialised education follows. In the beginning, this system worked well. However as the time goes on, a very serious discontinuity between general education and specialised education began to appear. Facing this situation, the Council said that this dual system should be abolished. The Council proposed that the freedom to design the curriculum should be given to each university.

The most important proposal in this report is perhaps the indication of importance of self-check and self-assessment. I can say that this is the first sign of recognising the importance of assessment in this country. Before this, this country was not aware of the importance of assessment in higher education.

Another important proposal is the expansion of graduate programmes. Traditionally, Japanese's higher education sector placed higher emphasis on undergraduate courses. The Council said that priority should be shifted from undergraduate to graduate programmes as in other developed countries in the world. In 1993, another report came out and the Council says that all the universities in this country have to open the gate wider so that they could accept part-time students. Accepting part-time students into a Japanese university was unheard-of before this time.

Then in 1998, another very comprehensive report was put forward. I named this the second highlight report. This report also included many important proposals. It pointed out the necessity of re-examining undergraduate teaching, together with promotion of faculty development. Another important proposal is the last one shown in this slide. For a long time, Japanese universities were considered to be so-called an ivory tower. Even the government and politicians were not able to say anything critical about their performance and management. The Council was very brave in saying that officers such as the presidents and deans have to have strong leadership to realize efficient management. This indication has given a strong impact on the management system of Japanese universities. This again is the right indication. The most important proposal is perhaps the indication of importance of setting up of third party evaluation system. This report came out in 1998. As I explained earlier, we had the first sign related to the importance of assessment in 1991. So, seven years after the first sign, the Council proposed to establish the third party evaluation system. Receiving this proposal, Ministry of Education moved very fast. In the year 2000, National Institution for Academic Degrees which is a Japanese version of CNA in UK, was reorganized so that they can take the responsibility of newly established accreditation and evaluation scheme. The name of the institution was changed to National Institution for Academic Degrees and University Evaluation(NIAD-UE). The other two agencies were also recognized as the official organizations. The official name of these organizations is certified evaluation institution.

This second one is Japan University Accreditation Association of which abbreviation is JUAA. The membership of this organisation comes mainly from private universities with long history. The third one is Japan Institution for Higher Education Evaluation abbreviated as JIHEE. Its membership comes mainly from private universities founded in recent years. These three organisations were certified as the official evaluation institutions for four-year universities. Each university can choose one out of these three. The first implementation was carried out in the year 2004.

Now I'd like to summarise what I have introduced to you up to this point. Keywords are these three: diversity, individuality and third-party evaluation. The University Council kept saying that each university should have its individuality. The Council stressed the importance of introducing the third party evaluation system. The official name of the third-party evaluation was set as "certified evaluation and accreditation". It is a rule now that all the universities in this country, nearly 750 of them, have to go through this evaluation and accreditation once in seven years. Seven years is one cycle. National Institution for Academic Degrees and University Evaluation finished the first cycle in 2011. I can say that, this new scheme has gone fairly well. But there is one problem. In this scheme research assessment is not compulsory. Only teaching performance is looked at. A question arises. Can we compete with universities in other countries in the world by only carrying out assessment on teaching performance? The answer is definitely no. So what the Ministry of Education did was to devise new funding schemes, such as promotion of CEO (centre of excellence) programmes, promotion of WPI (world premier institute) and many others. Through these funding schemes, research oriented universities are now able to obtain sizable amount of research fund to keep their international level research going. What the Ministry of Education has been doing is combining these two schemes. Teaching side is strictly assessed by newly introduced evaluation and accreditation system, and by introducing the competitive funding system, the Ministry tries to enhance the quality of research.

Now I will move on to the second part. Here I would like to show you results of post event questionnaire survey carried out by National Institution for Academic Degrees and University Evaluation. Different types of question were given to the universities which had received assessment by NIAD-UE. The first category of questions is on assessment criteria set up by NIAD-UE. The first question is; "Are criteria set up by NIAD-UE appropriate in order to assure the quality of your research and teaching activities?" Many questions of this type were given. By the way, Y05 represents the year 2005 and 11 is for year 2011. This is for the first cycle. This T indicates a total number of higher education institution which have received the assessment by NIAD-UE; 5 in year Y05 and 38 in year 09. The universities were asked to give responses in five grades. Five for very positive, four for positive, three for neutral, two for somewhat negative, one for very negative. There's one thing we have to be very careful about. Culturally perhaps, Japanese are quite reluctant to give negative answers to this type of official surveys. So three, although it is neutral, we have to interpret this as slightly negative. I want you to remember this.

Let's have a look at this one. Overall responses are very good, so they are very satisfied with this. What about next question? Responses are very positive. However, when it comes to the question of, "If this criteria help your university gain societal support for their research and teaching activities?", responses are not very positive. They are doubtful about its effect. This shows that NIAD-UE has to revise the assessment criteria in this line. The next one is on the quality of self-assessment report. Each university has to write out a self-assessment report in order to receive the assessment. A question was asked; "Do you think that the self-evaluation report you wrote was for general public easy to understand?" Slightly negative response. Some universities are not very confident about the quality of their own report.

Then next. The effect of the final assessment report prepared by NIAD-UE. The response is good. But when it comes the story of understanding and the support of general public, the response is generally negative. There are some very negative responses. The next question; "Were you able to obtain new perspectives for the activities of research and teaching through the final report?" The responses are fairly negative. Some are very negative. So we can say that the final report is very good for reviewing what universities did in the past, but that it cannot give useful suggestions to the future direction.

The next is on the impact of carrying out self-assessment. The response is good, but for the question ; “Did, the majority of your staff come to understand the importance of carrying out self-assessment?” The responses are very negative. This means that universities are having a problem in persuading their members to understand the merits of this assessment scheme. We see a similar situation for the question; “Did the majority of your staff come to understand the importance of carrying out research and teaching activities systematically all across your institution?” The response conjures up a problem at Japanese universities. Traditionally, Japanese universities have a very strong autonomy particularly at the faculty level. In order to establish efficient management within a university, we have to eliminate excessively strong autonomy of each faculty. Probably, the responses we see here may have been put forward by officers of universities such as Presidents, Deans, committee chairpersons, and so on. That is a reason why responses are very negative. This shows that officers are having a difficulty in making the general staff mobile towards the direction taken by the country.

The next question; “Were you able to improve the quality of management of your institution as a whole through the final report?” Again responses are not that positive. What about the development of new challenges in research and teaching activities? Very negative. They say that certainly for the past achievements, the indications of the final report are useful, but not for the enhancement of future or new challenges.

Now I'll come to a conclusion. Generally speaking, the current assessment scheme has been well-received by higher education institutions. But it seems to be difficult for higher education institutions to convince their staff of the merits of the assessment, although the officers have understood them. I feel that the assessment criteria need to be modified so that research and teaching activities of higher education institutions can gain understanding and support from the students as well as general public. Further efforts are necessary to enhance unique and innovative future challenges of each institution. Future is the keyword.

This is the final slide. In 1998, the University Council put forward a report titled "How should higher education institutions in Japan be in 21st century?" The report stressed the importance of individual institution's effort to generate its uniqueness and individuality and of carrying out research and teaching activities in a competitive environment. These are mottos set for the university reform in this country. With respect to competitive environment, I must confess that we do not have it yet. Has the current assessment scheme I just explained created a competitive environment in higher education sector in this country? The answer is “No” . I always compare our current assessment scheme with the UK research assessment exercise and teaching quality assessment. There are big differences in the two countries' scheme with respect to creation of competitive environment. What this country has to do is to hammer out a new assessment scheme, which enables to build up a competitive environment. This is the key to success for universities reform in this country. This is my personal view. One of the solutions may be to carry out discipline-based assessment like the one UK has been doing. The key element is the quality of research. We haven't had a look at it yet. Also students learning outcomes are important. These two have to be incorporated into a new assessment scheme in order to create a competitive environment. I was in UK when they had a big fuss about the results of research assessment exercise. I think it's happened in the field of history. The story was Oxford Brookes University versus Oxford University. Oxford Brookes is a small university. The social pecking order for Oxford Brookes is not that high, but in that year, Oxford Brookes got higher mark than Oxford University. That was a big news in all the media in England. A system like this will certainly create a competitive environment. I would like to see this happen also in this country.

Thank you very much indeed, for your attention.

How has the current assessment scheme been introduced and how has it been received by HEIs?

Advisor, Ministry of Education, Culture,
Sports, Science and Technology
Tutomu Kimura

1

1. Proposals related to university reform in recent years

— From proposals of University Council —

Emergency Council for Education

「Higher education is one of the weak points of Japan」

1987 University Council founded
(1987 ~ 2000)

2

Outlines of proposals

1) 1988.12

- Fundamental issues related to graduate courses
- Diversification of the missions of PhD courses
Fostering highly-professional human resources
(Not only academics with high-profile)

3

2) 1991.2 : The 1st highlighted report

- Drastic improvement of teaching
~ Emphasis on teaching for the 1st time
- Deregulation on curricula
→ Abolishing the difference between general education and specialized education
→ Creating unique curriculum by individual university
- Self-check and self-assessment ~ The 1st sign of importance of assessment

4

- Foundation of national degree awarding institution
Moving into life-long learning society
Development of various higher education institutions
- Expansion of graduate programmes
~ Prioritizing graduate schools

3) 1993.9

- Acceptance of part-time students

4) 1996.10

- Introduction of limited-term employment for academic staff ~ Increasing mobility of academic staff

5

5) 1998.10 : The 2nd highlighted report

- Restructuring of undergraduate teaching
FD
- Advancement of teaching and research at graduate schools
- Setting up of professional graduate schools
Law schools to start with
- Introduction of efficient management system
Leadership of president and deans
Setting up of advisory board

6

- Introduction of multi-faceted university evaluation system

→ A third-party evaluation system recommended.

2000 National Institution for Academic Degrees
→ National Institution for Academic Degrees and University Evaluation (NIAD-UE)

Japan University Accreditation Association (JUAA)
Japan Institution for Higher Education Evaluation (JIHEE)

~ Certified evaluation institutions

2004 The introduction of the certified evaluation and accreditation system.

7

Half-way summary

Diversity, Individuality, Third-party evaluation

Certified Evaluation and Accreditation (7 years' cycle)
Research evaluation ~ Not compulsory
Others (e.g. teaching, management) ~ Absolute : In reference to each institute's own missions

Can we compete with universities in other countries? Definitely no!

Introduction of competitive research grants such as COE programme or World Premier Institute

8

2. Results of post-event questionnaire survey by NIAD-UE

9

1. Assessment criteria set up by NIAD-UE

- Appropriate in order to assure the quality of R&T activities

	5	4	3	2	1	T	
Y 05	0	4	1	0	0	5	5~Very positive
Y 06	3	8	0	0	0	11	3~ Neutral
Y 07	5	29	5	0	0	39	1~ Very negative
Y 08	11	2	0	0	0	13	
Y 09	2	31	4	1	0	38	
Y 10	8	18	4	0	0	30	
Y 11	0	7	0	0	0	7	
T	29	99	14	1	0	143	

- Appropriate in order to accelerate improvement of the quality of R&T activities

	5	4	3	2	1	T	
Y 05	1	3	1	0	0	5	
Y 06	2	8	1	0	0	11	
Y 07	5	31	3	0	0	39	
Y 08	1	11	1	0	0	13	
Y 09	4	28	5	1	0	38	
Y 10	7	20	3	0	0	30	
Y 11	0	7	0	0	0	7	
T	20	108	14	1	0	143	

10

- Appropriate in order to gain societal support for R&T activities

	5	4	3	2	1	計
Y 05	1	3	1	0	0	5
Y 06	1	6	4	0	0	11
Y 07	4	19	14	2	0	39
Y 08	1	9	3	0	0	13
Y 09	1	21	15	1	0	38
Y 10	4	15	11	0	0	30
Y 11	1	5	1	0	0	7
T	13	78	49	3	0	143

11

2. Quality of self-assessment report

- Was able to produce a report for general public easy to understand

	5	4	3	2	1	T
Y 05	2	1	2	0	0	5
Y 06	0	8	3	0	0	11
Y 07	6	23	8	2	0	39
Y 08	11	0	2	0	0	13
Y 09	8	20	9	1	0	38
Y 10	4	17	9	0	0	30
Y 11	2	5	0	0	0	7
T	23	84	33	3	0	143

- Was able to produce a satisfactory report

	5	4	3	2	1	T
Y 05	2	2	1	0	0	5
Y 06	0	9	2	0	0	11
Y 07	5	27	6	1	0	39
Y 08	0	11	2	0	0	13
Y 09	5	23	8	2	0	38
Y 10	6	20	4	0	0	30
Y 11	2	5	0	0	0	7
計	20	77	23	3	0	143

12

● Improved the quality of R&T activities

	5	4	3	2	1	T
Y 05	2	1	2	0	0	5
Y 06	1	8	0	2	0	11
Y 07	3	25	11	0	0	39
Y 08	0	12	1	0	0	13
Y 09	7	17	12	2	0	38
Y 10	2	21	7	0	0	30
Y 11	0	5	2	0	0	7
T	15	89	35	2	0	143

● Promoted new challenges in R&T activities

Y 05	0	4	1	0	0	5
Y 06	1	4	5	1	0	11
Y 07	2	13	22	2	0	39
Y 08	1	8	4	0	0	13
Y 09	3	13	19	3	0	38
Y 10	3	13	14	0	0	30
Y 11	0	4	3	0	0	7
T	10	59	68	6	0	143

19

5. Outcomes of this assessment if it continues

● Will assure the quality of R&T activities

	5	4	3	2	1	T
Y 05	1	2	2	0	0	5
Y 06	2	6	3	0	0	11
Y 07	6	21	12	0	0	39
Y 08	1	11	1	0	0	13
Y 09	7	27	3	1	0	38
Y 10	9	17	4	0	0	30
Y 11	0	5	2	0	0	7
T	26	89	27	1	0	148

● Will be able to obtain understanding and support of students for R&T activities

Y 05	1	1	3	0	0	5
Y 06	1	3	7	0	0	11
Y 07	1	11	26	1	0	39
Y 08	0	5	7	1	0	13
Y 09	0	12	23	3	0	38
Y 10	3	8	17	2	0	30
Y 11	0	3	4	0	0	7
T	6	43	87	7	0	143

20

● Will be able to obtain understanding and support of general public for R&T activities

	5	4	3	2	1	T
Y 05	0	3	2	0	0	5
Y 06	1	4	6	0	0	11
Y 07	2	15	21	1	0	39
Y 08	0	8	3	2	0	13
Y 09	3	15	17	3	0	38
Y 10	5	10	14	1	0	30
Y 11	0	4	3	0	0	7
T	11	59	66	7	0	143

● Will promote new challenges in R&T activities

Y 05	2	2	1	0	0	5
Y 06	1	4	6	0	0	11
Y 07	1	21	16	1	0	39
Y 08	3	6	4	0	0	13
Y 09	3	13	20	2	0	38
Y 10	3	13	14	0	0	30
Y 11	0	5	2	0	0	7
計	13	64	63	3	0	143

21

Summary

Generally speaking, the current assessment scheme has been well received by HEIs, but it seems to be difficult for HEIs to convince their staff of merits of the assessment. Assessment criteria need to be modified so that R&T activities of HEIs can gain understanding and support from the students and general public. Further efforts are necessary to enhance unique and innovative future challenges in R&T activities of HEIs.

22

In 1998 University Council put forward a report titled “How should HEIs in Japan be in 21st century ? — Proposals of reform plans—” The report stressed the importance of individual institution’s effort to generate uniqueness and individuality in its R&T activities in competitive environments. Has the current assessment scheme created competitive environments in the HE sectors in this country ? The answer is probably no. What the country has to do next is to hammer out a new assessment scheme which enables to build up competitive environments. One of the solutions may be discipline-based assessment like the one UK has been doing. Key elements are quality of research (R) and students’ learning outcomes (T).

23

Thank you so much for your kind attention

24

Presentation 1

TUNING Competences Frameworks: Key Tools for Educational Reform and the Assessment and Comparison of Learning in Global Perspective

International Tuning Academy, University of Groningen

Robert Wagenaar



Introduction

During the last two decades the playground of Higher Education has changed fundamentally. Globalization and Information and Communication Technology have given a further push to its internationalization. The required level of quality and effectiveness of higher education programmes are no longer determined at local or national level only, but are also referenced internationally today. This does not imply that it is felt like that by the majority of academic staff. One can observe a striking difference between opinions and viewpoints of governmental authorities, management of higher education institutions, its faculties and departments and academic leaders on the one hand and the average academic on the other. One can also see differences between world regions.

Nevertheless, whether one likes it or not, in particular for the more renowned institutions in every country competition in terms of attracting academic staff, young researchers / PHD-students and master and bachelor students has moved from the national to the international scene. Academics as well as students identify and select higher institutions which serve their interests best. Due to search engines, portals, web presentations, and new forms of communication such as Facebook and other social media, this process has become a global one. As a result the student body has changed in a large and growing number of institutions as has often – partly - the language of instruction. Academics are noticing that students have become more demanding regarding the content of educational programmes as well as the learning and teaching process. Not only because they have a wider and easier access to information about higher education institutions, what is taught there, how it is taught and by whom, but also because this information is easily exchanged via social media. This involves risks, for all involved, because images and reputations can easily be built or destroyed, rightly or wrongly. This is also the case for the quality of education which is offered. At present, there is more focus on this issue than ever before. Universities develop so-called quality cultures as a result of external pressure in particular. A key question in this respect is what and who decides what high quality programmes are and on which basis?

Besides what has been said above, and besides personal development, there is another dimension which has a growing impact on the content, implementation and modes of delivery of a higher education programme: its relevance for society. Relevance is understood here in terms of preparing for citizenship and for employability. In particular employability seems to be given more and more weight and is therefore competing with the actual interests and abilities of the student. This is understandable in a situation where an economic global crisis is impacting so many, but it might lead to wrong choices, possibly followed by a growing number of drop-outs. This can be a costly affair in more than mere financially. Although the chances for obtaining employment at a suitable level after graduation for one programme might be better than for another, this does not imply that less successful programmes in this respect

do not have an obligation regarding the transition to society of its graduates. Education is simply not intended to be 'art for the sake of art'. Degree programmes are also not intended to mirror the academic profiles of the teaching staff itself in today's dynamic world. This has implications for the design and delivery of programmes as well as the for the competences which are developed and the desired outcomes.

Searching for a new paradigm

In 2001 the Tuning Educational Structures in Europe project (hereafter: Tuning) was launched with these notions in mind. At the end of the 1990th a growing concern developed among the initiators and developers of the European Credit Transfer and Accumulation System (ECTS) that one credit system would not be a sufficient answer to societal challenges and was not the only panacea for (trans)national mobility of students and cooperation between higher education institutions in and outside Europe. It was concluded that the emphasis should be much more on the outcomes of the learning process, but also that more attention should be given - within degree programmes - to transferable or general skills. It was also thought necessary to stipulate the role of higher education institutions as major contributors to the welfare of society. This fitted in a trend in which higher education institutions were forced to show that they are accountable, responsible, and sustainable. This should not apply to the higher institution only but also to the individual degree programmes on offer. A number of conclusions were drawn from this analysis:

- ▶ Higher Education structures and programmes and qualifications should be reformed on a large scale to be able to respond to the needs of society;
- ▶ Academics should be given a key role in this process;
- ▶ The reform process would require the development of internationally shared reference points / standards at disciplinary / subject area level;
- ▶ A language for communication should be developed which would be understood by all major stakeholders, that is academics, students, graduates, (potential) employers of graduates as well as professional organisations;
- ▶ All stakeholders, including (potential) employers and professional organisations, but in particular graduates should be (indirectly) involved in the process of curriculum design and quality enhancement;
- ▶ The focus should be on diversification of degree programmes by profiling and stimulating flexibility;
- ▶ The reform should facilitate national and international mobility and the recognition of periods of study, including qualifications for obtaining access to the next level of programmes.

The concept of 'competences' was chosen as a means to communicate with stakeholders despite awareness of the different meanings throughout the world attributed to the term. The basic idea was and is that the role of education is primary to make the student / learner more competent as a result of a learning process. Competences should be understood - according to Tuning - as a representation of a dynamic combination of cognitive and metacognitive skills, knowledge and understanding / insight, interpersonal, intellectual and practical skills, and ethical values. In other words, high level competences required to operate with confidence and success in a leading capacity in society. Besides competences Tuning also introduced as part of its methodology the concept of learning outcomes. Learning outcomes in Tuning state the level of competences to be developed in a course unit, module or degree programme.

From the onset Tuning made a distinction between general or generic competences and subject specific (disciplinary) competences. This was done to raise awareness about - in particular - the generic competences. Although they might have been taught already as part of a degree programmes, it was noticed at the time that in most cases they were not made explicit in the course material. This may have changed somewhat over time but it still seems to be an important point of attention in many degree programmes.

Having said this, a more far reaching conclusion was drawn from the analysis made; the reform process would require a paradigm shift in the teaching, learning and assessment process. As a consequence of focusing on the outcomes of the learning process in terms of the competences to be developed, the switch should be made from input

or content based learning to outcomes or output based learning. To phrase this differently: to move from a more staff-centred approach, based on the concept of learning objectives to a concept of student oriented learning, based on the concept of learning outcomes. A learning outcome is understood here as *a statement of what a learner is expected to know, understand and be able to demonstrate after completion of a process of learning*. A learning objective outlines the material the teaching staff intends to cover or the questions related to the discipline that the class will address. This last approach means in practice that the focus is on the teaching process (not the learning process) and on knowledge transfer of the academic staff member to the students. Student centred learning is an approach or system that supports the design of learning programmes which focus on the learners' achievements, accommodates different learners' priorities and is taking into consideration student workload (i.e workload that is feasible within the duration of the learning programme). It accommodates for learners' greater involvement in the choice of content, mode, pace and place of learning.

From this perspective the Tuning initiative developed two main action lines: a methodology or approach to develop current, high quality and relevant degree programmes for all levels, bachelor, master and doctorate, and internationally established reference points or subject area based competences frameworks to support the (re-) design, implementation, delivery and quality enhancement of degree programmes. This paper focuses on the latter.

TUNING competences frameworks

Having outlined above that degree programmes are no longer developed and delivered in its own right (only), but should be referenced against at least national, but preferably internationally established reference points, Tuning took the initiative for developing international competences frameworks. More or less at the same time the Quality Assurance Agency in the UK initiated the development of national based benchmark papers. In both cases a bottom-up approach was applied by making a group of academics responsible for setting-up a subject area based framework, respectively benchmark. The approach applied was slightly different, although in both cases it was based on a process of discussion and reflection. With respect to Tuning, lists of key competences were formulated which were used as a basis for consultation among stakeholders. For all subject areas / disciplines involved in the Tuning project at the time a common list of 30 generic competences was drawn up and for each individual subject area a subject area list of key competences was produced (see below). A consultation of stakeholders took place in 2001-2002 and was repeated in 2008 with more subject areas (9 instead of 7) and with a slightly adjusted and improved list of generic competences. The stakeholders consulted were academics, graduates, employers and in 2008 also students.

The consultation was based - in both cases - on different variables, that is, first:

- the degree of importance: the relevance of the competence, in the opinion of the stakeholder (for work in their profession);
- the level of achievement: the achievement of this competence as a result of having taken this university degree.

To evaluate these two variables, the respondents had to use a scale: 1 = none; 2 = weak; 3 = moderate; 4 = strong. In addition they had to:

- rank the generic and subject specific competences; based on the categorisation of the five most important generic and the five most important subject specific ones. The competence that was ranked highest in the survey was allocated five points, four for the second and so on, with one point for the last in the selection. If the competence was not chosen in the survey, it scored zero points.

The outcomes of the consultation were very informative and revealing and were used as input for producing the Tuning competences frameworks for the subject areas involved. What was learned in particular was the need in society for graduates with better developed generic competences. Although doing well regarding the acquisition of knowledge of and insight into the subject area, students did less well than thought necessary in developing abstract thinking, analyzing and synthesizing skills (seen by the respondents as the most important competence) as well as

applying knowledge in practice, problem solving, learning abilities, and written and oral communication skills. Also more attention should be given – in particular according to employers and graduates - to leadership and teamwork competences in the learning process. Furthermore, it was advised to offer more attention to creativity and the development of an entrepreneurial spirit. These kind of consultations of the same stakeholder groups were repeated in other regions of the world with adjusted lists (based on the situation, culture and opinions of the academics of that region that drew up the lists) but with more or less comparable outcomes.

The consultations and competence lists were only one of the basic elements for developing the competences frameworks. Others were and are a description of the academic field country by country and synthesized at regional level– at first instance European and later also other regions of the world. Furthermore the typical degrees offered were mapped, as well as the social and professional needs, including the (potential) employability field, and if possible future trends were identified. The last items showed us differences between countries and higher education institutions. It was a confirmation that degree programmes have and should have different profiles, in particular at master level, and that they are also partly regionally bound.

Constructing subject area based competences frameworks

Since 2001 Tuning has built up much experience regarding the construction of (inter)national competences frameworks. As part of the European Tuning projects a large number of documents called *Reference Points for the Design and Delivery of Degree Programmes in [name of subject area]* - covering a large set of disciplines by now - have been developed. These were validated by external peers before being finalized and published. In the European context the term *Reference points* was used instead of, for example the term ‘standards’ (which is in use in Russia and Australia to express the same idea). Although Tuning intends to be standard setting, it wants to avoid at any cost that these reference points are regarded as prescriptive. In the dynamic world of today, they simply cannot be set in stone and therefore revision is needed every six years – preferably on the basis of a new consultation round of stakeholders. The development of the reference points was based on a common format, independent of the subject area involved.

For each competences framework to be developed, a group of 12 to 15 international experts was established. Its members were selected from different countries and were representing their higher education institution in the field involved. At its start Tuning focused only on traditional research oriented mono-disciplines to be able to develop its approach and to avoid unnecessary complications. At a later stage multi-disciplinary and interdisciplinary degree programmes as well as more applied disciplines were covered as well.

The process started with the mapping process as described above and the consultation of stakeholders. As was already explained, to prepare for the consultation process a common list of generic competences was drawn up by the different groups together and each group developed its own list of so-called key subject area competence statements which should ‘frame the subject area’. This last list contained on average of 25 statements. That list was established on the basis of a collection of ideas and expectations regarding degree programmes in the subject area based on an open discussion. On the basis of the long list consensus was sought on the short list (key competences) to be used for the consultation process. Because the focus was on competences to be covered in the discipline as a whole, the first cycle / bachelor and the second cycle / master were not distinguished. To allow for the consultation process, each university draw up a list of relevant employers for its field, a list of graduates which graduated within the last 3 to 5 years and a list of academics to be consulted. In the first European consultation round (2001-2002) no (mature) students were consulted. The Tuning America Latina project decided in 2004 to involve also students. This has been standard procedure since. After finishing the consultation process, its results were analysed by each subject area group. This led to a redefining of the original subject specific competences lists. The next step in the process was to design academic and professional frameworks for each of the cycles, preferably formulated as descriptors and to

build consensus on the most relevant competences for each of them, combining both what is common for academic recognition and what is different (the specific features). It was also checked whether the subject area frameworks were consistent with another recent development at the time, the establishment of a Qualifications Framework for the European Higher Education Area. This framework was endorsed in Europe by the ministers of Education in 2005 as part of the Bologna Process. Below we will come back to the complementary relationship of these so-called overarching or meta-frameworks and the subject area based frameworks.

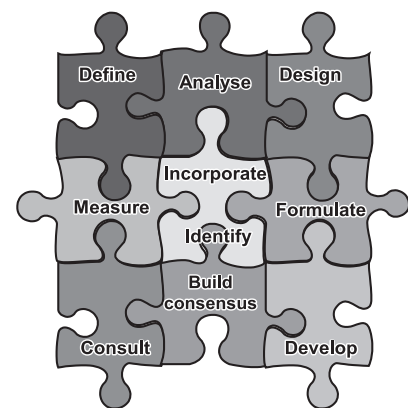
The following step in the process of establishing competences frameworks was to check whether these frameworks allowed for a realistic student workload. In other words: to ‘measure’ the required student workload to achieve the desired competences levels, to be expressed in terms of credits (besides learning outcomes). This step was followed by an open discussion to identify the most suitable modes and approaches for teaching, learning and assessment to develop the competences identified. This reflection took place against the proposed change of paradigm – student centred learning – as described above. The outcome was an overview of different possibilities and strategies for each of the subject areas from which the academic staff responsible for offering the degree programmes in each institution could make their own selections and combinations. The last step was to relate the competences to consistent mechanisms for quality control and enhancement.

This approach can be visualized in the following image:

As part of the Tuning America Latina 2 Project (2011-2013) ideas about the development of competences frameworks were further refined and deepened, by introducing the concept of meta-profiles. Both reference points and meta-profiles have the aim to identify and describe the core (elements) of a discipline / subject area. However, the meta-profile approach is slightly different. While in the original approach the focus is on identifying the core or key competences, both generic and subject specific, in the second approach the focus is on the clustering of generic and subject specific competences to derive to so-called meta-competences. The interlinked groups of meta-competences then serve as the basis for defining a meta-profile (competences framework) which captures the essence of the discipline in more general terms. When the meta-profile is decided, it can be used as a template for constructing individual degree programmes.

This new method which was and is being applied later in Tuning projects in other regions of the world such as Russia, Africa and Central Asia, offers us a more sophisticated way forward, because the existing template of collecting a long list of competences and then boiling it down to the more essential ones, lacked sufficient structure.

Two main approaches have been developed for the grouping of competences. The first way is to cluster the most related competences in a feasible number of groups, minimum 5 to 8 maximum. After having done so, a label for each group is decided which reflects best its content and purpose. However, it is also possible to work the other way around, defining labels for the group’s first and then using these as a basis for clustering the competences. Each group or meta-competence will contain a mix of generic and subject specific competences. This is fully in line with the Tuning philosophy, which requires that these are developed together. Although named generic and transferable skills or competences these are or should always be developed in conjunction with the main field(s) of study, and not in isolation. To illustrate this point: the competence *abstract thinking, analyzing and synthesizing* in history is based on a different theoretical and methodological framework than for example the one used in physics or mathematics. This also applies - in general - for oral and written skills, leadership, teamwork, entrepreneurial spirit etc. because each discipline has its own academic culture and paradigm.



Overarching competences frameworks

Competences frameworks are not only a fundamental basis and reference for the design of new and improvement of existing curricula, but also for the assessment and comparison of learning in a national and global context. Different levels of competences frameworks can be distinguished. While Tuning has focused on the subject area level, others, such as governments and quality assurance organisations, have initiated overarching or meta-frameworks which intend to cover all levels of learning or part of them. According to Tuning these complementary frameworks are a requirement for the restructuring and/or enhancing of the quality of higher education sector and its qualifications.

Qualifications Frameworks are not a new phenomenon, but in recent years their content and structure have changed. In principle every country has its qualifications framework or system which – in the past or still - describes the different types of qualifications offered, its interconnections, as well as pathways to progress within the system. This traditional model has been replaced now in many countries by a model which focuses on the outcomes of a learning process, be it formal or informal. As its basis descriptors are used which describe the expected / achieved learning in terms of competences and (indicated in terms of level) in learning outcomes. This new type of qualifications framework has been developed at transnational level (Europe) as well as national level.

At European level two meta-frameworks have been developed which are the Qualifications Framework for the European Higher Education Area (QF for EHEA), initiated as part of the Bologna Process, and the European Qualifications Framework for Life Long Learning (EQF for LLL). Both can be defined as an overarching framework that makes transparent the relationship between European national higher education frameworks of qualifications. The first one is based on the so-called Dublin Descriptors. These were developed in the same period in which Tuning developed its first competences frameworks (2001-2005). In March 2002 at the official Bologna seminar *Working on the European Dimension of Quality*, which took place in Amsterdam, the main conclusion was that general descriptors for the different cycles and reference points at subject area level should go together. To quote from its published report: “There is a widely-shared consensus that the ‘Dublin Descriptors’ , defining key outcomes for Bachelors and Masters programmes in general are useful. These generic descriptors are complementary to the more specific outcomes of the Tuning project (…), which have been developed at the level of areas of knowledge (‘disciplines’)” . The Dublin descriptors outline the essential components of any degree programme that leads to the completion of a Bologna cycle, that is bachelor, master and doctorate. They are based on the following *dimensions*:

- Acquiring knowledge and understanding
- Applying knowledge and understanding
- Making informed judgments and choices
- Communicating knowledge and understanding
- Capacities to continue learning
- Contributing to original research (doctorate only)

Besides the ‘Bologna’ QF- EHEA, the European Union established the EQF for LLL which contains eight levels, covering learning achievements at all educational levels. As far as higher education is concerned, the top 4 levels (that is 5, 6, 7 and 8) in this Framework are compatible with the three cycles, plus the short cycle (Associated Degree), included in the Qualifications Framework for the EHEA. Hence, their outcomes correspond to those foreseen in the Dublin Descriptors, even though they are expressed in a slightly different language. The EQF distinguishes three main categories to order its outcomes based descriptors: *knowledge, skills and competences*. Competences should be read here as wider competences, intending to describe responsibility and autonomy.

Within each subject area, discipline or professional sector, the QF for EHEA Dublin Descriptors and /or the EQF for LLL level descriptors can be applied and adapted according to the specific way that learning is acquired in that sector. Thus the Dublin Descriptors/EQF descriptors form general reference points at the European level in which any specific Degree Programme can be situated.

In particular during the last five years, the development of competences / learning outcomes based National Qualifications Frameworks (NQFs) has gathered considerable momentum. To establish such a framework, each country sets out its own qualifications framework according to its educational structures and traditions. A National Qualifications Framework can be described as a single description, at national level or level of an education system (for example, Scotland, England and Wales), which is internationally understood and through which all qualifications and other learning achievements may be described and related to each other in a coherent way. Good examples are the qualifications frameworks of Australia, South-Africa, Thailand, a growing number of European countries, as well as a model developed by the Lumina Foundation for the USA. A national framework enables students, employers and quality assurance and accreditation agencies to ‘compare and contrast’ the learning achievements of students and benchmark (i.e. position) them according to other national and regional frameworks.

The Australian Qualifications Framework (AQF) distinguishes – with regard to the higher levels – the following items: purpose and volume of learning (expressed in time) besides the categories knowledge, skills, applications of knowledge and skills. The Lumina USA Degree Qualifications Profile is – as the QF for the EHEA – based on dimensions: Broad, Integrative Knowledge, Specialised Knowledge, Intellectual Skills, Applied Learning and Civic Learning. In this last model more emphasis is given to the learning process as a classifying principle.

All these overarching competences frameworks have in common that they are based on broad descriptors which are general by definition because they include all types and orientations (applied, research driven etc.) of qualifications or other learning achievements.

Bridging meta-competences frameworks and Tuning subject area based competences frameworks

In 2008 the Tuning management team concluded that it would be useful to bridge the gap between the meta-level and the subject area level by developing an intermediate framework level by grouping academic programmes in terms of domains or sectors. A sector or domain is understood here as a combination of related fields of study which are based on more or less comparable learning profiles. Tuning distinguishes six sectors: Humanities, Creative and Performing Disciplines, Engineering, Natural Sciences, Health Care and Social Sciences. The basic aim for developing Sectoral Qualifications Frameworks (SQFs) was and is to produce a common set of statements about expected achievement levels for students in any and all of the disciplines represented, as well as seeking to define what these disciplines have in common. The framework intends to show what sets them apart from disciplines in other sectors and give them their distinctive character as a sectoral grouping. At the start of the project period there was much doubt among the disciplinary experts whether it would be possible to find sufficient common ground. During the project these hesitations were slowly replaced by enthusiasm about the opportunities a SQF would offer.

From 2008 to 2010 a first project was implemented: the development of a Tuning Sectoral Qualifications Framework (SQF) for the Social Sciences. The sector was represented by the following academic areas: Business Studies, European Studies, Education Sciences, Occupational Therapy and Social Work, Law, Psychology and International Relations. The project designed a SQF which covers the higher education sector levels 5 to 8 as well as the preceding levels 3 and 4. Its development proved to be a pioneering and innovative experience, but most of all a major step forward towards linking the different initiatives – QF for the EHEA, EQF for LLL, NQFs and Tuning Subject Area based Reference Points, - so far.

The Social Sciences SQF-project was followed by a project in which two SQFs were developed, one for the Humanities and another one for the Creative and Performing Disciplines. While the Social Sciences SQF project group limited itself to define competences statements covering the broad categories of knowledge, skills and (wider) competences, the other two SQF project groups went a step further. Initiated by the Creative and Performing Disciplines *dimensions* were identified to organize the competences statements. The catalyst for generating mutual

belief in the validity of a joint sectoral framework came from the realisation of the central role played in the Creative and Performing Disciplines by the dimension of *Creativity* (making, performing, designing, conceptualising) in all of the disciplines represented. Also for the Humanities such a central dimension could be identified in the concept of the *Human Being* (human condition, experience and expression).. In the framework of the OECD TUNING-AHELO project which aimed to develop a *Conceptual Framework of Expected/Desired Learning Outcomes in Engineering* (2009) a set of dimensions for organizing the competences was decided as well. It is intended to develop SQFs based on dimensions for the other three other sectors in the near future. The following table offers an overview of the dimensions identified for the three sectors covered so far:

Table 1: SQF Dimensions

Humanities <i>Dimensions</i>	Creative and Performing Disciplines <i>Dimensions</i>	Engineering <i>Dimensions</i>
The Human Being	Making, Performing, Designing, Conceptualising	
Cultures and Societies	Re-thinking, Considering and interpreting the Human	Basic and Engineering Sciences
Texts and Contexts	Experimenting, innovating & Researching	
Theories and Concepts	Theories, Histories and Cultures	Engineering Analysis
Interdisciplinarity	Technical, environmental and Contextual issues	Engineering Design
Communication	Communication, Collaboration & Interdisciplinarity	Generic Skills
Initiative and Creativity	Initiative & Enterprise	Engineering Practice
Professional Development	Professional Development	Professional Development

The development of Tuning Sectoral Frameworks based on dimensions seems to be a breakthrough. They not only bridge the two European meta-frameworks and national frameworks globally but also the meta-profiles / reference points at subject area level. The three sectoral frameworks or profiles developed so far offer the necessary precision which is required for degree programme design, delivery, quality assurance and enhancement and the recognition of degrees and periods of studies.

The following table offers descriptors of the first two dimensions of the Tuning SQFs for the Humanities and the Creative and Performing Disciplines. It illustrates the kind of information about what is actually (or should be) taught and learned in the framework of degree programmes which belong to the sectors involved.

This information should be used as a reference for the designing of individual degree programmes. The expression 'reference' is used here to express that each degree programme will and should have its own individual profile. The profile of a degree is not only defined by the competences framework but also by the mission of the institution, the role foreseen for its graduates in society, the particular strengths of the department offering the degree in terms of particular expertise(s), as well as the financial means available. According to the philosophy of student centred learning and therefore that of Tuning, degree profiles should be developed by the team of academics involved in offering the degree / qualification. This team, which includes preferably one or more mature students, formulates the learning outcomes followed by the outline of the degree programme and its individual modules. This structure should

Table 2: Relation EQF and Tuning SQF dimensional approach

Sector at level 6 EQF	Knowledge	Skills	Competences
EQF	Advanced knowledge of a field of work or study, involving a critical understanding of theories and principles	Advanced skills, demonstrating mastery and innovation, required to solve complex and unpredictable problems in a specialised field of work or study	<ul style="list-style-type: none"> • Manage complex technical or professional activities or projects, taking responsibility for decision-making in unpredictable work or study contexts • Take responsibility for managing professional development of individuals and groups
CPD: Making, Performing, Designing, Conceptualising	Have advanced knowledge of the processes and concepts underlying creation and/or performance in their specific discipline	Have the advanced skills necessary to create, realise and express their own creative concepts	Be able to draw upon the knowledge and skills gained within their studies to act and respond creatively in different situations
CPD: Rethinking, Considering and Interpreting the Human	Appreciate how the practice and/or creation generated within their discipline both stems from, and shapes, our humanity	Demonstrate interpretative skill and a reflection of the human dimension in their creative practice	Be able to draw upon experience gained within their studies to operate with an ethical awareness and to encourage the development and foster the well - being of other individuals and groups
HUM: The Human Being	Have a critical understanding of the human condition, experience and expression in its various forms and environments	Be able to use disciplinary knowledge to understand and interpret contemporary societal challenges	Be able to understand and respect the individual human in his/her personal, cultural and social dimension
HUM: Culture and Societies	Have knowledge and critical insight into how human behaviours, institutions and modes of expression emerge from and interact with ideas, beliefs and values	Be able to draw on knowledge of the relevant field to identify and define, with guidance, significant problems and areas of enquiry with respect to social and cultural interaction	Be aware of the role of humanities and a humanistic perspective in society, and demonstrate an ethical commitment to their use to achieve social cohesion and sustainability

allow for covering the development and achievement of all learning outcomes defined in such a way that in terms of student workload the programme is realistic and feasible. For the development of new programmes and enhancement of existing programmes Tuning has developed a ten steps approach, which is included in the publication *A Tuning Guide to Formulating Degree Programme Profiles*. Bilbao, Groningen, The Hague, 2010.

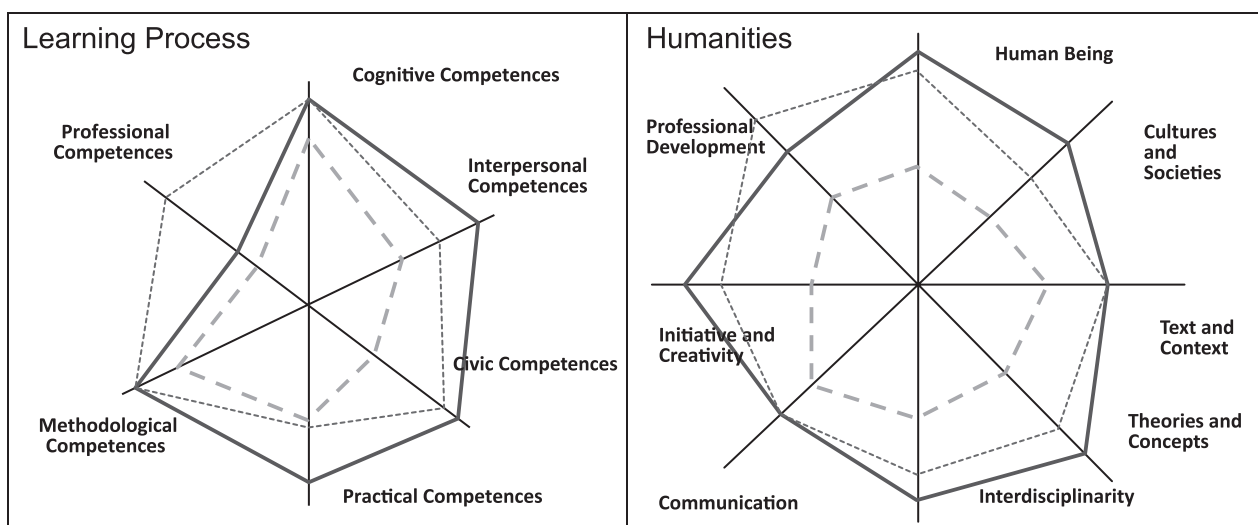
Tuning competences frameworks: tools for comparison and assessment of learning

The AHELO project has shown us that a Tuning competences framework is the foundation of an assessment framework for allowing comparisons regarding the level of achievement of learning in a particular subject area / discipline in transnational perspective. An assessment framework measures the level of achievement of individual learners but also of a system. This system can be the department within an institution, a local or regional system or a national system. In combination with meta-competences frameworks and sectoral frameworks, the Tuning competences frameworks are also important tools for recognition of periods of learning as well as of qualifications.

With the exception of the TUNING-AHELO conceptual frameworks, which were developed in a global setting, all other Tuning competences frameworks were defined regionally: Europe, Latin America, Africa, Russia. We can observe that in the different Tuning projects two main types of meta-profiles, or combinations of these have been developed. The first one focusses on the learning process; the second one derives from the subject area or an academic sector. Both are based on dimensions.

The dimensions identified for the 'learning process approach' are for example: Cognitive Competences, Interpersonal Competence, Civic Competences, Practical Competences, Methodological Competences and Professional Competences. The dimensions of the subject area approach are reflected in the table above.

Both models can be used to develop and to position individual degree programmes according to their own profile. These can be visualized as spiderwebs against the background of agreed competences statements of the Tuning subject area based competences frameworks. Depending on its mission and role and its academic strengths the different dimensions identified are more or lesser developed than the levels indicated in the meta-profile. This web visualizes individual profiles of a bachelor, a master and a PhD-programme:



Legend:
 Green (broken line) = First cycle / Bachelor
 Red (dotted line) = Master
 Blue (solid line) = Doctorate

As was stated earlier, Tuning has developed competences frameworks already for a large number of subject areas. All these frameworks cover the first and second cycle of higher education, that is the bachelor and master, and often also the doctorate. It is important that frameworks are developed which cover more than one level. It forces the designers / experts to decide on the level of achievement that can or should be developed for each of the competences levels involved. It will guarantee that a good balance is defined. Below, an example derived from the subject area of Architecture, is offered. This academic field decided to develop its competences framework on the basis of the SQF for the Creative and Performing Disciplines, not the one for Engineering. The reason for this

choice was that the experts of the European Tuning subject area group of architecture identified themselves much more with the dimension of creativity than the dimension of modelling (typical for Engineering). Descriptors were developed for the levels 6 to 8 of the EQF, covering bachelor, master and doctorate.

Table 3: Levels of mastering competences, the example of Architecture:

ARCHITECTURE			
EQF Categories:	Knowledge	Skills	(Wider) Competences
Architectural Creation & Architectural Creativity			
Level/ Dimension	<i>Graduates of Architecture are expected at level ...</i>		
6 / Conceptualising, Designing, Materialising Architecture	To have advanced knowledge of the processes, concepts and cultural values guiding architectural creation	To have the advanced skills to formulate critically , elaborate creatively and translate innovatively into spatial forms their own architectural concepts	To be able to draw upon the knowledge and skills gained within their architectural studies so as to act and respond critically and creatively in different situations
7 / Conceptualising, Designing, Materialising Architecture	To have highly advanced knowledge of the processes, concepts and cultural values guiding architectural creation, some of which will be at the forefront of their field	To have developed to a high professional level their ability as architects to formulate critically , elaborate creatively and translate innovatively into spatial forms their own architectural concepts	To emerge as well-developed personalities, able to draw upon the knowledge and skills gained within their architectural studies so as to act and respond critically and creatively in situations that are complex, unpredictable and require new strategic approaches
8 / Conceptualising, Designing, Materialising Architecture	Knowing in depth all the relevant methods and techniques of inquiry related to a particular field of study of architecture	Integrating previous experience so as to demonstrate original creative insights in the domain of architecture Functioning with complete creative autonomy	Comprehending the transferability of their research capabilities to other fields Displaying professional, creative and scholarly integrity

From the above it can be learned that different phrasing is used to indicate each level. The Tuning experience has learned – as is also shown in the table - that level is indicated by complexity and scope.

When developing a measurement instrument it is not only very important to have agreement on the level of mastery or achievement. This level will be expressed in the form of a learning outcome for that particular competence. That learning outcome, however, will still be rather general, and therefore it is necessary to break it down in (for example 3) indicators. These indicators in turn are supported by (for example 5) descriptors or assessment criteria, which cover the level of attainment. This subdivision is made for each generic and subject area / disciplinary competence to be assessed / measured. Also in this respect, Tuning has developed examples of good practice. With regard to the generic competences the publication of Aurelio Villa Sánchez & Manuel Poblete Ruiz, eds., *Competence-based learning. A proposal for the assessment of generic competences*. Bilbao, 2008 is referred to. It is also advised to consult this publication when developing or enhancing an individual degree programme.

Conclusion

This paper started with the assumption that the higher education sector is challenged by dynamisms of society, such as globalism and information and communication technologies, which have raised (global) competition between higher education providers in terms of outcomes of achievement and level of appropriateness of what has been learned by its graduates to serve today's society/ies.

An analysis of the situation has shown that higher education institutions should take their responsibility to respond to the needs of society. These needs are most of all well-educated graduates which have the knowledge, skills and competences to operate successfully in that society. To educate such graduates a change of paradigm of the educational model in use is required, which means in practice a move from a staff-centred approach to a student oriented approach. This last model implies that stakeholders groups are involved in the design, delivery and enhancement of the quality of degree programmes. It also implies that a language is used which is well understood by all parties. Such a language has been proposed and developed by Tuning, by using the concepts of competences and learning outcomes. Furthermore, Tuning has outlined a methodology for (re-)designing degree programmes, and it has developed competences frameworks both at sectoral and subject area level, against which individual degree programmes can be referenced.

It has been demonstrated in this paper that not only these Tuning competences frameworks at subject area level (meta-profiles) but also overarching competences frameworks, such as the Tuning SQFs and the European and National Qualifications Frameworks, are essential instruments for reforming the higher education sector and its degree programmes. In particular the Tuning meta-profiles allow for comparison of (the level of) learning. Therefore, they form the backbone of any instrument for measuring learning at sectoral / subject area level in a comparative setting, locally, nationally, regionally and globally. Level descriptors and indicators as developed by Tuning allow for fair measurement of performance and comparison of learning.

Why is it important to compare learning achievements of programmes and systems with another? It will first and foremost show us their relative strengths and weaknesses. This is important information in today's world, which requires transparency and accountability. It will offer the precise data to enhance degree programmes in a sophisticated way, by allowing us to focus on its weaker elements. When the performance indicators are linked - as part of the assessment instruments - to civil and employment needs and requirements, it will also make it possible to tailor the programmes better to those. However, it has to be stressed that according to Tuning, higher education institutions and their academics should remain responsible for content and form of their higher education programmes. Nevertheless, serving society best is a role of academics for which they can and should take full responsibility as the Tuning experience has shown us. The new approach as explained in this paper, will allow academia to show what high quality programmes are. It also allows to prove that it is doing its job well.

International Symposium on Education Reform 2013
Tokyo, December 10, 2013

**TUNING Competences Frameworks:
Key Tools for Educational Reform and
the Assessment and Comparison of
Learning in Global Perspective**

Robert Wagenaar




rijksuniversiteit groningen

1

TUNING Competences Frameworks

Content presentation

1. Global challenges
2. Searching for a new paradigm
3. TUNING competences frameworks
4. Overarching competences frameworks
5. Bridging meta-competences frameworks and Tuning subject area based competences frameworks
6. Assessment of learning: measuring performance
7. Conclusions



2

1. Global challenges

- Globalization
- Information and Communication Technology
- Competitive Internationalized Higher Education market
 - Academic staff / Researchers
 - Bachelor, Master and PhD students
 - Language of instruction
- Opportunities on the labor market
 - Role of economic crises

3

2. Searching for a new paradigm

Finding responses to the challenges (1):

- Higher Education structures and programmes and qualifications should be reformed at a large scale to be able to respond to the needs of society;
- Academics should be given a key role in this process;
- The reform process should require the development of international shared reference points / standards at disciplinary / subject area level;
- A language for communication should be developed which would be understood by all major stakeholders, that is academics, students, graduates, (potential) employers of graduates as well as professional organisations;

4

Searching for a new paradigm

Finding responses to the challenges (2):

- All stakeholders, including (potential) employers and professional organisations, but in particular graduates should be (indirectly) involved in the process of curriculum design and enhancement;
- The focus should be on diversification of degree programmes by profiling and stimulating flexibility;
- The reform should facilitate national and international mobility and the recognition of periods of study, including qualifications for obtaining access to the next level of programmes

5

Searching for a new paradigm


Agreed common language: "Competences" and "Learning outcomes"

Tuning definition of competences

- Competences represent a dynamic combination of cognitive and metacognitive skills, knowledge and understanding / insight, interpersonal, intellectual and practical skills and ethical values.
- Fostering competences is the object of an educational programmes.
- Competences are formed in various course units and assessed at different stages.

[competences are obtained by the student]

6

Searching for a new paradigm 


What is a Learning Outcome according to Tuning?

Level of competence is expressed in terms of Learning outcomes (required or expected achievement):

- Statements of what a learner is expected to know, understand and be able to demonstrate after completion of learning.
- They can refer to a single course unit or module or else to a period of studies, for example, a first, a second or third cycle programme.
- Learning outcomes specify the requirements for award of credit.

[learning outcomes are formulated by academic staff]


7

Searching for a new paradigm 

**Reforming Higher Education
PARADIGM SHIFT REQUIRED
From Staff Centred to Student Centred Learning !**

Student centred learning:
An approach or system that supports the design of learning programmes which focus on learners' achievements, accommodate different learners' priorities and are consistent with reasonable students' workload (i.e workload that is feasible within the duration of the learning programme). It's accommodates for learners' greater involvement in the choice of content, mode, pace and place of learning.

8

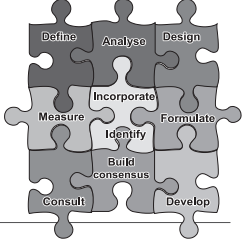
3. TUNING competences frameworks 

Developing *References Points for the Design and Delivery of Degree Programmes in*


- >Standard setting
- >Non-prescriptive

Process:

- Establish group of 12-15 international experts
- Describe Academic Field
- Define list of 30 Generic Competences
- Define list of Subject Specific Competences
- Map typical degrees
- Map potential employability field



9


TUNING competences frameworks 

The procedure:


- Defining Competences Statements
- Distinguishing Generic and Subject Specific Competences
- Consultation of stakeholders: academic staff, graduates, employers (and students)

↓

Outcomes surveys input for defining Tuning Competences Frameworks



10

TUNING competences frameworks 

Consultation based on different variables:


- the degree of importance: the relevance of the competence, in the opinion of the stakeholder (for work in their profession);
- the level of achievement: the achievement of this competence as a result of having taken this university degree.

To evaluate these two variables, the respondents had to use a scale: 1 = none; 2 = weak; 3 = moderate; 4 = strong.

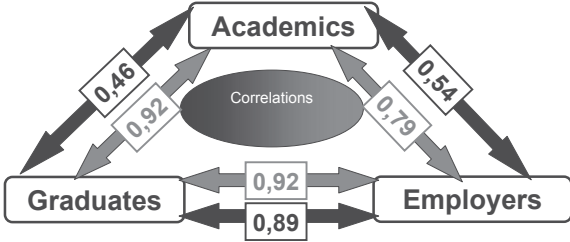
and second:

- ranking of generic competences: based on the categorisation of the five most important ones according to academics, graduates, students and employers.

11

TUNING competences frameworks 

EUROPE 2001 **EUROPE 2008**



12

TUNING competences frameworks

EUROPE 2001 EUROPE 2008

Academics, Top 6

1 Basic general knowledge	1 Ability for abstract thinking, analysis and synthesis
2 Cap. for analysis and synthesis	2 Ability to apply knowledge in practical situations
3 Capacity to learn	3 Knowledge and understanding of the subject area and understanding of the profession
4 Creativity	4 Ability to identify, pose and resolve problems
5 Cap. for appl. knowledge in practice	5 Capacity to learn and stay up-to-date with learning
6 Critical and self critical abilities	6 Capacity to generate new ideas (creativity)

13

TUNING competences frameworks

EUROPE 2001 EUROPE 2008

Employers, Top 6

1 Capacity to learn	1 Abil. to appl. knowledge in pract.
2 Cap. for appl. knowledge in practice	2 Ability for abstract thinking, analysis and synthesis
3 Cap. for analysis and synthesis	3 Ability to identify, pose and resolve problems
4 Problem solving	4 Knowledge and understanding of the subject area and understanding of the profession
5 Concern for quality	5 Ability to work in a team
6 Teamwork	6 Ability to plan and manage time

14

TUNING competences frameworks

EUROPE AFRICA AMERICA LATINA

ALL GROUPS: Ranking, Top 5

2 competences in common:

- Ability to translate knowledge into practice
- Ability for conceptual thinking, analysis and synthesis

15

TUNING competences frameworks

From Reference Points to Meta-Profiles

- >Reference points approach: identifying core or key competence of academic field
- >Meta-Profile approach: Clustering of generic and subject specific competences to derive to meta-competences

Two approaches for grouping:

- >Clustering followed by labeling
- >Labeling followed by clustering
- >Meta-competences are a mix of generic and subject specific competences
- >Number of clusters: 5 minimum to 8 maximum

16

5. Overarching competences frameworks

Complementary instruments:

- Meta-Qualifications Frameworks (EQF / QF for EHEA)
- National Qualifications Frameworks
- TUNING Sectoral Qualifications Frameworks: Social Sciences, Humanities, Creative and Performing Arts
- TUNING subject area meta-profiles or reference points for large range of disciplines
- Diploma Supplement (as an instrument to publish content and outcomes of degree programmes)

Under construction:

- TUNING Sectoral Qualifications Frameworks for other domains/sectors

17

Overarching competences frameworks

Word wide perspective: Tuning and Qualifications Frameworks

International environment

18

Overarching competences frameworks

European Qualifications Framework for Lifelong Learning: Level 7 (Master)

Knowledge	- Highly specialised knowledge, some of which is at the forefront of knowledge in a field of work or study, as the basis for original thinking and/or research - critical awareness of knowledge issues in a field and at the inter-face between different fields
Skills	- Specialised problem-solving skills required in research and/or innovation in order to develop new knowledge and procedures and to integrate knowledge from different fields
(Wider) Competences	- Manage and transform work or study contexts that are complex, unpredictable and require new strategic approaches - Take responsibility for contributing to professional knowledge and practice and/or for reviewing the strategic performance of teams

19

6. Bridging frameworks

THE COMPETENCES FRAMEWORK PYRAMIDE

20

Bridging frameworks

21

Bridging frameworks

Structuring Competences and Learning Outcomes by using Dimensions

Examples of general categories / dimensions:

Qualifications Framework for the EHEA	Lumina USA Degree Profile
Acquiring knowledge and understanding	Broad, Integrative Knowledge
Applying knowledge and understanding	Specialised Knowledge
Making informed judgments and choices	Intellectual Skills
Communicating knowledge and understanding	Applied Learning
Capacities to continue learning	Civic Learning
Contributing to original research (doctorate only)	

22

Bridging frameworks

Examples of sectoral / subject area dimensions

Humanities Dimensions	Creative and Performing Disciplines Dimensions	Engineering Dimensions
The Human Being	Making, Performing, Designing, Conceptualising	
Cultures and Societies	Re-thinking, Considering and interpreting the Human	Basic and Engineering Sciences
Texts and Contexts	Experimenting, innovating & Researching	
Theories and Concepts	Theories, Histories and Cultures	Engineering Analysis
Interdisciplinarity	Technical, environmental and Contextual issues	Engineering Design
Communication	Communication, Collaboration & Interdisciplinarity	Generic Skills
Initiative and Creativity	Initiative & Enterprise	Engineering Practice
Professional Development	Professional Development	Professional Development

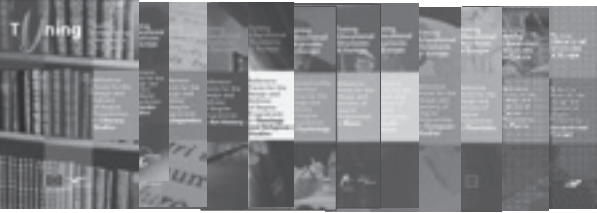
23

Bridging frameworks

Sector at level 6 EQF	Knowledge	Skills	Competences
EQF	Advanced knowledge of a field of work or study, involving a critical understanding of theories and principles	Advanced skills, demonstrating mastery and innovation, required to solve complex and unpredictable problems in a specialised field of work or study	<ul style="list-style-type: none"> Manage complex technical or professional activities or projects, taking responsibility for decision-making in unpredictable work or study contexts Take responsibility for managing professional development of individuals and groups
CPD: Making, Performing, Designing, Conceptualising	Have advanced knowledge of the processes and concepts underlying creation and/or performance in their specific discipline	Have the advanced skills necessary to create, realise and express their own creative concepts	Be able to draw upon the knowledge and skills gained within their studies to act and respond creatively in different situations
CPD: Rethinking, Considering and Interpreting the Human	Appreciate how the practice and/or creation generated within their discipline both stems from, and shapes, our humanity	Demonstrate interpretative skill and a reflection of the human dimension in their creative practice	Be able to draw upon experience gained within their studies to operate with an ethical awareness and to encourage the development and foster the well-being of other individuals and groups

24

Tuning Subject Area Meta-Profile publications ...



Other subject area brochures: Business Administration, Educational Sciences, Gender Studies, History, Mathematics, Nursing, etc.

Recent publications: Art History, Linguistics, Literature and Culture, Theology and Religious Studies.

Also published: Tuning AHELO conceptual frameworks for Economics and Engineering (first cycle)

25

Bridging frameworks

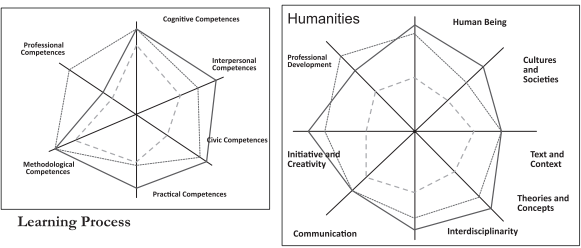
Degree programme profiles:

- > Own features
- > In accordance with Overarching / Meta-framework
- > Based on Sectoral Qualifications Framework (if available)
- > Based on Meta-profile / Reference points subject area / discipline

26

Bridging frameworks

Profiles can be based on two types of dimensions: the Learning Process and the Subject Area. Examples:



Legend: Bachelor (broken line), Master (dotted line), Doctorate (solid line)

27

Bridging frameworks

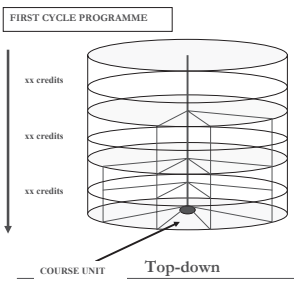
Key elements academic-professional profile

- IDENTITY** → gathers the essence of what is - "should be" - the degree holder.
- FUNCTIONS** → detects the occupations and tasks which can be carried out by the graduate.
- CONTEXTS** → focuses on the environment in which the graduate is able to function successfully.
- EDUCATION** → defines the main expected learning outcomes in terms of competences –generic and specific.

28

Bridging frameworks

Consequences for degree programme design and enhancement:



Degree programme based on the Tuning methodology:

- Programme based on profile, sets of competences to be obtained, desired learning outcomes to be achieved, student workload based credits to be awarded
- Programme design is team work, based on consultation, discussion, cooperation
- Learning outcomes / competences to be developed are basis for credit allocation
- Teaching, learning and assessment approaches respect credit allocation: feasibility key factor

29

6. Assessment of learning: measuring performance

Comparing levels of achievement of learning:

- ▣ Locally: in and between classes in same institution
- ▣ Nationally: between institutions / external reviews
- ▣ Regionally: Limited number of countries, e.g. EU, East-Asia, Latin America, North America
- ▣ Globally, e.g. AHELO

Outcome: Reliable way to compare effectiveness of learning strategies and approaches in relations to level of outcomes (taking into account mission and profile of one own's degree programme)

30



31

measuring performance
Tuning

MEASURING ACHIEVEMENT OF COMPETENCES

METHODOLOGY:

- Definition of the **competence**.
- Indication of the **levels of achievement** (mastery) for that competence (General Learning Outcome).
- For each level of achievement, determination of **indicators** (progression route) (Subdivided Learning Outcome).
- For each indicator, in a **matrix**, statement of the level attained from 1 to 5 phrased as descriptors or assessment criteria.

An example of a generic competence : *Critical Thinking*

32

Measuring performance
Tuning

Levels of Achievement : CRITICAL THINKING

- **FIRST LEVEL OF ACHIEVEMENT:** Asking oneself questions about surrounding life and actively participating in discussions on it, analysing the judgements made and reflecting on the consequences of one's own and others' decisions
- **SECOND LEVEL OF ACHIEVEMENT:** Analysing the logic of own and others' judgements, weighing their personal and social implications
- **THIRD LEVEL OF ACHIEVEMENT:** Arguing the pertinence of judgements made and analysing the consistency of own behaviour, based on underlying principles and values

33

LEVELS OF MASTERY	INDICATORS	DESCRIPTORS				
		1	2	3	4	5
First level of mastery: Ability to asking oneself questions about surrounding life and actively participating in discussions on it, analysing the judgements made and reflecting on the consequences of one's own and others' decisions	Showing critical spirit	Never questions the situation or conditions of own life.	Questions certain situations in own life.	Shows critical attitude towards conditions of own life.	Questions and explores reality, reflecting on life.	Formulates own judgements and evaluations based on systematic reflection on reality.
	Distinguishing fact from opinion, interpretations, evaluations, etc. in others' argumentation	Accepts as own judgements or decisions based on opinions, evaluations, etc. as though they were objective facts.	Accepts without question judgements or decisions based on opinions, evaluations, etc. as though they were objective facts.	Questions judgements or decisions based on opinions, evaluations, etc.	Distinguishes objective facts from opinions and evaluations.	Correctly analyses judgements or decisions based on opinions, evaluations, etc.
	Actively participating in discussion	Remains passive during discussions.	Finds it hard to participate in discussion situations.	Actively participates in discussion.	Participates constructively in discussions, contributing to construction of rich, shared reflection.	In discussions serves as constructive point of reference for others.
	Foresseeing the practical implications of decisions and approaches	Unaware of the effects of decisions and proposals.	Ignores the practical implications of decisions and proposals.	Foressees the practical implications of decisions and proposals.	Analyses the pros and cons of the effects of decisions proposed.	Gives importance to proper evaluation of the pros and cons of decisions and proposals.
Reflecting on the consequences and effects that one's decisions have on others	Doesn't think about the consequences of own actions.	Simply accepts others' observations and criticisms of own behaviour.	Reflects on the consequences and effects that own decisions have on others.	Recognises and accepts own mistakes.	Asks for, weighs and takes into account others' feedback on own conduct.	

34

Measuring performance
Tuning

Same approach can be applied for subject area / disciplinary competences:

- > Define subject area competence
- > Indicated levels of mastery (achievement) for succeeding course units (Define progression route or pathway) (General Learning Outcome)
- > Determine the indicators for achievement of (each) level defined (Subdivided Learning Outcome)
- > Define descriptors / assessment criteria

Outcome: framework to measure disciplinary based competences / Learning Outcomes



An example of a subject specific competence in the field of Architecture: *Conceptualising, Designing, Materialising*

35

Measuring performance
Tuning

ARCHITECTURE			
EQF Categories:	Knowledge	Skills	(Wider) Competences
Architectural Creation & Architectural Creativity			
Graduates of Architecture are expected at level			
Level/ Dimension			
6 / Conceptualising, Designing, Materialising Architecture	To have advanced knowledge of the processes, concepts and cultural values guiding architectural creation	To have the advanced skills to formulate critically, elaborate creatively and translate innovatively into spatial forms their own architectural concepts	To be able to draw upon the knowledge and skills gained within their architectural studies so as to act and respond critically and creatively in different situations
7 / Conceptualising, Designing, Materialising Architecture	To have highly advanced knowledge of the processes, concepts and cultural values guiding architectural creation, some of which will be at the forefront of their field	To have developed to a high professional level their ability as architects to formulate critically, elaborate creatively and translate innovatively into spatial forms their own architectural concepts	To emerge as well-developed personalities, able to draw upon the knowledge and skills gained within their architectural studies so as to act and respond critically and creatively in situations that are complex, unpredictable and require new strategic approaches
8 / Conceptualising, Designing, Materialising Architecture	Knowing in depth all the relevant methods and techniques of inquiry related to a particular field of study of architecture	Integrating previous experience so as to demonstrate original creative insights in the domain of architecture Functioning with complete creative autonomy	Comprehending the transferability of their research capabilities to other fields Displaying professional, creative and scholarly integrity

36

 7. Conclusions 

- Overarching competences frameworks and Tuning meta-profiles are essential instruments for reforming the higher education sector and its degree programmes.
- Tuning meta-profiles allow for comparison of (the level of) learning.
- Tuning meta-profiles form the backbone of any instrument for measuring learning at sectoral / subject area level in a (trans)national setting.
- Level descriptors and indicators allow for fair measuring of performance and comparison of learning.
- Degree programmes require a facelift to meet the challenges of today's and tomorrow's world.

37

Thank you for your attention !



<http://www.unideusto.org/tuningeu/>

38

Presentation 2

The AHELO Feasibility Study: Study Results and the Conclusions of the Technical Advisory Group (TAG)

Vice President, National Center for Higher Education
Management Systems

Peter T. Ewell



The purpose of this brief paper is to describe the design and results of a Feasibility Study of the Assessing Higher Education Learning Outcomes (AHELO) initiative conducted by the OECD in 2009-2012, and to draw some conclusions about what was learned through the Feasibility Study from the perspective of its Technical Advisory Group (TAG). The TAG was established early in the Study to provide technical guidance to the AHELO Feasibility Study and, based on this experience, to draw conclusions on the conduct of a possible Main Study.¹

Background and Early Development of AHELO.

OECD's Directorate for Education began planning for an international assessment of learning in higher education in the spring of 2007. An initial "experts meeting" was held in Washington, DC in April of that year and involved a group of invited assessment and policy representatives from a half dozen countries. While discussions at that meeting remained general, it was agreed that such an assessment would have value for participating countries and that OECD should design and implement an exploratory study in a few countries in order to identify the challenges involved. This was followed by larger experts meetings in Paris in early July and in Korea in late October.

By the end of the third meeting, the basic shape and timetable of the AHELO initiative had emerged. The assessment would involve several subject matter domains producing results at the individual institutional level for benchmarking purposes. Like the PISA assessments already offered by OECD, they would prominently feature task-like production measures. Two components of the assessment would examine outcomes in selected fields of study with international relevance. After some discussion, Economics and Engineering were chosen as the primary subjects of interest. Another component of the proposed study would be a direct examination of "generic skills" like critical thinking and problem solving that might apply to all fields. One reason advanced for the latter was that "valid and reliable" instruments were already available in this area such as the U.S. Collegiate Learning Assessment (CLA). The experts recognized that such generic skills are widely viewed as critical for individual success, but that any assessment of them should properly be set in the context of particular fields of study. It was also recognized that many fields like history, literature, and law, are so culturally conditioned that international benchmarking would make little sense. Finally, recognizing the vast differences in the contexts for higher education across countries including curricular structure, educational values, and predominant pedagogies, a final important component would be a "Contextual Dimension." This would consist of background surveys completed by faculty and institutional administrators, as well as a student survey to be administered at the same time as each cognitive assessment.

¹This paper's author served as the Chair of the TAG and its contents draw heavily on reports of the TAG issued throughout the Feasibility Study's history.

Before establishing the detailed design of the Feasibility Study, additional decisions needed to be made about AHELO's purpose and how it should be governed. With respect to purpose, OECD argued that the need to align higher education outcomes in key areas across boundaries in a time of growing graduate mobility was becoming imperative. This is a major objective of the Bologna Process in Europe, and is also reflected in initiatives like the "Tuning Project," which is trying to coordinate academic standards across institutions in different countries in multiple subjects. Reasons for individual institutions to participate reflect the same motivations as institutions administering national standardized assessments like the CLA: such information can be useful for strategic planning and supporting external quality assurance reviews, as well as for benchmarking performance on locally developed measures of student achievement.

With respect to governance, AHELO is located within OECD in the Institute for Management in Higher Education (IMHE), a unit primarily concerned with providing services to institutions in common areas like strategic planning and assessment throughout the world. Policy within this unit is set by the Education Policy Committee (EdPC), which with the OECD Secretariat, assumed broad responsibility for the AHELO initiative. To provide more detailed guidance to the initiative, a Group of National Experts (GNE) was created consisting of representatives of participating countries and jurisdictions, as well as "observer" countries that expressed interest in the initiative but did not participate in the Feasibility Study. Members of the GNE are a mix of policy and ministry representatives and individuals with expertise in international testing.

The AHELO Feasibility Study was originally envisioned to be a two to three year effort involving a small group of countries. The primary questions that this effort was designed to answer were: a) could valid and reliable assessment instruments be developed in multiple languages across quite different country contexts; b) could such an effort be managed politically through OECD's complex multinational governance structure; c) could the testing process be effectively implemented through the use of country and institutional coordinators, assessment administration manuals, and local training programs and; d) would the resulting data be able to be analyzed and prove useful to participating countries and institutions?

The Design of the AHELO Feasibility Study.

Each country participating in the Feasibility Study was asked to select a set of six to ten institutions chosen to reflect the diversity of that country's higher education system. The sampling approach for each institution then involved choosing 200 students who were nearing the end of their three or four year period of tertiary study. Because the principal purpose of the Feasibility Study was to try out an assessment approach and draw conclusions about implementation, the resulting data were never intended to support comparisons across institutions or countries. To achieve maximum efficiency, all assessments were "spiraled" so that different students completed different parts of the whole. All students participating in AHELO completed a ninety minute to two hour assessment instrument in one of the domains of interest, plus the short survey designed to gather information about student backgrounds and educational experiences. All of the assessments were administered electronically through secure computer links. Constructed response task (CRT) items were scored by trained reviewers in each country using specially-designed rubrics.

The core of the Generic Skills component of AHELO was a modified version of the CLA, the task-based assessment administered by the New York based Council on Aid to Education (CAE). Representatives of CAE attended each of the initial planning meetings for AHELO and CAE was chosen as one of the prime contractors because its CLA assessment was at that time the only generic skills assessment that used a constructed response format. Two CLA task prompts were selected for further development after an invitational meeting in New York attended by prospective countries—"Lake to River," which deals with the pros and cons of proceeding with a dam project with uncertain safety consequences and "Catfish," which deals with ascribing the likelihood that an observed biological anomaly might be due to a commercial polluter. In addition to completing one of these tasks, each student also answered a short battery of multiple-choice questions (MCQs) on generic competencies drawn from the Graduate Skills Assessment

designed by the Australian Center for Educational Research (ACER).

The AHELO Economics assessment was developed by the Educational Testing Service (ETS) using a combination of existing multiple-choice test items drawn in Economics and various international inventories of economic content knowledge including results of the Tuning project and Subject Benchmarks created for the discipline by the Quality Assurance Agency of the United Kingdom. Construction of the assessment was guided by an assessment framework created by a group of subject matter experts. This group also reviewed the assessment itself after draft items were developed. Rather than concentrating on strict content knowledge, the focus of the assessment was on students' ability to "think above content" and use the concepts and language of the discipline effectively. The assessment itself consisted of a set of constructed-response tasks plus a multiple-choice battery drawn from re-worked items on the Economics GRE.

The Engineering assessment was developed by an international partnership of assessment development organizations led by ACER, and was guided by existing national competency frameworks, a Japanese engineering licensure examination, and the results of the Tuning project in Europe. Because Engineering consists of various sub-fields, a decision had first to be made about which should be examined. The panel of subject-matter experts chose Civil Engineering, largely because it is the most conceptually straightforward. In addition, a number of more general skills in Engineering shared by all sub-fields such as analysis and design, as well as basic scientific concepts underlying all of them, were assessed. Construction of the assessment was also guided by an assessment framework, constructed by the subject matter experts. Like the Economics assessment, a focus of the Engineering assessment goes beyond mastery of content to examine students' ability to "think like an engineer." The design of the assessment itself was influenced heavily by the Engineering licensing examination in Japan. Like the assessment in Economics, the Engineering assessment consisted of a mix of multiple-choice items and constructed response tasks.

Although the original plan for the Feasibility Study called for only five or six countries, a total of seventeen countries eventually participated in twenty-five fields—nine in Generic Skills, seven in Economics, and nine in Engineering. All countries participated in contextual data collection, which involved the ten minute student survey administered in conjunction with the assessment, a survey of faculty, a survey of department chairs (in Economics and Engineering) and a survey of institutional administrators. Additional contextual data included a range of descriptive materials about curriculum and each country's higher education system assembled by a study coordinator in each country. Testing began in the spring of 2012 and was concluded at the end of May, with scoring and analysis completed by the end of the summer of 2012.

Creation and Role of the TAG.

The need for an advisory body responsible for reviewing and upholding technical standards for the AHELO Feasibility Study was recognized by the OECD Secretariat and interested countries from the outset of the AHELO initiative. This need was affirmed by the three expert group meetings held in 2007 in Washington, Paris, and Seoul.

The TAG was formally established in 2010 with eight members drawn from assessment and policy experts from throughout the world. The TAG reported to both the OECD Secretariat and the AHELO GNE. The Terms of Reference of the TAG were specified broadly, but essentially established a role that consisted of a) reviewing draft materials on all aspects of the Feasibility Study and suggesting mid-course corrections, b) providing recommendations on the eventual conduct of an AHELO Main Study and, c) providing a definitive recommendation on the feasibility of AHELO at the conclusion of the study. A fourth responsibility was added in Phase II of the Feasibility Study when the TAG was charged with serving as the expert group for the Generic Skills strand and the Context Dimension. Finally, the Terms of Reference established that the GNE could call on the TAG for advice on technical "or other matters"—a charge that allowed the TAG to consider policy and implementation questions with increasing frequency as the Feasibility Study progressed.

The TAG met eight times in the course of the Feasibility Study, three of which were face-to-face meetings and

the balance conducted via teleconference. Most meetings consisted of updates on progress guided by a review of documents and covered all facets of the study including the development of assessment frameworks, instrument development, sampling approaches, country coordination, assessment administration procedures, scoring arrangements for CRTs, analysis plans, and reporting arrangements. Recommendations for mid-course guidance of the Feasibility Study were developed by the TAG in the course of these reviews. After each meeting, the Chair of the TAG drafted a report, which was then forwarded to the GNE and the Secretariat. The Chair of the TAG also met with the GNE after each face-to-face meeting to report on issues and recommendations, and the Chair of the GNE also observed the final face-to-face meeting of the TAG in October of 2012.

The TAG's Overall Assessment of the Feasibility Study. The AHELO Feasibility Study constituted an unprecedented multi-national data collection effort at the higher education level. Data on student learning outcomes were collected in three domain strands in seventeen different countries or systems, using assessment instruments comprising both production-focused CRTs and forced-choice MCQs. Data were collected on a wide range of contextual factors by means of surveys completed by students, faculty members, institutional coordinators and national project managers. Numerous implementation challenges including translation, contextualization, sampling, electronic test administration, CRT response scoring, data cleaning, statistical analysis, and reporting were, for the most part, met and successfully overcome. To be sure, some countries/systems experienced more difficulty than others and, because of this, levels of success varied. Nevertheless, all participating countries reported they learned something from the experience and would do it again. Just as important, the Feasibility Study generated a range of important findings about student learning at the higher education level, as well as dozens of lessons about how a Main Study should be implemented.

That said, some things went particularly well in the AHELO Feasibility Study and a few did not go so well. Most have implied lessons for any AHELO Main Study.

What Went Well. The following were particular strengths of the Feasibility Study:

- Assessment Administration. Electronic administration of assessment on a global scale, and in multiple languages and jurisdictions, presented the Feasibility Study with an enormous challenge. This challenge was met admirably. Only one significant failure in administration occurred over scores of testing sessions at hundreds of institutions. The technical infrastructure underlying this achievement, the thorough training regimens put in place for institutional coordinators, and the robust administration procedures established all contributed to success here.
- Technical Aspects of the Data Analysis. The data yield of the Feasibility Study was large and complex, resulting from the administration of six different instruments to many different kinds of respondents. In the face of this, efforts to provide sound analyses were exemplary from a technical standpoint. The analysis plans were sound, the statistical techniques employed were proper and well executed, and appropriate and effective “work-arounds” were put into place when analytical problems (such as missing data or malfunctioning items) were encountered.
- Instrument Design for Purpose-Built Instruments. All of the instruments designed especially for the Feasibility Study were of exemplary technical quality including the MCQs and CRTs for Engineering and Economics and the three surveys comprising the Contextual Dimension. All were developed through reference to adequate and helpful Assessment Frameworks and were informed by knowledgeable expert groups (in the cases of Engineering and Economics) or considerable background work (in the case of the Contextual Dimension). Moreover, these instruments were produced quickly with little re-work, were designed to a high technical standard, and were piloted as well as could be expected in the short timelines available.
- Overall Coordination. Management and coordination of an enterprise as complex as the AHELO Feasibility

Study involved massive challenges of maintaining consistent procedures across five continents, seventeen unique cultural-political contexts, and numerous time zones. The administrative arrangements that were put in place to run the Study met these challenges with clear direction and minimum confusion. Where the inevitable problems were encountered, they were for the most part resolved quickly and smoothly.

Things that Did Not Go So Well. At the same time, some aspects of the Feasibility Study did not go so well. As a consequence, they constitute areas that must be particularly examined as the initiative moves forward.

- **Resources and Time.** The AHELO Feasibility Study was seriously under-resourced and was implemented on far too short a timeline. More resources and time could have enabled such important features as more cognitive interviews and pilots of newly-build instruments, full-scale field trials of administration and scoring arrangements, and more time for de-briefing and collective discussion of obtained results.
- **Uneven Student Cooperation Rates.** The validity of any assessment depends in part on obtaining a sufficient number of students chosen as part of the sample at each institution to participate. Countries and institutions achieved substantially different levels of cooperation in the AHELO Feasibility Study. In some countries (Colombia and Mexico, for example) almost all students completed the assessments. In others (Norway and The Netherlands, for example) so few did so that obtained results were not valid enough to use. This mixed track record means that unusual attention to obtaining needed levels of student cooperation will be needed for a Main Study.
- **CRT Difficulty and Contextualization.** While the CRTs used by the Engineering and Economics assessments were of high technical quality, they were simply too difficult for many students to effectively engage and perform well. At the same time, the CRTs used in Generic Skills based on the CLA proved excessively “American” in an international context. As above, more time for piloting and field trials might have revealed both of these situations at an earlier stage—in time for them to be rectified.
- **Contractual Arrangements.** The AHELO Feasibility Study began with separate contracts between the OECD Secretariat and the two principal contractors—ACER and CAE. These independent contractual relationships resulted in poor communication among the contractors and occasional duplication of effort. Furthermore, no tendering process was used to procure or develop instruments for the Generic Skills strand. By the time this situation was addressed by re-structuring contractual arrangements so that CAE was a subcontractor of ACER under the Consortium, past history meant that it was difficult to establish a true culture of partnership.

Some Particular Lessons from the Feasibility Study. Experience with the AHELO Feasibility Study offers additional lessons that should be taken forward for the AHELO Main Study:

- *There should be more opportunities for stakeholder participation in assessment design and in the analysis of assessment results.* There were many points in the Feasibility Study at which the wisdom of practitioners and the national and institutional levels could have been better collected and used for improvement. While the many efforts to contextualize instruments and administration procedures were admirable and, for the most part, successful, a more collaborative approach might have yielded even greater benefits.
- *A full-scale try-out of all instruments and administration arrangements could enable stakeholder participation in a “design-build” process that would both pilot these designs and enable more stakeholder engagement in making them better.* This is especially the case for reporting results and sharing data with countries and institutions. Many institutional participants were somewhat disappointed by the lack of attention to their needs for information resulting from the study. Institution-level reports with more detailed breakdowns across student populations would have been beneficial, as well as more fully documented institutional and country data files. A project-wide Quality Monitor should also be established, as well as a National Quality Monitor for each participating country/system. This is consistent with international standards in conducting such studies.


- *More information should be made available about the costs and benefits to countries and institutions of participating in AHELO.* Two primary questions will probably be raised by any country/system considering whether or not to join an AHELO Main Study: “what is it likely to cost us?” and “what are we likely to learn?” Because the AHELO Feasibility Study is only just concluded, little can be said about the second question at this point. But some information about costs is available. The direct monetary costs of developing, adapting, and administering the various instruments are known through OECD contracting records. Many costs incurred by institutions and systems for such activities as sampling, student recruitment, test administration, scoring and data reporting, and coordination/oversight are similarly known. But many are not documented because they constitute less tangible costs, for example the time devoted to AHELO by institutional and system personnel. As a consequence, a systematic effort to collect data about both direct and indirect costs should be included in any future Main Study.
- *Tools such as “Readiness Criteria” should be developed and put in place to allow potential participants and the OECD to determine whether institutions and jurisdictions can actually undertake and benefit from AHELO.* An explicit set of country and institutional readiness criteria should be established to govern institutional participation in any AHELO Main Study. These criteria should include the provision of a student population sampling frame, sufficient computing infrastructure and IT personnel to support computer-based testing, commitment to participation in training, and effective internal management. It should also include a formal commitment to carry out Study protocols and to abide by the AHELO Technical Standards.
- *Further work is needed about how to effectively assess Generic Skills in multiple disciplinary and national/cultural contexts.* A major design choice for the AHELO Main Study is whether or not to include a dedicated Generic Skills strand. The existence of these competencies independent of discipline or field of study is a contested issue in the field of higher education assessment. Some generic competencies transfer relatively well across domains, other generic competencies are developed, applied, and assessed much more appropriately within the contexts of particular domains. Results of the Feasibility Study on Generic Skills CRTs suggest that these tasks might perform better if they were better contextualized. The two Generic Skills CRTs used in the Feasibility Study were contextualized to a “real world” problem-solving situation. However obtained results suggest that the manner in which these tasks were culturally situated and perceived varied substantially across countries and systems. How appropriate contextualization of Generic Skills should be accomplished in any future AHELO Main Study is still a matter for consideration. One option is to continue down the path of including “discipline-specific generic” components in each disciplinary assessment. This was done in Engineering in the Feasibility Study and, to some extent in Economics. If further development along these lines is pursued, these “discipline-specific generic” competencies should be more appropriately aligned with one another to ensure that they address some parallel content. If a decision is made to continue with a separate Generic Skills strand, the performance tasks might be situated in the context of broad disciplinary groupings like the sciences, social sciences, humanities and fine arts.
- *To provide meaningful information for improving teaching and learning, a mix of item types is required in international assessments at the higher education level.* Another design choice about instrumentation is whether production based CRTs should be included in an AHELO Main Study at all. Decades of research have shown that CRTs will never perform as well in terms of reliability as a battery based solely on MCQs. Results of the Feasibility Study confirm this conclusion for all three domains. The question for an AHELO Main Study is whether the use of CRTs adds enough validity to be worth this inevitable price in lost reliability. On this question, results of the Feasibility Study in Engineering suggest that some of the most important information that could drive improvements in teaching and learning was obtained through the CRTs. The major drawback of including CRTs is substantially increased costs. If the main purpose of AHELO is held to be instructional improvement, the inclusion of CRTs will undoubtedly increase the usefulness of results. On the other hand, if the

main purpose is to provide the most reliable international benchmarks of institutional performance with respect to student learning outcomes, the greater reliability and lower cost of adopting an approach based solely on MCQs may be preferred.

- *An acceptable response rate based on an accurate probability sample is required to assure comparability of results across institutions.* In the Feasibility Study, probability sampling or a census of students was used by almost three-quarters of participating institutions. For the remaining institutions, it is not apparent why such a sample was not used. For the Main Study, participating institutions should be required to compile a list (or lists) of eligible students (or groups of students) and to use either probability sampling or a census. That said, there should be some flexibility regarding the choice of probability sampling method. For example, cluster sampling of class groups may be reasonable when the number of eligible students is large. It may also be reasonable for AHELO to impose a fixed minimum response rate threshold, at the level of the country or the institution, below which data will be excluded from the data analysis. Finally, measures to increase response rates should be actively researched before any new AHELO data collection.
- *AHELO should be better located and integrated with the international scholarly community examining student learning outcomes and the policies and practices that support better learning.* The past decade has seen a sharp increase in policy and scholarly interest in improved academic performance in higher education. Evidence of this can be seen in the Bologna Process and Tuning in Europe, the Spellings Commission and interest in accreditation in the U.S., the rise of qualifications frameworks in many nations, and the emergence of multinational ranking initiatives like U-Map and U-MultiRank. AHELO represents an opportunity to better align the emerging scholarly and policy dialogue about quality.
- *All of this will require more time and adequate resources.* The AHELO Feasibility Study experienced serious resource shortfalls which, in the course of implementation, negatively affected many of its components. This occurred incrementally and its effects were complicated by the fact that the project included more countries than a “feasibility study” should probably have included. A similar under-resourced condition cannot be allowed for a Main Study. The OECD and participating countries will need to ensure adequate resources in moving forward. If this cannot be guaranteed, implementation will have to wait until it can.

Moving Forward?

The OECD has concluded that the results of the Feasibility Study were sufficiently positive that a Main Study should be conducted and has already distributed a paper describing its main features and inviting country participation. This verdict by no means assures that AHELO will become more broadly operational, however. The real question that will govern moving forward is not whether or not it is possible to conduct these assessments. It is instead whether or not the effort is cost effective. The cost side is by this point readily apparent. The AHELO Feasibility Study cost more than nine million Euros to implement, most of which was borne by participating jurisdictions and institutions. Benefits to participants, on the other hand, have been mixed—largely depending upon the amount of effort jurisdictions and institutions invested in undertaking local data analyses and disseminating the results. Consistent with OECD’s hopes for a Main Study, participating institutions may find internationally benchmarked assessment results helpful in their strategic planning efforts. Participating countries, meanwhile, will be able to rehearse how various kinds of results can be used to evaluate higher education performance. Whether either of these actors, as well as a host of non-participating institutions and nations, will come to believe that full implementation is worth the investment remains to be seen.



The AHELO Feasibility Study: Study Results and the Conclusions of the Technical Advisory Group (TAG)

Peter T. Ewell
National Center for Higher Education
Management Systems (NCHEMS)

NIER Conference
Tokyo, Japan
December 10, 2013

1



Background and Early Development of AHELO

- Parallel Efforts in PISA and PIAAC
 - Expert Meetings in 2007 in Washington, Paris, and Korea
 - Domains: Economics, Engineering, and Generic Skills—As Well as a Contextual Dimension
 - Decision to Launch a “Feasibility Study”
-

2



Design of the AHELO Feasibility Study

- Seventeen Countries (Jurisdictions)
 - Six to Ten Institutions per Country
 - Two Hundred Students per Institution who are Nearing the End of their Academic Programs
 - Ninety Minute to Two Hour Testing Period
 - Mix of Multiple Choice and Constructed Response
 - Test Items “Spiraled” Across Test Takers
-

3



The AHELO Generic Skills Assessment

- Administered in Nine Countries
 - Two Constructed Response Prompts Based on the Collegiate Learning Assessment (CLA)
 - Multiple-Choice Items Drawn from the Australian Graduate Skills Assessment
 - Contextualized and Translated Under Direction by Each Country
 - Scored Locally in Each Country
-

4



Examples of Generic Skills

- Cognitive:
 - Communications (Written and Oral)
 - Critical Thinking and Problem Solving
 - Non-Cognitive
 - Empathy and Ethics
 - Working with Others
 - Taking Responsibility
-

5



Graduate Attributes: Top Choices for Employers

Percent Reporting “Very” and “Somewhat” Important:

- Ethical Judgment and Integrity – 96%
 - Comfortable with Diverse Backgrounds – 98%
 - Demonstrated Capacity for Professional Development – 94%
 - Interest in Giving Back to Communities Our Company Serves – 71%
-

6

Teaching Generic Skills: Common Approaches

- A “General Education” Block of Courses
- Embed These Skills in Disciplinary Teaching
- Particular Pedagogies:
 - Collaboration and Group Work
 - Internships and Fieldwork
 - Capstone Courses and Projects
 - Role of the “Co-Curriculum”

7

The AHELO Economics Assessment

- Administered in Seven Countries
- Constructed Response Prompts Based on Tuning and UK Subject Benchmarks in Economics
- Multiple-Choice Items Based on the ETS Graduate Record Examination in Economics
- Focused on Student Ability to “Think Above Content”
- Scored Locally in Each Country

8

The AHELO Engineering Assessment

- Administered in Nine Countries
- Confined to Civil Engineering
- Constructed Response Prompts Based on Tuning
- Multiple-Choice Items Based on a Japanese Engineering Licensure Examination
- Focused on Student Ability to “Think Like an Engineer”
- Scored Locally in Each Country

9

The AHELO Contextual Dimension

- Designed to Gather Background Information in Terms of Which to Explain (and Adjust) Cross-Country Variation
- Elements of the Contextual Dimension
 - Student (Test Taker) Survey
 - Faculty Survey
 - Institutional Background Survey
 - Country Background Information

10

Some Findings

- All Assessments Performed at or Above Expected Levels of Reliability
- Constructed Response Items Less Reliable than Multiple-Choice Items [as Expected]
- Constructed Response Items Clearly Tapped an Independent Dimension of Student Knowledge
- Some Evidence of Convergent Validity for All Assessments

11

The Technical Advisory Group (TAG)

- Charged with Reviewing the Technical Adequacy of All Aspects of the AHELO Feasibility Study:
 - Assessment Frameworks and Instruments
 - Sampling and Test Administration Procedures
 - Analysis and Reporting
- Charged with Making Final Recommendations on Feasibility and the Future Conduct of the AHELO Main Study

12

The TAG's View: What Went Well

- Assessment Administration
- Technical Aspects of the Data Analysis
- Instrument Design for Purpose-Built Instruments
- Overall Coordination

13

The TAG's View: What Did Not Go So Well

- Resources and Time
- Uneven Student Cooperation Rates
- Contextualization and Difficulty of Constructed Response Tasks
- Contractual Arrangements

14

Lessons for the AHELO Main Study

- More Opportunities for Stakeholder Participation in Assessment Design and Analysis
- Full-Scale Try-Out of Instruments and Administration Procedures
- More Up-Front Information on Costs and Benefits
- Development of Country/Institution "Readiness Criteria" to Assist Decisions to Participate
- More Work on the Assessment of Generic Skills

15

Lessons for the AHELO Main Study (Continued)

- Continue Mix of Item Types (Multiple-Choice and Constructed Response)
- Use Census or Probability Sample for All Institutions
- Better Integrate AHELO into International Discussions of Quality in the Scholarly Community
- All Will Require More Time and Adequate Resources

16

Moving Forward?

- OECD Has Concluded that AHELO is Feasible and is Planning to Move Forward
- Country and Institutional Decisions About Participating will Depend on Individual Decisions About Costs and Benefits
- Cost Side Known, Benefits Not Yet Assured
- We Will Know within a Year Whether and How AHELO Will Move Forward

17

Presentation 3

AHELO in Japan and Its Prospects

Dean, Graduate School of Engineering, School of
Engineering, Tokyo Institute of Technology

Kikuo Kishimoto



This is the last presentation before the break. My name is Kishimoto. I'm from the Tokyo Institute of Technology. I will talk about AHELO and its prospects in Japan. As has been discussed, our country has participated in the engineering field, and we have conducted various activities in regard to that. First, I will give you an overview of what we have been doing in the engineering arena. Second, I will give you an idea of the kinds of questions that were created for the examination. And third, I will talk about the significance and prospects of our participation.

First, here is the outline of the initiative. As Dr. Ewell mentioned, its purpose was to measure university education learning outcomes worldwide through a common test. The final report is being translated, and is expected to be published in March 2014; the original version can be downloaded from the site that you see on the slide. Regarding what is being measured, from the perspective of Tuning AHELO, the framework was created and broken down into the questions on the test. A similar methodology was applied to both economics and engineering. This is the structure of Japan, this is AHELO and the national center, and validation was done. It was conducted under the framework of the university reform initiative of MEXT, and my university, Tokyo Institute of Technology, served as its representative. The AHELO consortium was commissioned by the OECD to carry out this project, and we utilized the test -- We asked the students to solve the problems on the test and analyzed the results. The AHELO consortium created the actual test. Therefore, we did not merely participate in AHELO, but have worked on the creation of the test and were thus able to get an overview of the initiative. The AHELO initiative had two phases: creating the test and conducting it in a small scale. The areas of improvement and scoring criteria for the test were selected. In the second phase, its scientific feasibility and practicality were evaluated, a final meeting was held in March 2013, and the report was produced.

We now go to the engineering arena. What was the framework for the creation of the tests? On May 4 and 5, 2009, experts met in Brussels to discuss the topic. As you can see, the agenda included nine items. The overall discussion concerned the engineering domains. What kind of specialized academic domains are there? What are the major degrees and occupations? The report was created under the Tuning framework. Agenda item number 5, the definition of expected and desired learning outcomes, was very important. What, then, is a learning outcome? As Prof. Wagenaar has said, this is the definition of the learning outcome; specifically, how do we select it? This was the challenge that was given to us. We decided to select the learning outcomes as follows. We decided to use ABET (Accreditation Board for Engineering and Technology) engineering criteria 2000. These are the criteria used for accreditation in engineering areas, so we decided to utilize the framework there. This is the equivalent of the Washington accord criteria; in Japan, it's the Japanese criteria -- JABEE: Japan Accreditation Board for Engineering Education. In Europe, EUR-ACE was utilized. This is also an accreditation organization and they have learning outcomes. We referred to the outcomes of the first cycle of bachelor degrees.

Because this was a feasibility study, everything was considered equal, and with the framework, we made some considerations. The Tuning-AHELO conceptual framework was thus derived; this chart shows the conceptual framework of the EUR-ACE used in Europe on the left-hand side. As you can see, the standards are categorized into knowledge and understanding, engineering analysis, engineering design, investigations, engineering practice, and transferable skills. As for ABET, if the criterion deals with the ability to apply the knowledge of mathematics, science, and engineering, it was categorized as a) to correspond to the categories of EUR-ACE. We decided to select the common parts, and a Tuning-AHELO framework was set up, as you can see on the right-hand side. After identifying the common areas, we decided to pick several of them in creating the scheme. Our selection included the basic engineering sciences, engineering practice, engineering design, and engineering analysis. These are engineering skills, and they were included in the Tuning-AHELO framework of learning outcomes.

The consortium's work was to create a test from the framework. This is how we did it. The generic skills in engineering are required across the board, and above them are the engineering fundamentals. In civil engineering, we put in the specific skills needed, and above that, engineering analysis, engineering design, and engineering practice were put in place.

For this area, the test we created was composed of multiple-choice questions to examine the level of basic knowledge and skills. On the other hand, for the areas related to activity, constructed response questions were created, as mentioned in the earlier presentation. What I mean is that the students would select 25 out of 30 items and solve one of the three items for the constructed response tasks. I would like to give you an idea of the specific questions given. We developed the questions based on the conceptual framework, and the consortium prepared the draft. As stated in the previous presentation, the multiple-choice questions were proposed by Japan. We had with us the provisions of the first professional engineering examination given by the Institution of Professional Engineers of Japan, and we translated it. Also, within the Japan Society of Civil Engineers, there is a civil engineering licensing examination to test the ability of those who graduated from the civil engineering department. We translated that as well and made a proposal to the consortium. We extracted the questions which were appropriate for AHELO. On the other hand, the constructed response tasks were proposed by Australia, and the appropriate questions were selected.

At the meetings of the international experts, the proposal from the consortium was submitted. They chose from the proposed questions and reviewed the selection to raise them to the internationally accepted level. The Japan-specific questions drawn from the professional engineering examination and civil engineering licensing examination were excluded from the common test, as these would not be appropriate for such. The areas that are not taught so much in other countries were also excluded. For example, some countries said they did not need countermeasures for earthquakes because they did not have any earthquakes. If the meetings had been scheduled after the Great East Japan earthquake, we could have made a strong proposal. At that time, however, countermeasures for earthquakes were considered irrelevant and, therefore, excluded. Also, there were countries that did not have railways, so the railway questions were considered inappropriate for them. It was very difficult to select the appropriate questions for the test. To illustrate specific questions, the multiple-choice questions look like this. This is called a truss; when there is a framework and a certain load is put on it, what kind of power would have to be applied? The students would pick the answer from choices A, B, C or D. The professional engineering examination has five choices, but we decided to cut the number of choices to four, even though it would alter the degree of difficulty.

We have a lot of experience in the creation of these types of questions. Through these questions, we are able to determine whether the basics have been acquired by the students. This is an example of the CRT (constructed response task) questions. On the top part is the address of the website; if you download that, you'll be able to read the original text. There's Hoover Dam, and this question is related to it. We posted a photograph of the proposed site of the dam. A sketch or plans of the proposed reservoir were also provided. This is the set of questions. The questions were not limited to four; there were more.

The first question is "Explain why this is a good dam site for hydroelectric power generation," and it asks you to

discuss at least two aspects. There are various answers to this, and the person who will score the test will check whether the answer is appropriate, half-scored, or unacceptable. The second question is “Explain the two main design features that contribute to the structural strength and stability of the Hoover dam.” There are various other questions; for example, “In the midst of constructing this dam, you found a defect. How would you respond to that?” These types of questions help a future professional engineer determine whether he or she has learned enough to be able to respond to such problems appropriately. The details of each question revealed the learning outcomes that we wanted to assess. For the first example, it was the features that were suitable for a dam that would be used to generate hydroelectric power. To come up with the appropriate answer, they will need to have sufficient professional knowledge and the necessary competencies. The aspects range from A to F; it’s not just one. Therefore, if they write one aspect, that’s one point; two aspects, two points; and so on. The learning outcome assessed by the second question was the ability to determine which features contributed to the structural strength and stability of Hoover Dam. Two of several possible answers are “arch shape” and “robustness of the material in the canyon walls.” I forgot to mention this earlier in the first question; here it is written that the dam location will have minimal social impact. Well, you might answer that there are no residents nearby, but such a response was not initially included in the model answer. Many of the Japanese students wrote similar answers to this, so we incorporated that as a correct answer, too. We needed to discuss all the areas—the kinds of answers or what the students will think of when answering these questions. Next, we asked the students what they thought about the multiple-choice questions. They said most of the questions were about what they had learned at the university, and that it was not difficult to solve them. They also said the questions were one-shot; they appeared to be wide and shallow, and could use more depth.

But does this reflect the actual score outcome. I think the impressions of the students and the actual results were not the same. Regarding the constructed response test, in the first phase, we had the students write their answers on paper, and they wrote a lot of sentences. In the second phase, we used the computer, and the answers were very short and abrupt. So, the content of the response was different when they used a pencil or a computer, and I think that’s another issue that we need to tackle. But first of all, the Japanese students said that they found it good that the questions in the first phase of the constructed response test concerned a practical issue, as they often deal with theoretical questions at the university. They also said that they enjoyed the ethical questions, too. Other comments said they found it interesting that it was a well-thought-out question that “made us think of the cause and then respond to it. I think it will be useful in practice in the future.”

Other comments were that perspectives such as those adapted to CRT-type questions were important. One student said, “I feel we should have more group discussions and case studies in our classes, meaning that we haven’t had the opportunity to learn them before, so we would like to learn these approaches to problem solving better.” The result of the feasibility study is one thing, but that the students showed this type of response is an important element when we think about the future education of the students. I would like to briefly summarize the significance of participation. First of all, we were able to participate in the discussion regarding the framework. And in our country, we have an accreditation institution for licensing engineers. Thus, by participating, we could express our views to make the framework consistent with the criteria used for licensing engineers and actual conditions in the educational arena in Japan. We will also be able to express the thoughts behind the creation of such questions in the international arena. In the consortium of colleges or within the expert meeting, our opinions were heard with respect. The reason for this is that Japan’s engineering is respected globally and is supported by the graduates who have become engineers. So, regarding engineering education, we found out that there is a strong interest in our country; therefore, we ought to contribute more in these areas. Furthermore, we were able to work not just with the engineer experts, but with education experts as well. Thus, we realized that it is important to not limit ourselves to working only with engineers; we should also work with experts in other fields, such as education. We participated in the engineering domain of the feasibility study, and we feel that it is necessary to explain and pass on the findings, achievements, and insights to other subject areas. Regarding the multiple-choice questions, we found out that the questions Japan proposed were

at the internationally accepted level. We tried a question that was made somewhere else (not in Japan) and it was not used. So, the questions that our colleagues have been making for long years have been proven good for assessing elsewhere what we wanted to assess here in Japan. Regarding the constructed response test, there were difficulties in developing questions and in scoring. The questions will be translated into Japanese. Since the language frameworks of English and Japanese are different, are we going to translate it as it is? In Japan, we don't express the subject that much, but if we don't express the subject within the sentence, let's say for an ethical question, what happens if there's a defect in the dam? If no subject is mentioned, the answers will be derived from somebody else's opinion, and not from the engineers.

That is one of the experiences that we learned from. Regarding the constructed response task, the ratio of students who got the right answer has a relationship with the ratio of students who scored high in the multiple-choice questions. But I think we can conclude that combining the two types of questions is quite effective. I would also like to touch on AHELO's potential impacts. We were able to reconfirm Japanese engineering education from a global perspective. We need to maximize AHELO's education improvement effect, so that by disclosing all this information, such as the questions and scoring criteria, we will be able to maximize its global effect. And let me reiterate that in the future, I think these efforts need to be continued in various places. However, as pointed out previously, we also need to devise a way to make this sustainable from the perspective of manpower and costs. In other words, let's move this forward gradually without rushing into getting results. I have run out of time so I will have to skip my summary. But I have put in bullet points of what I have been speaking about on my slide. That is all. Thank you for your kind attention.

International Symposium on Educational Reform 2013
13:00 - 17:30, Tuesday, December 10, 2013 at Auditorium, MEXT

AHELO in Japan and Its Prospects

Kikuo Kishimoto
Dean, Graduate School of Engineering, School of Engineering
at the Tokyo Institute of Technology
AHELO Engineering Expert Group

1. Task Overview (Engineering)
2. Questions for Examination
3. Significance of Participation
4. Future Prospects
5. Summary



1

1. Task Overview (Engineering)

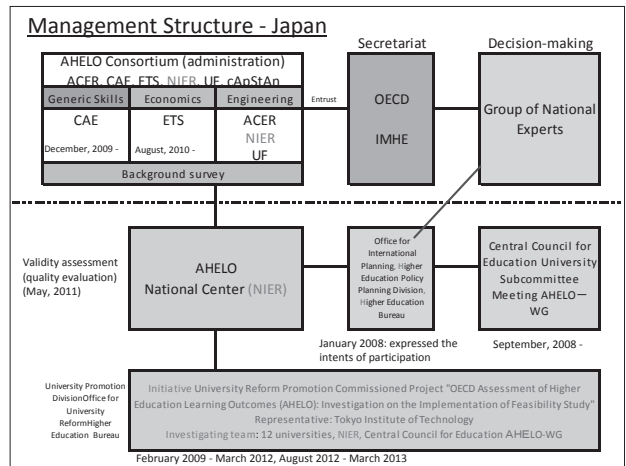


2

OECD-AHELO Assessment of Higher Education Learning Outcomes

- Purpose: To measure university education learning outcomes through a common test around the world.
- The feasibility study implemented between 2008 and 2012 intended to determine whether it was possible to globally assess learning outcomes.
 - Implemented for subject areas (Engineering and Economics) and generic (+ background survey)
 - Japan participated in the field of Engineering (12 universities, 504 students, 196 university teaching staff)
 - Final Report (Japanese edition is scheduled for publication in March, 2014)
 - AHELO Feasibility Study Report Volume 1 – Design and Implementation (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>)
 - AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf>)
 - AHELO Feasibility Study Report Volume 3– Further Insights (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf>)
- The employed competencies framework was Tuning-AHELO (Engineering and Economics)

3



4

OECD-AHELO Feasibility Study The Experimental Study to Assess Feasibility - 3 Phases

Phase 1 January, 2010 – June, 2011	Initial proof of concept [Validity assessment - qualitative evaluation] ● Is it possible to develop an assessment tool (measuring instrument = test) to draw reliable conclusions regarding learning outcomes while considering the diversity and specificity of each country? ⇒ Development of questions and scoring rubrics, small-scale field implementation, hearing < Contents of questions and its translation, revision of scoring rubrics >
Phase 2 - December, 2012	Scientific feasibility and proof of practicality [Validity assessment - quantitative evaluation] ● Is the assessment tool valid and reliable? ● Is it virtually possible to ask university students for their participation and to implement assessment appropriately? ⇒ Large-scale field implementation and scoring
Final meeting March, 2013	Whether to move AHELO forward with full-scale implementation is examined based on the results of (1) and (2).

- 3 areas (Generic Skills, Economics, Engineering + background)

http://www.oecd.org/document/22/0,3746,en_2649_39263238_40624662_1_1_1_1_00.html

5

Tuning in AHELO Engineering

Main agenda matters at Experts' Meeting in Brussels (May 4 - 5, 2009)

1. Academic/specialized domains of engineering
2. Major degrees in engineering
3. Major occupations
4. Categorization and ordering of learning outcomes statements
5. Definition of expected/desired learning outcomes
6. Expected/desired learning outcomes in each subject area
7. Definition of level indicators
8. General statements for engineers
9. Approaches to learning, teaching, and assessment

6

What are Learning Outcomes?

"Learning outcomes are statements of what a learner is expected to know, understand, and/or be able to demonstrate at the completion of a process of learning."

AHELO: Selecting learning outcomes in engineering

Selection will be done referring to ABET Engineering Criteria 2000 (member countries of Washington Accord use the same criteria while Japan uses JABEE Criteria) and EUR-ACE Learning Outcomes for First Cycle Bachelor Degrees.

=> "A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering"
 OECD Education Working Paper No.60
 (http://dx.doi.org/10.1787/5kghtchn8mbn-en)

7

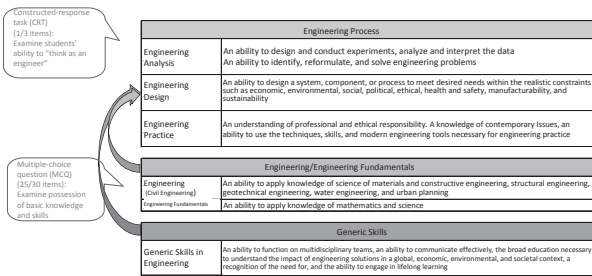
A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering pp.28-29

EUR-ACE framework for accreditation criteria of engineer education	ABET accreditation criteria of engineer education	Tuning-AHELO framework of learning outcomes
Knowledge and Understanding	a) An ability to apply knowledge of mathematics, sciences, and engineering	Basic and Engineering Sciences
Engineering Analysis	b) An ability to design and conduct experiments, analyze and interpret the data e) An ability to identify, reformulate, and solve engineering problems	Engineering Analysis
Engineering Design	c) An ability to design systems, components and processes to address needs within realistic constraints by issues on economy, environment, society, politics, health and safety, productivity, and sustainability	Engineering Design
Investigations	-	(Integrated into "engineering - analysis")
Engineering Practice	f) An understanding of professional and ethical responsibility j) A knowledge of contemporary issues k) An ability to use the techniques, skills, and modern engineering tools necessary for engineering practice	Engineering Practice (Including a part of "Transferable Skills")
Transferable Skills	d) An ability to function as a member of a interdisciplinary team g) An ability to communicate effectively h) Wide range of education to understand impacts of engineering solutions within the global, economic, environmental, and social contexts i) An ability to recognize the need for and engage in independent lifelong learning	Generic Skills (Including a part of "Knowledge and Understanding")

(http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en)

8

List of Tuning AHELO Key Concepts in Engineering



A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering (http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en) (Japanese edition is scheduled for publication in March, 2014)

ENGINEERING ASSESSMENT FRAMEWORK (Materials of the 8th GNE on November 18 & 19, 2013) (http://www.oecd.org/dataoecd/2/1/54619202.pdf) (http://www.oecd.org/dataoecd/2/1/54619202.pdf) (http://www.oecd.org/dataoecd/2/1/54619202.pdf)

9

2. Questions for Examination



10

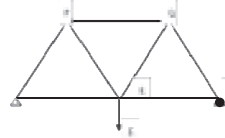
Development of Questions based on the Conceptual Framework

- The Consortium prepared the draft
 - Japan proposed multiple-choice questions (MCQs)
 - First-Step Professional Engineer Examination by the Institution of Professional Engineers, Japan
 - Civil engineering licensing examination of Japan Society of Civil Engineers
 - Australia proposed constructed-response tasks (CRTs)
- International experts' meeting:
 - Makes a choice from proposed questions
 - Reviews selected questions to improve them to the internationally accepted level

11

Scope of Learning Outcomes and Definition of Standards An Example of MCQ

http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf (pp.265-266)



A load P is applied to a Warren truss as shown below. If the self-weight of the members is ignored, which of the following statements is correct?

- Compressive force exists in both the upper-chord member (p-q) and the lower-chord member (r-s).
- Tensile force exists in both the upper-chord member (p-q) and the lower-chord member (r-s).
- Compressive force exists in the upper-chord member (p-q), while tensile force is applied to the lower-chord member (r-s).
- Tensile force exists in the upper-chord member (p-q), while compressive force is applied to the lower-chord member (r-s).

Focused competencies: engineering (systematic understanding of the key aspects and concepts of majoring branch of engineering)

* MCQs were reviewed to meet the internationally accepted level based on civil engineering licensing examination of Japan Society of Civil Engineers and First-Step Professional Engineer Examination by the Institution of Professional Engineers, Japan.

12

**Scope of Learning Outcomes and Definition of Standards
An Example of CRT**

<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf> (pp.252-264)



The Hoover Dam is a 221-metre high concrete arch-gravity dam in the Black Canyon of the Colorado River in the United States of America. It was built to provide irrigation water, to control floods, and to provide water for a hydroelectric power station at the base of the dam.
Figure 1: Hoover Dam
Figure 2: The site proposed for the dam in about 1921.
Figure 3: A sketch of the proposed reservoir.
Figure 4: Plans
Figure 5: Plans

1. Explain why this is a good dam site for hydroelectric power generation. You should discuss at least two aspects.
2. Explain the two main design features that contribute to the structural strength and stability of the Hoover dam.
3. The maximum electrical power generated by the turbines at the Hoover Dam is 2.08×10^9 W. What is the approximate amount of water that flows through the turbines at this power output, if the power station operates at 90% efficiency? Choose the closest option. A) 1038/sec, B) 1048/sec, C) 1068/sec, D) 1078/sec Note: $1 \text{ W} = 1 \text{ J/s}$, where $1 \text{ J} = 1 \text{ Nm}$.
4. Imagine that a new dam is being planned today in a different location. Briefly explain two environmental effects of the dam (which could also be upstream or downstream) that an engineer would need to consider in an environmental impact statement.

13

**Scope of Learning Outcomes and Definition of Standards
An Example of 1st Question**

- Learning outcomes to be assessed
 - Understanding of features suitable for a dam for hydroelectric power generation.
- Background competencies
 - Engineering (systematic understanding of the key aspects and concepts of majoring branch of engineering and the ability to apply the understanding)
 - Engineering analysis (the ability to identify engineering issues, create solutions, and solve the issues using existing methods)
- Answering aspects
 - a) Dam height/High potential energy
 - b) High flow rate of river
 - c) Lake capacity (size of water storage area and stable amount of water)
 - d) Minimal social impact
 - e) Characteristics of rock (hardness)
 - f) Narrow gorge

For each aspect, consensus was built during national/international scoring training on what extent of responses can get how many points.

14

**Scope of Learning Outcomes and Definition of Standards
An Example of 2nd Question**

- Learning outcomes to be assessed
 - Understanding of features that contribute to the structural strength and stability of the Hoover dam.
- Background competencies:
 - Engineering Design (an ability to understand and apply methods to design systems, components, and processes to satisfy requirements within the realistic constraints)
- Answering aspects
 - a) Arch-shape
 - b) Material in the canyon walls must be robust
 - c) The Heavy weight of concrete
 - d) Tapered shape of concrete wall/low center of gravity
 - e) Spillways and/or tunnels are equipped

15

Excerpt student comments (from the first phase)

[Multiple-choice question (MCQ)]

- Most of the questions were about what we learned at universities, and it was not difficult to solve them.
- All questions were one-shot, thus appeared to be wide and shallow. Questions can be more in-depth.

[Constructed-response task (CRT)]

- I found it good that the questions was on a practical issue, as we deal with more theoretical, abstract questions more often at university. I enjoyed the ethical question, too.
- It was a well-thought-out question that it made us think of the cause and then responses to it. I think it will be useful in practice in the future.
- If perspectives such as those adopted in the CRT-type questions are important, I feel we should have more group discussions and case studies in our classes. We haven't had them before, so we didn't have an opportunity to learn approaches to problem-solving.

16

3. Significance of Participation



17

Achievements from Participation in AHELO-FS

- Through the participation in establishing the framework for learning outcomes, we could express our views to make the framework consistent with criteria used for licensing engineers and actual conditions in the educational arena in Japan.
- Having participated in the question developing process (selecting subjects, making items, and developing/revising scoring rubrics), we could actively send information to the world from the perspective of Japan's engineering education.
 - In the world, there is a strong interest in and a high expectation from Japan's engineering education supporting Japanese science technology. Accordingly, it was recognized that Japan was expected to make international contribution in assessing learning outcomes in the field of engineering to respond to such an interest and expectation.
- Another thing that was realized through the participation was the importance of cooperative efforts among experts in the fields of engineering and education to improve educational systems within both national and international frameworks.
 - It is necessary to closely examine possible influences of such efforts in engineering on other subject areas, and to explain and pass on findings and achievements to other subject areas.

18

Insights Obtained from Experiences of Developing, Implementing, and Scoring the Examination

- **Multiple-choice question (MCQ) (Engineering fundamentals)** –
It was found that questions Japan proposed were at the internationally accepted level.
 - Professional Engineer Examination by the Institution of Professional Engineers, civil engineering licensing examination of Japan Society of Civil Engineers
 - No significant variation was identified among participating countries in terms of average score and distribution (first phase)
- **Constructed-response task (CRT) (Engineering processes)** –
Difficulties in developing questions and in scoring were shown.
 - Difficulties that can be got over by accumulating experiences.
 - Appropriate level of difficulty
 - Amount of information that can be dealt with within the given time of an examination
 - Translation Skills (word order, unfamiliarity with passive voice, omission of subject, and plural form)
 - Difficulties of uniformly measuring the ability to think
 - Difficulty in presenting the key point of a question: clearer the key point is made, more limited the scope of answer gets, and that inevitably shapes thinking.
 - Efforts to standardize scoring criteria actually narrow the acceptable range of a correct answer.

19

Insights Obtained from Experiences of Developing, Implementing, and Scoring the Examination

- **Conceptual framework “Engineering process is supported by learning engineering fundamentals”**
 - CRT and MCQ are highly correlated with each other (First phase).
 - Is it really necessary to include CRT-type questions in an examination?→ Yes
- **Meaning of CRT**
 - “Measuring competencies of engineering process using CRTs” is significant by itself in terms of experimental study. At the same time, presenting a clear message, “competencies are important in the engineering process,” has an essential implication as an initiative. Such international efforts are expected to be influential.
 - CRT can encourage university teaching staff to discuss in a cooperative and interdisciplinary manner about education programs necessary for students to gain such competencies, thus urging university teaching staff to embark on an effort for educational improvement.
 - Structuring education programs for the acquisition of competencies
 - Devising education methods (e.g., PBL, linkage among subject areas, and internship)

20

4. Future Prospects



21

AHELO's Potential Impacts

- Experts from different countries and universities discussed expected learning outcomes in engineering education, through which we could recognize in a tangible way that an internationally common view was emerging. The view was in turn successfully shaped as specific questions for examination.
 - Important materials for discussion on the international acceptability of engineering education in Japan
- To maximize AHELO's education improvement effect, it is indispensable to:
 - Disclose information on questions and scoring criteria, and associate it with education programs.
 - Promote the participation of teaching staff with various backgrounds from different countries in AHELO to deepen the understanding of the international framework for learning outcomes, to create a system useful to update, and to devise a system that enables information on effective educational approaches to be shared.
 - Devise a framework for survey (such as survey conditions and information disclosure) to improve validity of information obtained from AHELO for universities.

AHELO is expected to be a good opportunity for universities to work towards educational improvement in the international context.

22

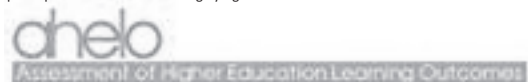
Future Prospects

- More experts will have opportunities to be involved in developing questions for examination and scoring rubrics through continued efforts. At the same time, it is necessary to make good use of the international assessment of learning outcomes to improve education by disclosing information on obtained findings and achievements.
- Expert groups are required to explore more effective international methods for assessing learning outcomes with a long view.
- It is necessary to identify such survey design and information disclosure that are useful to improve university education, and to take measures for their realization.
- It is necessary to devise a sustainable means from the aspects of human resources and cost and to identify an international assessment of learning outcomes that has significance for participating universities.

23

5. Summary

- It was recognized in a tangible way that an internationally common view was emerging regarding contents of learning outcomes in engineering education.
- It was realized that cooperative efforts among experts in the fields of engineering and education were important to improve educational systems within both national and international frameworks.
- In the world, there is a strong interest in and a high expectation from Japan's engineering education supporting Japanese science technology. Accordingly, it was recognized that Japan is expected to make international contribution in assessing learning outcomes in the field of engineering to respond to such an interest and expectation.
- A long view is required when discussing assessment methods for learning outcomes, especially regarding how and what are measured and compared with. For this matter, international expert groups need to provide further consideration. Japan's active participation in such efforts is highly significant.



24

Thank you for your attention.



25

Section 2: Panel Discussion

Report

Satoko Fukahori

Senior Researcher, Department for Higher Education Research, NIER

Informing Universities for Educational Improvement:
The AHELO Feasibility Study Experience in Japan,
Canada, and Australia

Panel Discussion

Facilitator

Motohisa Kaneko*

Professor, Research Center for University Studies, University of Tsukuba

Panelists

Robert Wagenaar

Director of Undergraduate and Graduate Studies, University of Groningen

Peter Ewell

Vice President, National Center for Higher Education Management Systems

Kikuo Kishimoto*

Dean, Graduate School of Engineering, School of Engineering, Tokyo Institute of Technology

Mary Catharine Lennon

Senior Research Analyst, Higher Education Quality Council of Ontario

Daniel Edwards

Principal Research Fellow, Australian Council for Educational Research

Satoko Fukahori*

Senior Researcher, Department for Higher Education Research, NIER

N.B.

- The * mark indicates that the original language of the speech was Japanese and that the transcript is a tentative translation based on the simultaneous interpretation provided during the symposium.
- The transcripts include changes made after the symposium for purpose of publication.
- The affiliations and professional titles of the speakers are as of December 10, 2013.

Report

Informing Universities for Educational Improvement The AHELO Feasibility Study Experience in Japan, Canada, and Australia

Senior Researcher, Department for Higher Education Research
National Institute for Educational Policy Research

Satoko Fukahori



1. Introduction

Assessment of Higher Education Learning Outcomes (AHELO) is an international initiative of the Organisation for Economic Co-operation and Development (OECD), aiming to assess the outcomes of university education using a common test across countries. A Feasibility Study was conducted between 2008 and 2012 to test whether it is possible to implement such an international learning outcomes assessment. With 17 participating countries and regions, the Feasibility Study investigated the cross-disciplinary strand of generic skills as well as the discipline-specific strands of economics and civil engineering. In addition to the tests in these fields, contextual dimension data were collected from all participating students. Japan took part in the Engineering strand together with eight other jurisdictions including Australia and Canada (Figure 1).

This paper will first outline the contextual background of the AHELO Feasibility Study, its purpose, and research design, and then describe the organizational structure and implementation process in Japan. Preliminary analysis of the study outcomes will then be presented, followed by a discussion on the significance of this international learning outcomes assessment.

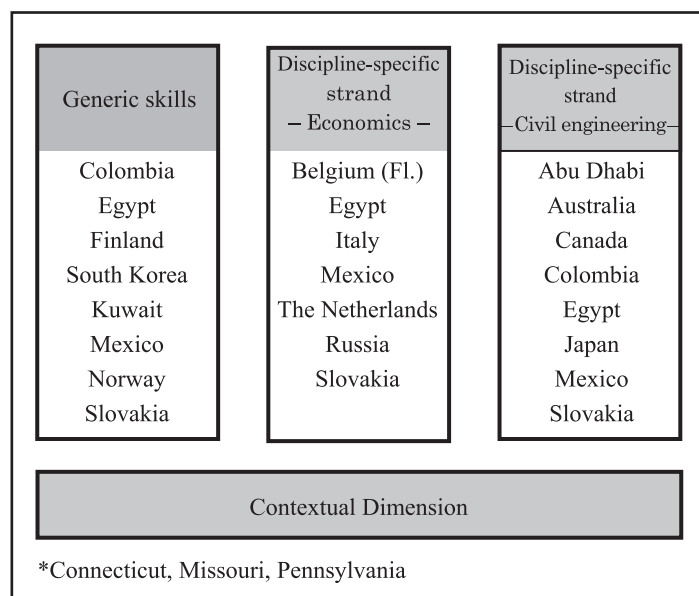


Figure 1. The AHELO Feasibility Study participating countries (by strand)

2. Background, purpose, and research design of The AHELO Feasibility Study

What is expected of graduates with bachelor degrees in terms of their knowledge, skills, and attitude (learning outcomes of bachelor degree programs)? Are higher education institutions successful in enabling students to acquire these qualities? As a growing number of students enter universities and as entry level qualifications and paths after graduation diversify, a global trend is emerging to examine the educational quality of higher education institutions from the viewpoint of learning outcomes. Furthermore, the increasing international mobility of students

and graduates is stirring an international interest in compatibility and comparability of academic credits and degrees between different countries.

The AHELO Feasibility Study was conceived in such a climate and developed with two goals in mind. The first goal was to verify the possibility of establishing an internationally comparable learning outcomes assessment by forming an international consensus on the knowledge, skills, and attitude students were expected to acquire through higher education, and to clarify whether it was possible to create assessment tools that validly and reliably measure students' levels of achievement.

The second goal was to determine whether it was possible at all to instigate participation from higher education institutions and students and conduct the assessment successfully. AHELO is not designed for uniformity or standardization of higher education. Nor does it expect to serve as an indicator for university ranking or governmental resource distribution. AHELO was an initiative aiming to improve the quality of higher education. The question was whether universities will find value in participating. Another question was whether it was possible, having explained the study's academic and political significance, to persuade students to partake in and undergo the assessment as seriously as they would in their actual courses knowing that the results would not be relevant to their personal academic records.

With these goals in mind, a research design was developed involving two phases. In Phase One, the study aimed to develop assessment tools and evaluate their validity. The plan was to develop cognitive instruments in "generic skills," "economics," and "civil engineering" as well as a questionnaire to gather "contextual" data from university students nearing graduation from bachelor degree programs. A small-scale field study was conducted to qualitatively verify the validity of the instruments. It was designed to include approximately 100 volunteer students from 10 invited diverse universities in each participating country and to administer the test as well as a questionnaire inquiring the participants' opinions about the validity of the test, followed by group discussions (focus groups) with academic staff and students.

Phase Two involved a large-scale field study using revised instruments and questionnaires from Phase One. The goal was to conduct a quantitative analysis to establish the validity and reliability of the assessment, as well as to verify whether implementation was realistic. The large-scale examination was designed to include approximately 1,000 randomly selected students from 10 universities in each participating country invited for participation. Students were administered the instruments and contextual dimension survey. Furthermore, institutional administrators and faculty were also administered contextual surveys (OECD, 2012: 77-95).

As the sampling design suggests, the sample groups in the AHELO Feasibility Study do not represent the higher education system as a whole in each country. Hence, the data collected is not suited to be used for "cross-national" comparative analysis. Sometimes the AHELO Feasibility Study is dubbed the "PISA for higher education," but it should be noted that sampling design is different from PISA, which employed two-stage stratified random sampling to select nationally representative schools and students, with 15-year-old students as the target population. AHELO is designed to produce information for the improvement of higher education, focusing on universities as the unit of analysis. The Feasibility Study in turn sets out to investigate the validity and reliability of the instruments developed for AHELO.

3. Organizational structures and implementation process in Japan

3.1 Organizational structures in international and national domains

Participating in the engineering strand of the AHELO Feasibility Study, organizations were arranged to pursue respective roles in the study as shown in Figure 2. Descriptions of the organizational structures in international and Japanese domains are as follows:

Beginning with the international structure, the AHELO Feasibility Study had a multilayered decision-making system involving the OECD Education Policy Committee (EDPC), the Programme on Institutional Management in Higher

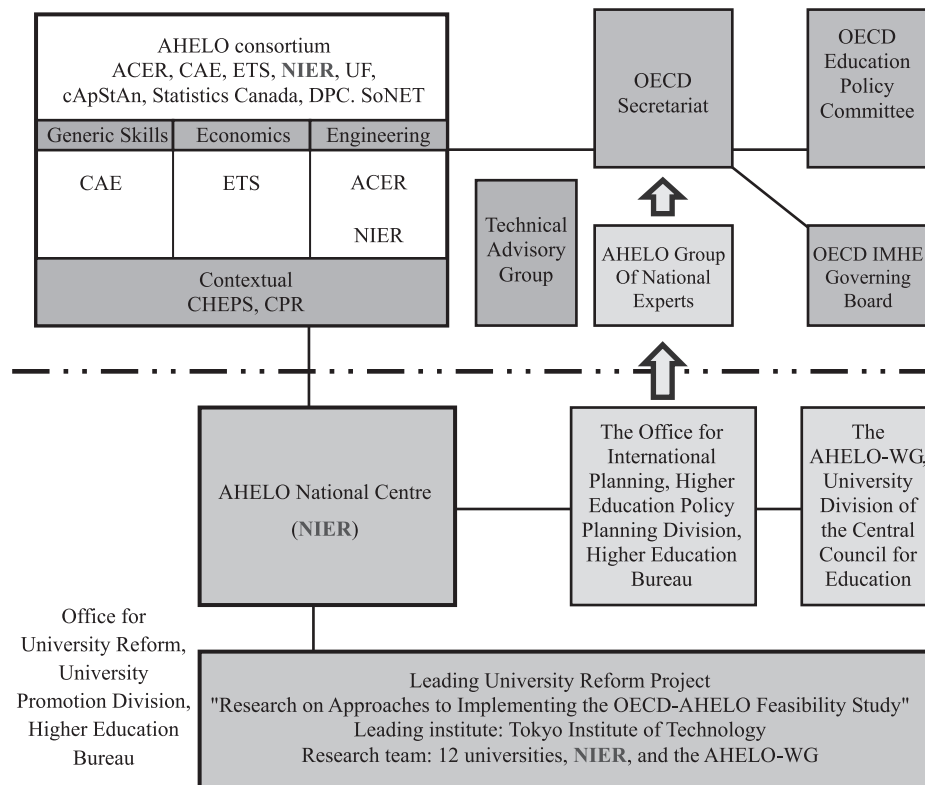


Figure 2. The AHELO Feasibility Study organizational structures (international and Japanese domains)

Education (later renamed as OECD Higher Education Programme; IMHE Governing Board or “IMHE GB”), and a Group of National Experts (GNE). The Secretariat undertook actual operations. EDPC, composed of representatives of OECD member countries, considered relevant issues to education policies in each country. IMHE, represented by IMHE GB, is a forum within OECD for discussions of issues regarding higher education, with higher education institutions and non-profit organizations dealing with problems in higher education as its members. The GNE is organized to discuss AHELO governance principles, joined by government and higher education institution representatives, mainly from participating countries of the Feasibility Study. As this organizational structure implicates, the Feasibility Study had to be conducted upon consensus among very diverse stakeholders associated with higher education.

The tasks of developing the instruments and questionnaire, managing implementation (examination) in participating countries, and analyzing the implementation data to verify validity and reliability were delegated to the AHELO Consortium, headed by the Australian Council for Educational Research (ACER) and comprising the following four groups: “Generic Skills” represented by the US Council for Aid to Education (CAE), “Economics” by the US Educational Testing Service (ETS), “Engineering” by ACER, the Japanese National Institute for Educational Policy Research (NIER), and University of Florence of Italy (Università degli Studi di Firenze, UF),¹ and the “Contextual dimension survey” undertaken by the Dutch Centre for Higher Education Policy Studies (CHEPS) and the Indiana University Center for Postsecondary Research of USA. Some technical specialists were also included in the AHELO Consortium; cApStAn Linguistic Quality Control Agency for translation, the International Association for the Evaluation of Educational Achievement (IEA) and the Data Processing and Research Center (DPC) for database development, SoNET systems for developing an online test system, and Statistics Canada for sampling. The Consortium held meetings regularly, and members pursued their own tasks on the basis of a shared understanding of AHELO principles and methodology as a rule.

¹ University of Florence is a coordinator organization of the European and Global Engineering Education academic network (EUGENE) with 76 member states.

Furthermore, a Technical Advisory Group (TAG) was established, consisting of experts in higher education evaluation and research, to provide GNE and the Consortium with third party expert advice (OECD, 2012: 96-100).

Regarding the organizational structure in Japan, the Office for International Planning, Higher Education Policy Planning Division, Higher Education Bureau of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) undertook administrative leadership in the AHELO Feasibility Study under the supervision of the AHELO working group, University Division of the Central Council for Education. NIER acting as the National Center carried out translation and validation of test questions as well as implementation of the assessment.

Furthermore, MEXT delegated to the Tokyo Institute of Technology a Leading University Reform Project “Research on Approaches to the Implementing the OECD-AHELO Feasibility Study” aimed at finding desirable ways of implementing the AHELO Feasibility Study in the engineering strand, and discussing what implications for university reform can be drawn from the experience. This research project proved to be instrumental in involving institutions to deliberate on the possibilities and challenges of AHELO in relation to higher education policies on quality assurance. It should also be noted that expert advice obtained from members of this research project enabled NIER to make proactive contributions to the AHELO Consortium in developing test items, as well as providing active feedback on the implementation process (Tokyo Institute of Technology, 2013).

3.2 Development of competence framework in the engineering strand

What learning outcomes should the assessment measure? “A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering” was developed by the Tuning Association based on the Tuning approach (OECD, 2011a)², and approved by the GNE to serve as the competence framework to be referenced for instrument development (OECD, 2012: 121-128).

Tuning refers to the methods and processes of defining competence frameworks that would outline the knowledge, skills, and attitudes students are expected to obtain through their degree programs in respective specialized fields, as well as designing degree programs based on the defined competence frameworks. The Tuning approach is unique in that it requires faculty to develop competence frameworks in consultation with stakeholders, including employers, graduates, and students using language that can be understood by all parties. This approach allows faculty to sustain academic ownership but at the same time be responsive to societal needs. The Tuning approach is unique in that it proposes a common framework that are sufficiently abstract so that diverse universities can share, but explicitly requires institutions to substantiate the competences into concrete learning outcomes that are measurable and attainable within a given timeframe. Degree programs should be developed in a way that are responsive to student needs and are in alignment with institutional missions, but at the same time guarantee the attainment of pursued learning outcomes. Because degrees and credits can be conferred only when the learning outcomes defined by individual institutions but are based on a common competence framework have been attained, Tuning serves both as an internal and external quality assurance system. (Gonzales and Wagenaar: translated by Fukahori and Takenaka, 2012).

The Tuning Project was initiated in 2000 by European universities with the purpose of substantiating the Bologna Process, aimed at establishing the European Higher Education Area. It is administered on a voluntary basis while being financially supported by the European Commission. Competence frameworks were defined in nine subject areas (Business Administration, Chemistry, Education Sciences, Earth Sciences, History, Mathematics, Physics, Nursing, and European Studies) during the first and second phases of the Tuning Project (2000-2004). Since then a diverse array of disciplines has joined the list. Tuning has been adopted in institutions in Latin America, USA, Canada, ,Russia, Africa, Australia, Central Asia, Thailand, and China (Tuning Association, 2013).

In engineering, even before the launch of the AHELO Feasibility Study, a number of pioneering efforts have been

²A similar conceptual competence framework based on the Tuning approach in the Economics strand (OECD, 2011b).

taken to establish the international comparability of engineering degree programs through mutual recognition of accreditation and cross border accreditation. For example, the Washington Accord was signed in 1989 between signatory countries such as the United States, the United Kingdom, etc. for the mutual recognition of accreditation systems in engineering education. The Japan Accreditation Board for Engineering Education (JABEE) joined the Accord in 2005 (IEA, 2013). In Europe, the EUR-ACE was launched in 2008 as an overarching cross border accreditation system for engineering education. Authorized accreditation agencies award EUR-ACE labels to engineering degree programs in recognition of their quality in terms of meeting EUR-ACE standards (ENAAEE, 2013). This trend is backed by an increasing number of engineers, whose activities extend beyond national borders, making indispensable a system to recognize the comparability of engineering qualifications in different countries.

In the AHELO Feasibility Study, the competence framework was identified by comparing the Washington Accord and EUR-ACE competence standards, and extracting their common features. As summarized in Table 1, the competence framework in civil engineering was categorized into the following five competence clusters: “Basic and Engineering Sciences,” “Engineering Analysis,” “Engineering Design,” “Engineering Practice,” and “Engineering Generic Skills.”

Table 1. Tuning-AHELO key competence framework in the engineering strand (Summary)

Competence cluster	Description
Basic and Engineering Sciences	Ability to apply knowledge in mathematics, science, and engineering.
Engineering Analysis	Ability to design and conduct experiments, and analyze and interpret data. Ability to identify, organize, and solve engineering tasks.
Engineering Design	Ability to design a system, factors, and processes to meet requirements under realistic conditions of economy, environment, society, politics, ethics, health, safety, production possibility, and sustainability.
Engineering Practice	Understanding of professional ethical responsibility. Knowledge of issues in modern society. Ability to utilize techniques and the latest engineering tools necessary for engineering practice.
Engineering Generic Skills	Ability to act as a member of an inter-disciplinary team. Effective communication skills. Extensive learning experience to understand the significance of engineering solutions in the contexts of international communities, economy, environment, and society. Positive attitude and capability to engage in lifelong learning.

The AHELO Consortium developed the engineering assessment instrument based on this competence framework. It was decided that the competence of “Basic and Engineering Sciences” would be measured by multiple-choice items suited to address the achievement of basic knowledge and skills, while the competences of “Engineering Analysis,” “Engineering Design,” “Engineering Practice,” and “Engineering Generic Skills” would be assessed by constructive response task items suited to address cognitive processes. Multiple-choice items were drafted based on the Civil Engineering Licensing Examination of the Japan Society of Civil Engineers (Organization for Promotion of Civil Engineering Technology, JSCE, 2013) and the First-Step Professional Engineer Examination by the Institution of Professional Engineers, Japan (The Institution of Professional Engineers, Japan, 2013). The constructive response tasks were prepared by the Australia team, drafting items designed to assess abilities in analytical reasoning and problem solving in real life settings. The instruments were adapted and finalized by an international expert committee, consisting of prominent experts in the field of engineering education in various countries (Australia, Japan, Italy, Germany, Sweden, the United States, and Mexico). Instruments were accompanied by scoring rubrics that describe the required scope and level of learning outcomes to be demonstrated by the students (OECD, 2012: 252-268).

3.3 Implementation

As stated above, the implementation of the AHELO Feasibility Study assessment took place in two phases. Phase One took place between January 2010 and June 2011, and aimed to conduct a small-scale implementation to qualitatively investigate whether the assessment instrument developed by the AHELO Consortium adequately addressed what students in each country had learned at university, as well as whether translated instruments had the quality to measure learning outcomes equal to the original; that is, if there were any discrepancies in the items in terms of the clarity and the levels of difficulty.

Universities were given different sets of items, with each set containing 20 multiple-choice items and one constructive response task, all chosen from the pool of items. The test was conducted in 60 minutes. After completing the examination, students were asked to complete a questionnaire asking their opinions about the validity of the assessment, followed by a group discussion with faculty. In Japan, the small-scale implementation took place between 16 and 25 May 2011, with 75 students volunteering from the 10 universities invited by the National Center to participate.

Phase Two was undertaken between July 2011 and December 2012, using instruments that were revised based on the outcomes of Phase One. It aimed to investigate quantitatively the assessment instrument's validity (whether the instrument actually measured the intended learning outcomes) and reliability (whether the instrument could reproduce the same results if repeated) on a large scale. Furthermore, it aimed to verify the possibility of maintaining certain levels of response rates with the cooperation of higher education institutions and students.

The instrument consisted of 25 multiple-choice items and one constructive response task. The combinations of items were determined based on Item Response Theory, and a total of 18 sets were created. Each student was given a set, and completed it in 90 minutes. In Japan, the National Center invited 12 universities (eight national and four private), and 504 students participated (target population: all civil engineering students; response rates: 12-100%; average: 65%). The implementation took place between 23 April and 25 May 2012. In addition, surveys were conducted with the universities and their academic staff (196 faculty members) (OECD, 2012: 147-172).

3.4 Scoring

In assuring the assessment's validity and reliability, it is imperative that scorers have the same understanding of the target competences and learning outcomes to be assessed, and carry out scoring on the same understanding. In the AHELO Feasibility Study, emphasis was placed on defining the criterion and levels of the scoring rubrics of the constructive response task as well as scorer training.

The small-scale implementation in Phase one adopted the scoring rubrics (first edition) prepared by the AHELO Consortium. In Japan, six engineering experts marked the responses of 75 students. Problems found in the rubrics through the scoring exercise were reported to the AHELO Consortium. For example, they requested additions to the criterion where potentially correct responses could not be accounted for, asked for clarifications of unclear points in the rubrics, and requested re-weighting of scores so that they would better correspond to their levels of importance. Based on such feedback provided from participating countries, the AHELO Consortium prepared revised scoring rubrics (second edition)

International scorer training (two two-day sessions) was held to prepare for the large-scale implementation of Phase Two. Lead Scorers from participating countries gathered at the training sessions. They practiced marking on sample responses taken from small-scale implementations in Japan and Australia, based on the revised scoring rubrics. They carried out discussions until the results of scoring assumed uniformity, and amended the rubrics as necessary (the third and final edition of the rubrics). It should be noted that providing sample Japanese student responses for the international scorer training was instrumental in assuring that the scoring rubrics corresponded well with the performance of Japanese students.

The international scorer training was followed by domestic training sessions in each country, under the leadership

of Lead Scorers. The AHELO Consortium developed an online scoring program for the purposes of training and actual marking, to ensure uniformity in the scoring process.

In Japan, 12 engineering experts, led by a Lead Scorer, engaged in a training and scoring exercise for three days. The training sessions were arranged in such a way that the Lead Scorer marked example answers beforehand, and every discrepancy in the marking results of the scorers was discussed among the scoring team members, until unified understanding of the scoring rubrics (final edition) was established.

The training yielded a high level of reliability in the actual scoring exercise. The scoring program was designed to have over 20% of the responses, randomly selected, to be assigned to two scorers, and this double-scoring yielded an 89.1% matching rate in the Japanese case. The discrepancies between two scorers were mediated at the discretion of Lead Scorers, while they provided the team with opportunities from time to time to ascertain the members' understanding of the scoring criterion and levels (OECD, 2012: 173-180; OECD, 2013a: 90-95).

4. Preliminary analysis of assessment results

4.1 Key concepts of assessment result analysis

As stated earlier, the AHELO Feasibility Study employed a convenience sample, where universities were invited by National Centers to participate. This being the case, the institutional sample is not representative of higher education systems in respective countries. As such, it is not appropriate to run a “cross-national” comparative analysis with AHELO data.

Furthermore, it is difficult, strictly speaking, to analyze the outcomes at the student level due to the fact that students took different sets of rotated items based on Item Response Theory. Although in principle, Item Response Theory presupposes the levels of difficulty and score distribution from existing data, enabling standardization and comparison between students answering different sets of items, because the AHELO Feasibility Study instruments are currently in the process of development, they lack the data that allow for such standardization. The decision to employ a research design based on Item Response Theory was a policy-related one, prioritizing the maximization of the number of items to be verified for validity and reliability while minimizing the burden on participating institutions and students, over drawing implications from the data analysis. It should be noted that employing a design that required sophisticated knowledge in psychometrics, resulted in making the instrument less accessible to content specialists, i.e. experts in engineering and education interested in interpreting the results of the tests.

The large-scale implementation was carried out with the purpose of quantitatively verifying the possibility of developing and implementing an internationally comparable learning outcomes assessment. From a psychometric analysis of the assessment outcomes, OECD reached the conclusion that the validity and reliability of the assessment were verified and that the assessment was implementable. OECD thus initiated discussions in preparation for the Main Study of the AHELO project (OECD, 2013b). However, no substantial proposal was made from the the AHELO Feasibility Study framework regarding what information AHELO could deliver that would effectively support higher education institutions in their efforts to improve education. In order to move on to an AHELO Main Study, it is imperative to gain the support from countries and higher education institutions. However, it has not yet been made clear as to what benefits the study will bring to stakeholders.

The research design employed in the AHELO Feasibility Study makes it inappropriate to conduct analyses on a “national” level, while it will take some time before sufficient information is accumulated so that AHELO may draw reliable estimates at the “student” level. What then can be said about the value in conducting a learning outcomes assessment for all the cost?

The following three major points were discussed in the aforementioned Leading University Reform Project “Research on Approaches to Implementing the OECD-AHELO Feasibility Study” regarding educational data that higher education institutions would expect from international learning outcomes assessments.

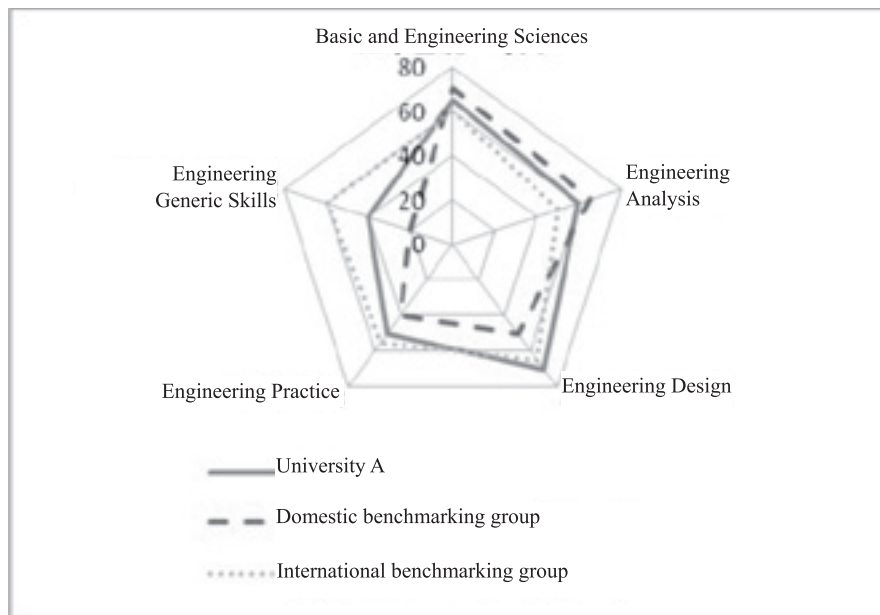


Figure 3. Competence profile (hypothetical)

The first point was about the need for data that served as international benchmarking. It enables universities to know how well their students perform in terms of their learning outcomes against an international standard. Their interests would not be in knowing the world ranking of Japanese universities “as a whole” or “on average.” Their interest would be in knowing the levels of their students against the standard of a given group of universities. It was not an issue whether the university group participating in benchmarking was representative of national higher education systems. Rather, they would value information about which specific universities comprised the group.

The second point was about the need for data that informed universities of their strengths and weaknesses. Universities expected information on student performance on the five competence clusters that the AHELO Feasibility Study instruments had aimed to measure, namely, “Basic and Engineering Sciences,” “Engineering Analysis,” “Engineering Design,” “Engineering Practice,” and “Engineering Generic Skills.” Universities would wish to know in which areas their students as a group excel and/or struggle, so that they may better tune their programs to student needs.

Such information would be made possible if students’ competence profile was presented as a radar graph showing the level of performance for each competence cluster, with benchmarking information for the national and international groups, as shown in Figure 3.

The third point was about the need for data that informed universities on the relationships between students’ learning environment, engagement, and test scores. Higher education institutions would utilize such data to address ways to improve their educational programs so that students may more effectively achieve their learning outcomes.

To what extent can AHELO satisfy higher education institutions’ expectations vested in the international assessment of learning outcomes? It must once again be emphasized that the AHELO Feasibility Study was designed to investigate if it is possible to implement an international learning outcomes assessment, and not directly to draw implications about student performance based on data analysis. Although the ultimate purpose of AHELO is to generate data that would inform institutions about educational improvement, the Feasibility Study data does not allow for such analysis. In particular, data collected in the Feasibility Study did not contain sufficient information to conduct analysis on competence clusters. Therefore, the results of a preliminary analysis will be presented here, not for the purpose of producing information on associations, but for the purpose of illustrating the kinds of information a future AHELO might provide.

In the sections below, the relationships between student feedback on the assessment instrument, perceptions about

their education, and test scores will be examined to discern the validity of the instrument. Then, the relationship between “student allocation of time,” “approaches to teaching and learning” at the university, and test scores will be examined. The scores used in the analysis are amalgamated scores of multiple-choice and constructive response tasks, which were obtained by estimating standardized scores based on results from multiple choice items that students have taken in common (average score 500, standard deviation at 100; 6,078 students from 70 higher education institutions in 9 countries).

4.2 Validity of instruments

To begin with, we will look at the validity of the instruments. Did the AHELO Feasibility Study successfully measure what it had intended to measure, namely what students learned through university education? We focus on whether the how students felt that the items were valid, and whether such perceptions correlated with test scores.

As shown in Figure 4, about 90% of students felt the instrument items were more or so “relevant to your current degree” (“Very much” 14%, “Quite a bit” 37%, and “Some” 37%), and the scores were higher among those who felt stronger compatibility. Similarly, 75% of students felt that the items asked questions “relevant to future professional practice” (“Very much” 7%, “Quite a bit” 24%, and “Some” 44%), and the scores were higher among those who agreed more with the statement. These results suggest that most students considered the AHELO Feasibility Study as a valid tool to assess the learning outcomes, and that the assessment was relatively successful in assessing the competences of students that were relatively well adapted to university education.

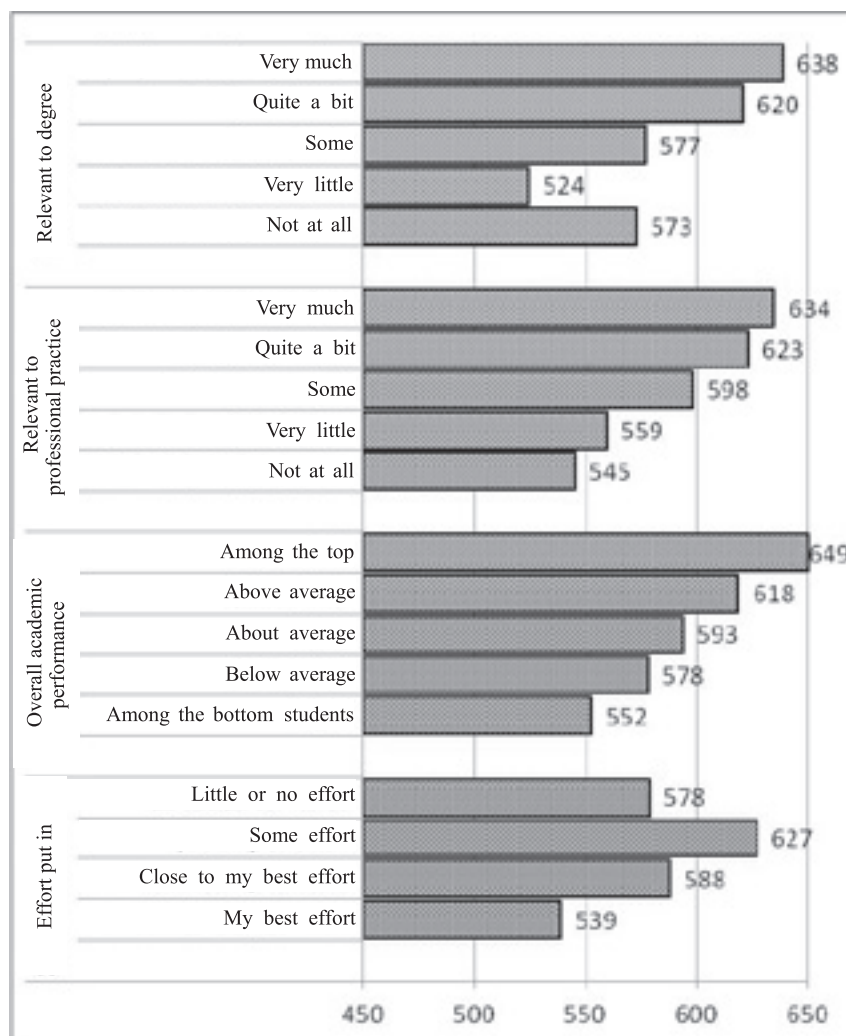


Figure 4. Validity of the instruments: students' perception and test scores

The test scores did not contradict these students' overall academic performance, as indicated by self-reported academic grades compared with other students in the same degree program (“among the top” 13%, “above average” 25%, “average” 24%, “below average” 22%, or “among the lowest” 16%). Regarding “the amount of effort put into taking the test” (“best effort” 15%, “close to best effort” 45%, “some effort” 40%, “little or no effort” 1%), the students who said they needed greater effort, that is, those who found the questions difficult, had poorer score results. These results again support the notion that the AHELO Feasibility Study assessment was successful in evaluating competence of students who were well adapted to university education.

Now let us turn to the relationship between students' test scores and student evaluation of the quality of their educational programs. About 90% of students considered the curricula they pursued to be “helpful in developing knowledge and skills they expect to use in future professional and working life” (“very much” 6%, “quite a bit” 30%, “some” 50%), again correlating between the extent to which they so felt and their scores. Regarding the students' evaluations of “their entire educational experience so far,” only half responded positively (“Excellent” 9%, “Good” 4%, “Fair” 41%, and “Poor” 6%), while the students with a better impression had higher scores. These results once again support the notion that the AHELO Feasibility Study assessment was successful in evaluating competence of students who were relatively well adapted to university education.

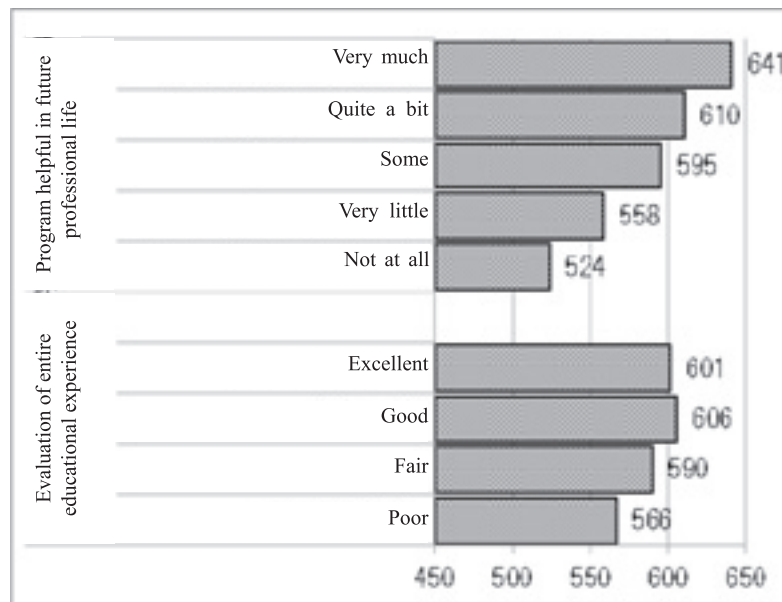


Figure 5. Validity of the instruments: evaluation of program and test scores

4.3 Student allocation of time

How can AHELO inform universities for educational improvement? In this preliminary analysis, let us explore the relationship between students' allocation of time and test scores.

Japanese higher education policies emphasize the importance of ensuring enough time for “preparation, course attendance, and post-class self-motivated learning” for students to accomplish high quality learning at HIGHER EDUCATION INSTITUTIONS (Central Council for Education, 2012). This comes from the awareness that the “total hours for learning” among Japanese students are no more than 4.6 a day on average (Center for Research on University Management and Policy (CRUMP), The University of Tokyo, National Survey of University Students).

In the AHELO Feasibility Study, students were asked their allocation of time during a typical mid-term week (7 days). The results were as follows (figures are weekly average): 18.9 hours for “attending courses (lectures, tutorials, seminars, etc.),” 5.7 hours for “engaging in practical exercises (laboratory work, fieldwork, etc.),” 6.8 hours for “course preparation (reading, doing homework, rehearsing presentations, etc.),” 0.5 hours for “part-time job relevant to their reading subjects,” 10.5 hours for “part-time job irrelevant to their reading subjects,” and 7.7 hours for “participating

in non-academic pursuits at university (on-campus journalism, students union, club activities, etc.).” The average time for these students engaging in academic pursuits amounts to 31.4 hours, or 4.5 hours a day. This figure matches the data from the research cited above.

Meanwhile, the students spend 18.2 hours a week for non-academic activities such as a part-time job irrelevant to their studies, and co-curricular activities on campus. It is notable that Japanese students rarely experience part-time jobs that are relevant to their studies.

To look at the relationship between the ways students allocate their time and their test scores, “attending formal classes” indicates that those who attend the formal courses for less than 10 hours (15%) tended to score low. This may suggest that this student group may be in need of attention for learning support. In terms of the “preparing for class,” the “0-5 hours” group, accounting for 56% of the total, had low scores, while the “11-15 hours” group accounting for 10% of the total achieved highest scores. Spending longer than 16 hours does not seem to contribute to better test scores. These results suggest that policies guiding students in the “0-5 hours” group to increase their self-initiated study to the “11-15 hours” level may be beneficial.

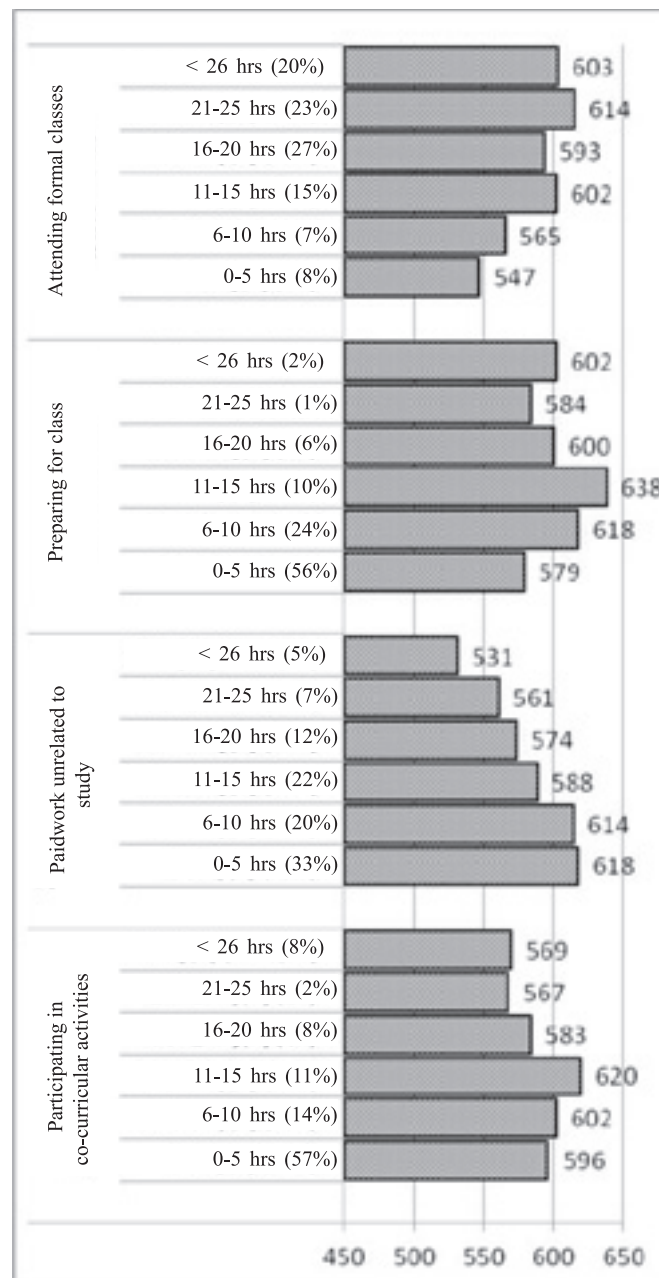


Figure 6. Allocation of time and test scores

On the other hand, engaging in part-time jobs irrelevant to their reading subjects has a negative impact. The longer hours students engaged in such part-time jobs, the lower their test scores. This tendency is starkly apparent for those who spend longer than 10 hours a week. It may be worthwhile to investigate the reason for students spending extremely long hours in occupations unrelated to their majoring studies and to adopt appropriate measures to extend student financial support. In contrast, research in Australia has found a positive impact on students' learning outcomes from engagement in part-time jobs relevant to their majors. Organizing opportunities for students to be involved in part-time jobs that are more relevant to their studies may be an option worth investigating.

Lastly, let us look at the time used for co-curricular pursuits. More than half of the students hardly take part in co-curricular activities, but high-achieving students are mostly found in the "11-15 hours" group (11%). Given that universities serve as a venue for learning as well as for socializing, it may be hypothesized that participation in non-academic on-campus activities brings about positive impact on students' learning outcomes through integration into a comprehensive university experience.

4.4 Approaches to teaching and learning

We now turn to the relationship between approaches to teaching and learning and students' test scores. In terms of the proportions of different course types involved in formal coursework by April of the fourth year, the average proportions of "lectures, where mainly the instructor speaks," "seminars or tutorials, where students are encouraged to discuss subject matter," and "group work, where students work together under supervision/facilitation from an instructor" were 74%, 20%, and 25%, respectively. Students experienced mostly lecture-based courses, and participatory-type courses such as seminars and group work were not prevalent.

As shown in Figure 7, cross-tabulations of course types and test scores show that students that take more "lectures"

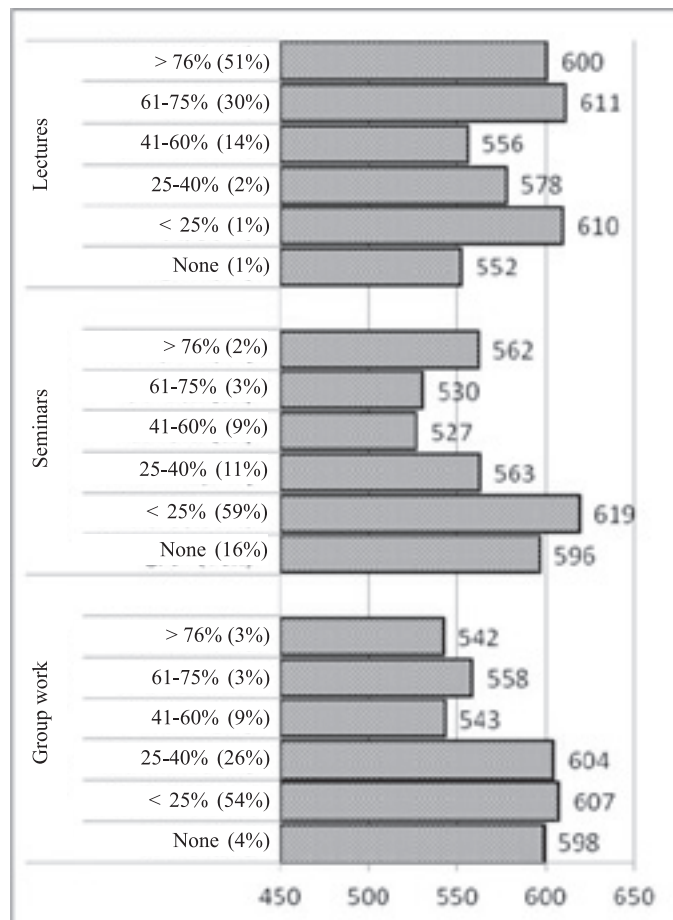


Figure 7. Approaches to teaching and learning and test scores (total)

(61% or more of all courses, accounting for 81% of students), fewer “seminars” (less than 25% of all courses, accounting for 75% of students), and fewer “group work” (40% or less of all courses, accounting for 84% of students) type courses tend to have higher test scores.

It is thus possible to draw the following two hypotheses. First, while universities with self-motivated, high-achieving students mostly keep the perpetual lecture-style without making much effort to change, ones with less academically motivated students are more creative in introducing student-centered participatory approaches, including seminars and group work.

Second, the educational impact of participatory approaches such as seminars and group work are difficult to assess. Encouraging self-motivated learning requires a well-developed curriculum design, thorough preparation, and excellent teaching skills on the part of teachers, the conditions for which may be premature in Japan as “lecture-style” courses are still dominant. Alternatively, it may be that by nature, the educational impact of participatory courses is difficult to translate into test scores.

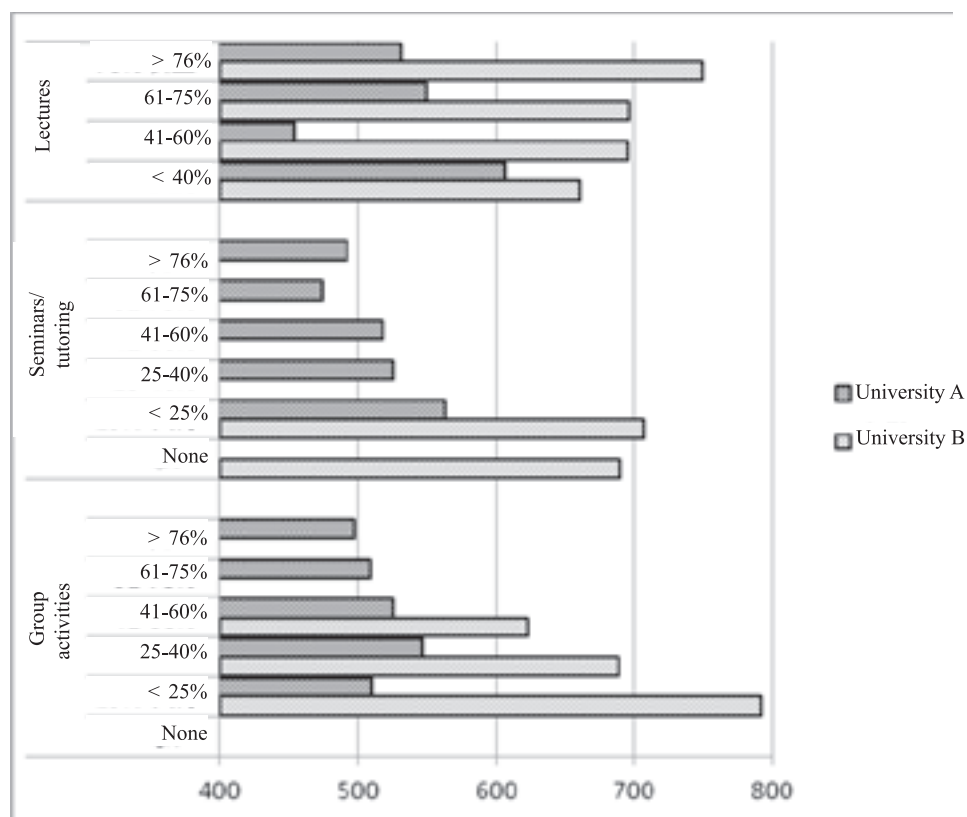


Figure 8. Test scores by approaches to teaching and learning (Universities A and B)

Figure 8 illustrates the difference between two universities, A ($n = 45$) and B ($n = 23$), in students' scores by approaches to teaching and learning. Due to the small sizes of the sample groups, the analysis is far from conclusive, but the observation seems to support the above two hypotheses.

University B, the higher achiever of the two, adopts more “lecture-style” classes and a fewer “seminars” and “group work.” Their students seem to adapt to this style relatively well. University A, on the other hand, provides fewer “lectures” and adopts “seminars” and “group work” in larger proportions. In University A, students taking 40% or less of their entire program had relatively higher test scores. However, no clear tendency is observed showing that students who take more seminars and group work perform better.

4.5 Informing universities for educational improvement

Examples from the preliminary analysis presented above show that the AHELO Feasibility Study has strong

potential for providing information universities can utilize for educational improvement. By producing data on achievement levels by competence clusters benchmarked against national and international peers, and by showing the relationships between students' learning environment, engagement, and test scores, AHELO can inform universities about their strengths and weaknesses and support strategic plans for improvement.

In order to fulfil this potential, AHELO would need to tackle the challenge of producing internationally benchmarked competence profiles of universities. Due to limitations in the data, the AHELO Feasibility Study analyses did not manage to create competence profiles indicating the strengths and weaknesses of universities. Further expert consideration will be necessary to re-design the assessment instrument so that this information will be made possible in a future AHELO.

As for international benchmarking, the National Centers in Japan, Australia, and Canada (Ontario) signed an agreement on data sharing, and have been conducting analyses using combined data. The agreement provides that the three parties mutually respect the confidentiality of the data that belong to each National Center, that the data should not be used for purposes other than for educational research, and that any of the data or analyses thereof should not be disclosed without written consent of other parties. In accord with this agreement, the three National Centers have endeavored to provide participating higher education institutions with international benchmarking data. This effort is expected to become an important precedent for protocols on international data sharing.

5. The significance of international assessment of learning outcomes

The AHELO Feasibility Study has been a significant step forward in that it successfully accumulated information necessary to initiate substantive discussions on the benefits the an international assessment of higher education would bring to stakeholders including, governments, higher education institutions, students, etc., and on the challenges that need to be addressed in order bring about those benefits.

In addition to this accomplishment, our participation in the engineering strand, and the experience gained from the participation in terms of instrument development, implementation, and scoring proved to be fruitful in itself. To close this report, three major accomplishments will be summarized here.

First, participation in the Feasibility Study provided us an opportunity to witness the actual process of internationally shared understandings on expected learning outcomes in the engineering strand develop, as experts from different countries and universities worked together. The assessment instruments and scoring rubrics were embodiments of the consensus. Active dissemination of information on what was learnt through this experience should guide future discussions on enhancing international comparability of Japanese education in engineering

Second, the fact that the Civil Engineering Licensing Examination of the Japan Society of Civil Engineers and the First-Step Professional Engineer Examination by the Institution of Professional Engineers Japan were adopted for an AHELO Feasibility Study multiple-choice items, and the fact that OECD has endorsed their validity and reliability for the international assessment of higher education learning outcomes signify the international comparability of these instruments. These facts signify international recognition that the Japanese standard for the “Basic and Engineering Sciences” competence are compatible with the international standard.

Third, participation in the Feasibility Study has given us updated knowledge about other components of the competence framework in civil engineering, namely “Engineering Analysis,” “Engineering Design,” “Engineering Practice,” and “Engineering Generic Skills.” Although there has been awareness of the importance of these competences, they have not yet been substantiated and widely adopted in educational programs in Japan. By shedding focus on measuring these competences through the AHELO Feasibility Study initiative, we have been able to send a clear message about their importance. The experience has also provided faculty and students concrete ideas about how to measure these competences.

Examples of student comments on the constructive response tasks are shown in Box 1 below. They symbolize the

powerful “awareness” that the AHELO Feasibility Study experience has brought about among participants regarding their university education. There are strong expectations for this international initiative to have far-reaching impact on teaching and learning approaches at universities.

Box 1. Students’ feedback on the written test

- Student A: “I found it good that the questions were on a practical issue, as we deal with more theoretical, abstract questions more often at university. I enjoyed the ethical question, too.”
- Student B: “It was a well-thought-out question in that it made us think of the cause and then respond to it. I think it will be useful in practice in the future.”
- Student C: “If what was tested in the written test is important, I feel we should have more group discussions and case studies in our classes. We haven't had them so far, so we haven't had the opportunity to learn approaches to problem-solving.”

References

- Central Education Council (2012), *Towards the Qualitative Change of University Education for a New Future*, Tokyo: Central Education Council.
- Tokyo Institute of Technology (2013), *Research on the Implementation of OECD AHELO Feasibility Study (final report)* (Leading University Reform Project sponsored by the Ministry of Education, Culture, Sports, Science and Technology).
- Center for Research on University Management and Policy (CRUMP), The University of Tokyo (2008), “*National Survey of University Students*” .
- Julia Gonzales and Robert Wagenaar eds. (2012). “*Tuning Educational Structures in Europe, Universities’ contribution to the Bologna Process: An introduction*” , translated by Satoko Fukahori and Toru Takenaka, Akashi Shoten.
- ENAEE (European Network for Accreditation of Engineering Education) (2008), *EUR-ACE (European Accreditation of Engineering Programmes) Framework Standards for the Accreditation of Engineering Programmes*. (http://www.enaee.eu/wp-content/uploads/2012/01/EUR-ACE_Framework-Standards_2008-11-0511.pdf)
- IEA(International Engineering Alliance) (2013), *Graduate Attributes and Professional Competencies* (<http://www.ieagrements.com/IEA-Grad-Attr-Prof-Competencies.pdf>)
- OECD (2011a), “A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering”, *OECD Education Working Papers*, No. 60, OECD Publishing. (doi: 10.1787/5kghtchn8mbn-en) (http://www.oecd-ilibrary.org/education/a-tuning-ahelo-conceptual-framework-of-expected-desired-learning-outcomes-in-engineering_5kghtchn8mbn-en)
- OECD (2011b), “Tuning-AHELO Conceptual Framework of Expected and Desired Learning Outcomes in Economics”, *OECD Education Working Papers*, No. 59, OECD Publishing. (doi: 10.1787/5kghtchwb3nn-en) (http://www.oecd-ilibrary.org/education/tuning-ahelo-conceptual-framework-of-expected-and-desired-learning-outcomes-in-economics_5kghtchwb3nn-en)
- OECD (2012), *Assessment of Higher Education Learning Outcomes Feasibility Study Report Volume 1 – Design and Implementation*. (<http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>)


Websites

- Organization for Promotion of Civil Engineering Technology, JSCE. *Japan Society of Civil Engineers qualification examination for certified civil engineer system - past examination papers*. (http://www.jsce.or.jp/opcet/02_testQ.shtml). (Accessed on November 01, 2013)
- The Institution of Professional Engineers, Japan. *Qualification examination and registration—information*. (http://www.engineer.or.jp/c_categories/index02019.html). (Accessed on November 01, 2013)
- JABEE. (<http://www.jabee.org/>). (Accessed on November 01, 2013)
- Tuning Association. *Tuning Educational Structures in Europe*. (<http://www.unideusto.org/tuningeu/>). (Accessed on November 01, 2013)

National Institute for Educational Policy Research
International Symposium on Educational Reform 2013
Tuesday, December 10, 2013 13:00~17:30

**Informing Universities
for Educational Improvement**
§
**The AHELO Feasibility Study Experience in
Japan, Canada, and Australia**

Satoko Fukahori
National Institute for Educational Policy Research



1

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.

Satoko Fukahori 2013.12.10 NIER Symposium 2

2

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”

Satoko Fukahori 2013.12.10 NIER Symposium 3

3

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.

Satoko Fukahori 2013.12.10 NIER Symposium 4

4

For the Panel Discussion:

What kinds of information can an international
Assessment of Higher Education Learning Outcomes
provide that will help universities improve their education?

Satoko Fukahori 2013.12.10 NIER Symposium 5

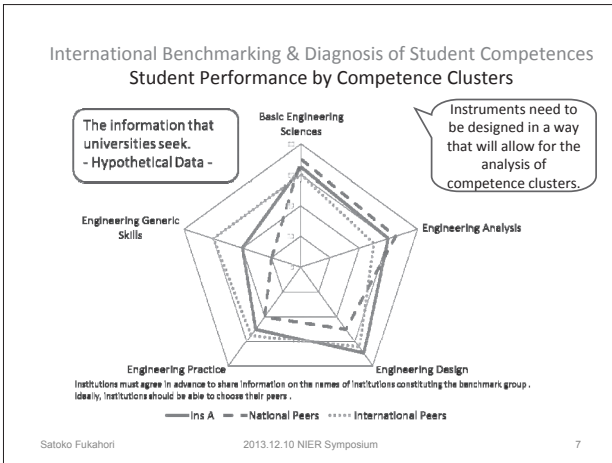
5

**What are the kinds of information that
universities seek from AHELO?**

- Consultation:
 - MEXT Commissioned Study on the OECD AHELO Feasibility Study, hosted by the Tokyo Institute of Technology (2009-2012).
 - 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?

Satoko Fukahori 2013.12.10 NIER Symposium 6

6



7

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

- The Central Council for Education (2012). *A Substantive Reform of University Education Towards Constructing a New Future*.

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.”

Satoko Fukahori 2013.12.10 NIER Symposium 8

8

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.

Satoko Fukahori 2013.12.10 NIER Symposium 9

9

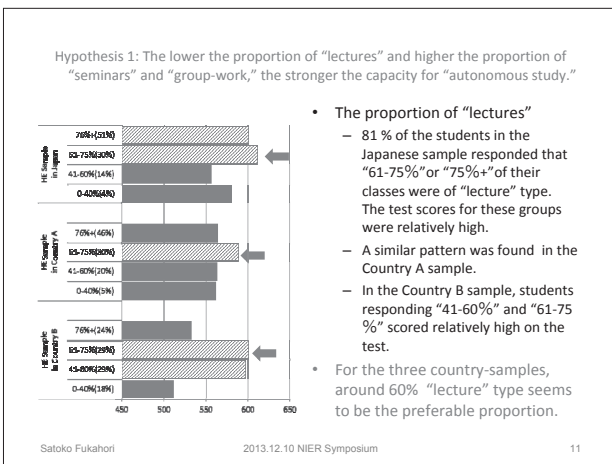
Sharing National Data

An MOU between Australia, Canada, and Japan

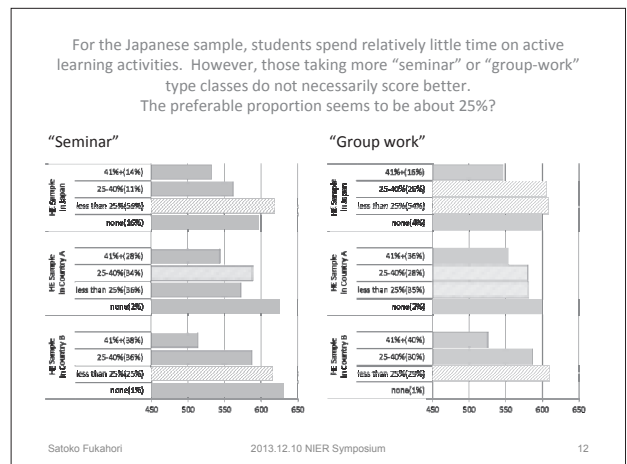
- For the purposes of educational research each Party desires to disclose and share its respective national Data with each other Party (“the Purpose”).
- The Parties agree that no Party shall disclose to any third Party, publish or otherwise disseminate any other Party’s national Data disclosed to it 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.

Satoko Fukahori 2013.12.10 NIER Symposium 10

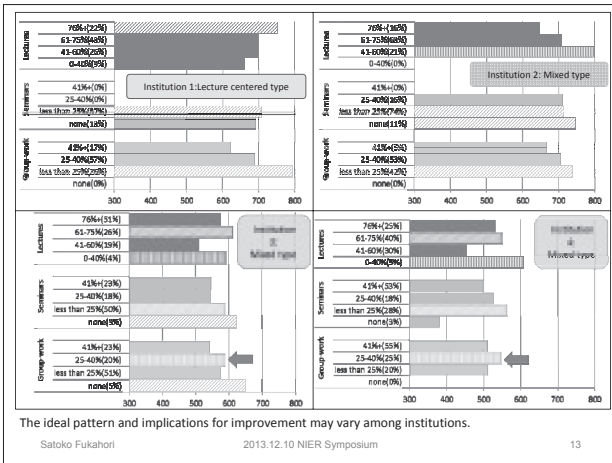
10



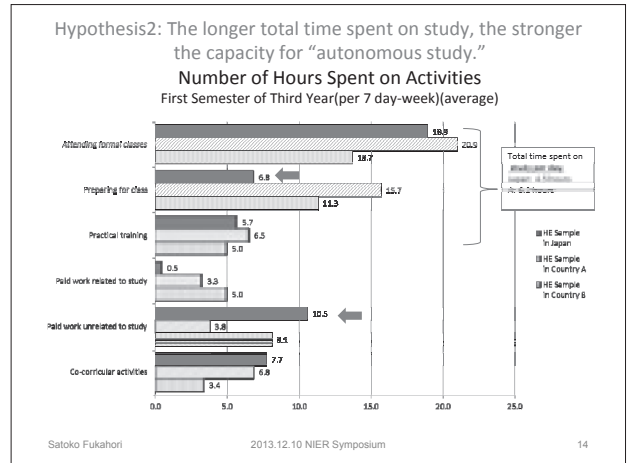
11



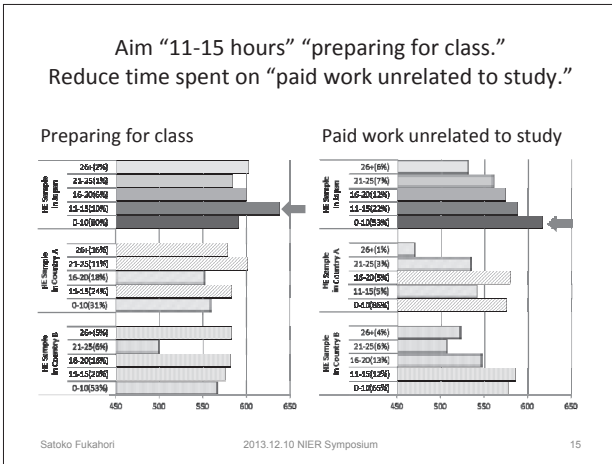
12



13



14



15

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.

Satoko Fukahori 2013.12.10 NIER Symposium 16

16

References

- OECD. *Testing student and university performance globally: OECD's AHELO*. (<http://www.oecd.org/edu/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdahelo.htm>)
- OECD (2011). *A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering*, OECD Education Working Papers, No.60, Paris: OECD Publishing.
- OECD (2012). *AHELO Feasibility Study Report Volume 1 – Design and Implementation*, Paris: OECD Publishing.
- OECD (2013a). *AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences*, Paris: OECD Publishing.
- OECD (2013b). *AHELO Feasibility Study Report Volume 3 – Further Insights*, Paris: OECD Publishing.
- Tuning Association. *Tuning Educational Structures in Europe*. (<http://www.unideusto.org/tuningeu/>)
- Gonzalez, J. and Wagenaar, R. (2008). *Tuning Educational Structures in Europe – Universities' Contribution to the Bologna Process*. Publicaciones de la Universidad de Deusto.

Satoko Fukahori 2013.12.10 NIER Symposium 17

17

Thank you for your attention!

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

Satoko Fukahori 2013.12.10 NIER Symposium 18

18

Panel Discussion

Facilitator Prof. Motohisa Kaneko

Panelists

Prof. Robert Wagenaar
 Dr. Peter Ewell
 Prof. Kikuo Kishimoto
 Ms. Mary Catharine Lennon
 Dr. Daniel Edwards
 Dr. Satoko Fukahori

■ **Motohisa Kaneko** : Thank you for giving me the microphone, but the burden is overwhelmingly big, because a huge number of questions have been received. It's going to be quite difficult to have all of these questions responded to. I believe that there are three basic issues. I would like to list those three major points. And I want to try to summarize all of the questions. I would like to apologize if I miss some of the questions in my summary at the outset, but please understand that there are time constraints. So, there are three fundamental issues. First of all, in order to do the test, in one way or another, the competency is conceptually summarized, and this is very relevant to Tuning meta-profiling or meta-competence or meta-skills. It's necessary to, in some way, conceptualize capability or competency. In the case of the United States, "generic skills" is the phrase that is frequently used. But the concepts themselves are based on certain logic. What is that logic? Tuning versus critical thinking is also frequently discussed in the United States. Is it possible to integrate these two concepts? Are they similar or the same? How are these concepts developed and how are they emerging? What would be the basis or logic if they are to be used empirically or demonstratively, especially when we try to use them as demonstrative tools? That is the first area. The second area concerns this: when we conduct such a large-scale test, what are the practical or logistical problems? You said that you were able to conduct the tests, but what were the practical and logistical results, especially in terms of cost? Dr. Ewell was mentioning this. How much expenditure or cost was incurred and what were the benefits? Who should be bearing the cost? How do we convince the general public that these tests are beneficial? How do we identify the effectiveness of these benefits or tests? Those are the questions related to the second issue.

The third issue is something that you are all interested in. What will happen in the future? The first sub-point is whether this can be developed under the umbrella of OECD on an international basis. Secondly, is it possible to apply this beyond the current scope to other areas? When I say "other areas," I mean other fields or subjects. For example, in the case of engineering, the area was limited to civil engineering, but even within the universe of engineering, there are various other subject fields. What about liberal arts? Is this applicable to liberal arts? Or further, inside Japan, this may sound closer to PISA in secondary education, but can we also expand this and include other tertiary education institutes within the scope?

Those are the three issues that have been raised, and the many questions submitted can be classified under these three major groups. Unfortunately, we have less than an hour for discussion. I wish I could ask each of you to answer all these questions, but I'd appreciate it if Prof. Wagenaar and Dr. Ewell could focus on the first question, which is the framework of capability or competency. Dr. Edwards will be asked to comment on the Generic skills assessment, which is based on CLA and ACE that were used to choose the problems or questions asked. If you compare them, can you point out the differences or anything that



Prof. Motohisa Kaneko

comes to your mind by way of comparison? And with regard to the second issue, the practical issue of costs and benefits, who will bear them? I would like to ask Ms. Lennon to share with us the Canadian case and Dr. Fukahori, to share with us the Japanese case. How was the arrangement done in Japan and how will it be done in the future pathway? Of course, we will have to ask Dr. Ewell to comment. Prof. Kishimoto, we would appreciate it if you could expound on whether it's possible to do this further in Japan and whether you think that it's possible to do this in other subareas of engineering other than civil engineering. You can touch on other points first, then the three, but for the three particular points that I raised, I will be asking you those questions.

Is there any problem with the questions? Okay, sure. So, who would like to start? Prof. Wagenaar? Please take the floor. From your side, we go one by one, following the seating order.

■ **Robert Wagenaar** : OK, this is a difficult question and I think I have to be open with you. From the very start I had my doubts about CLA, to be frank with you. And the reason is -- which was already explained by Peter -- that it is an instrument which has been developed in one context, the American context. And I suppose in one context it has proven its usefulness but the moment you move it to another context, it might lead to problems. And the CLA involved two issues here. Of course there are differences between cultures throughout the world. That's one issue. Secondly and I focused on this already during my presentation, it makes a difference whether it's one discipline or another discipline. So the people I work with, the many academics, thousands of academics throughout the world now, they seem to be convinced that if you develop generic competency, and they should be developed -- that's without any question, because it's important for society -- it should be done in the context of a domain of knowledge. If you develop them in the domain of knowledge, you should also test them, you should measure them in that domain of knowledge. That seems to be fair. But the fact that we are operating in different cultures, and you get the example of, for example, the competency conceptualisation. We have to find out whether there are differences, we still have to find out whether there are differences between different regions. I think for good reasons, Tuning organized its frameworks region by region. And at the moment, we are busy comparing what has actually been the outcomes, whether we have a common understanding in terms of the cultural and also the educational settings. And that will show us, whether we can actually compare those concepts. I think there's a good chance that it will be possible because we also learnt from engineering and economics that if the moment you, in particular engineering in this case, if you link them to a field of study, there're a lot of common grounds.

So I think there is prospect, and you have asked me about the future. I'm not sure or maybe I'm sure that AHELO was to a certain extent a risky affair. Because we have checked, we have tried to find common ground between 17 jurisdictions and some of those jurisdictions came in rather late and also that might have been a problem. I'm not sure whether we should not test this first region by region and then compare the regions and I brought this also up when we were discussing this in the TAG, and that step-by-step approach might probably be more promising than what we see at the moment.



Prof. Robert Wagenaar

There's also another model: we start in one region and we open that region up to other regions, in a sort of circular system. That might take time, but also might be more rewarding. A few things from the perspective of Tuning about the cost and benefits. Because we have been working very hard in trying to frame what the disciplines stand for, we learned a lot from each other when making these frameworks, working all over the world. We also learned that some systems are more effective in developing a certain set of competences than others, but to be sure, we have to compare in one way or another, to learn more about it. But most of all it's the responsibility, or should be the responsibility of the academic world. The issue, I wouldn't say the problem, but at least a challenge with regard to AHELO was that

it was a top down approach. It was initiated by the Ministries and the universities were invited to participate at random. But also, with regard to the students, as Peter showed us, we weren't sure that actually students showed up, in such a way that you get reliable outcomes. Maybe we should turn it around, maybe we should give a more prominent role to the universities, a more leading role in the next step. There will be a next step, I'm absolutely convinced because I think there's not much alternative if we are working in a competitive world, also we should be sure that what we are doing can be compared and therefore, be measured. But we have to find the right framework and I'm not convinced that the framework we have developed so far seems to be the right one.

■ **Peter Ewell** : We had no framework for generic skills. We began with an existing instrument, and there is nothing more “top-down” than beginning with something that should have been developed incrementally. I think we learned that lesson and I would not recommend doing this again if we were to go forward.

I think what you're seeing in many respects with Tuning and AHELO is the top-down and the bottom-up growing out of the discussions among teaching staff, is that they are meeting in the middle. I think that what we'll find is a situation--you've called it sector competencies or skills--where we can span these things a little bit more. We had many discussions in the TAG about what the architecture of an assessment might be that could do that and we did suggest two significant modifications to what was done. If there is to be a free-standing generic skills instrument, it should be heavily discipline “flavoured.” Ironically, the examination that the CLA was derived from was not about everyday situations. It was anchored in the disciplines and I think we need to honour this approach in moving forward.

The other thing that we recommended is that there has to be, regardless of discipline, a common core of items that is administered to everybody, so that you can see the way in which different disciplines perform. Getting back to the question of the conceptual basis of any AHELO instrument, it is hopeless to try to assess a domain which is not well specified. That's why you have to do the Tuning work first in order to get some kind of agreement on what you're really talking about. So I think there will be disciplines, and this is skipping to your last question, Professor Kaneko, which just cannot be done in AHELO. They are simply too culture specific and context specific. I would feel fairly confident about assessing all parts of engineering in AHELO and I think that there are many practice disciplines that could be done quite well. But Law? I don't think so. History? I don't think so. There are a number of disciplines that would be very, very difficult.

■ **Kikuo Kishimoto** : All three questions are very difficult to answer. So I would like to provide the answers to the extent that I can. One is whether the scope will expand from engineering. In terms of the learning outcome conceptual framework, what would be the basis for that? In the case of engineering education, as the students and engineers had high mobility, the conceptual framework had to be sorted out in the accreditation of the framework in that perspective. Well, of course, there's the question of whether what we have is perfect or not. Nevertheless, the opinions of the people of the universities, the employers, and the engineering students have been reflected in the discussion of what is the important learning outcome.

In that sense, the group in Europe and the Washington Accord Group have almost the same results, and what is required of engineers has been fixed to some extent. However, the capability required of engineers has changed with the times, and the framework must be deepened. We need to continue the discussion



Dr. Peter Ewell



Prof. Kikuo Kishimoto

from that perspective.

At AHELO, questions will be created and tested. First, we'll have to find out whether the learning outcomes can really be checked. We haven't had enough time to reflect on that, as there are still other things that we need to work on in terms of application capabilities. How we can measure that is one issue. As Prof. Wagenaar mentioned, the capability to work within a team is considered very important. But if the testing is measured on an individual basis, can we really measure the teamwork capability? With regard to communication capabilities, scientists and engineers communicate in various ways, including charts and graphs. It would be ideal to have a test in place with which we would be able to measure all of those capabilities. We may not go there 100 percent, but by continuing at this, I think that we would need better means of measurement.

The students are learning not only in one country, but in various countries. As they go from one university to another, the kind of education being conducted in each university must be considered. In terms of whether this would expand, Japan intends to provide globally competitive education programs in engineering. We thus need to check where we are and what kind of performance can be expected from the students enrolled in our programs. I think that this should be done with other universities, within the cooperative framework.

We have been working with various universities under the AHELO initiative, and I feel that other university professors may want to do this kind of cooperative framework. But who would be paying for that? And who would be providing resources to conduct this kind of initiative? These are very big challenges. In addition, our collaboration with other universities and international consortiums has shown us that Asian countries want to undertake this kind of initiative. So, how do we proceed with that? Well, in the AHELO initiative, OECD has taken the lead. When the name OECD is attached to an initiative, it's easier to find people who want to participate. Also, if we are able to create university frameworks, maybe on a regional basis, it would be better. That is the comment that I would like to make. Thank you very much.

■ **MC Lennon** : Thank you. I have been asked to speak a little bit about some of the practical and logistical results, if the cost benefits actually justify some of the work that we've done. I think that as a feasibility study it's naturally a slightly more expensive endeavour than a full scale activity would be. But have the costs actually justified the work? I know in Canada, because we don't have a Federal government that's responsible for Education, we run provincially, a number of the provinces did not choose to participate, because they felt it was a very costly endeavour, because many of the provinces don't actually have that many institutions to compare. Some of them have one institution per province.

So that upfront cost and the uncertainty of what was actually going to be provided back to the provinces, was certainly a consideration that we dealt with. It was Ontario's choice to participate and entirely take on the costs associated with the project, because we were not only investigating the learning outcomes assessment through the OECD's AHELO project, but we were involved in a number of other initiatives as well. So we recognised that there's a value in investigating and investing in learning outcomes by participating in Tuning projects, by participating in implementation activities for institutions to actually embed learning outcomes in their projects.

So ultimately we of the province felt that there was actually good reason to invest in this activity, and the impression or the desire of the province was that we were hoping to get some international comparative information to really see how we benchmark ourselves across the world. And as it turned we were going to be able to have some comparative information, and the OECD chose not to present that to us in a formal way, but it was something that we as the Australians, the Japanese and the Canadians, felt was a really important part of this study, and so collaboratively we decided to seek out that comparative information.



Ms. Mary Catharine Lennon

Our Institutions were very interested in participating and they did so free of charge, we gave them a very nominal fee to engage in the activities, and they did so with gusto, they were really quite happy to participate in it, and felt that they were going to be given significant information back to help them inform their teaching and learning and to help their students actually and engage in the international world that we live in, particularly in the engineering programmes in Ontario. They're a very international student population, they're a very international faculty, and we know that they're very mobile as well. So they were very, very interested in participating with the international comparisons.

I think overall, the province, the faculty and the institutions were quite satisfied with what they received based on the costs that they'd put into it. And though the institutions were very much looking forward to having a bit of a better idea about the competencies, the certain clusters of competencies and how they compared to each other and compared to other jurisdictions. We realised that as a feasibility study, there are a number of issues that still need to be worked out. And we realise that there's costs associated with that. So they, though disappointed with certain aspects, they didn't lose heart in the project, and having read the report and received both the institutional reports from the OECD, and the jurisdictional report that's recently been presented to them, they're pleased with the results, they're interested in working with them and, but mostly they're very much looking forward to our collaborative report that we're going to be working with.

So the benefits I think are there for the faculty and certainly for the institutions. And I think they're continuing to want to engage in these sorts of projects. The question I suppose, is whether or not the province itself has as much to gain out of it, other than supporting the institutions and their activities, without the opportunity to have international comparisons, it's a question as to whether or not the province benefits as much as perhaps it might, if the comparisons were available.

■ **Daniel Edwards** : Thank you. I'll speak to the three points, but starting with the conceptualising the competencies and starting with the generic skills, I think Professor Ewell and Wagenaar have already sort of covered a bit of that. But I think I'd like to juxtapose the way that the framework for engineering was developed and then the instrument versus the way that the generic skills were developed that were pretty much the opposite way, so the generic skills, as we've said, the test existed, but a framework was retro-fitted on top of it, whereas the way that, in civil engineering is where I did a lot of, had more of the involvement, as Professor Kishimoto has said, it involved employers, it involved students and it involved academics most importantly and educators in identifying these competencies through the Tuning project and then I guess ret-tuning it to fit AHELO. And then thinking about once we had these competencies, what sort of questions would be best to put on top of the framework and actually give us some outcomes or give us a measurable, questions that we could find in the outcomes. That issue didn't happen with generic skills, because of the way that things were designed, there's no necessary blame involved, but it also points to the difficulties with generic skills as a concept. And I think if you have a look at the civil engineering framework, you will see that there are special noted engineering generic skills and it's one of the areas of the framework and I think and I agree with Professor Wagenaar in particular, who's said a number of times today, how important it is to think about the discipline based generic skills. There are engineering generic skills that just simply aren't generic skills, in sociology for example. But then there are other generic skills which span all disciplines, but perhaps are used slightly differently by people in different disciplines.

So I think it's a very difficult and contested area, even in terms of something relatively more simple like civil engineering, we produced the framework, if you got another ten academics together, they would produce a different framework. There would be overlap of course there's overlap, but there's no right or wrong



Dr. Daniel Edwards

way of doing this. We're doing some other work in medicine at the moment and I'll talk about that at the end but we have a similar issue. You can never really nail it as we might say, there's always someone who's got a different opinion and a valid opinion to change things.

I was involved in AHELO as part of the consortium but also as the National Project Manager for Australia, so in terms of some of the logistical issues, I'll just mention it briefly, I think one of the key take-homes for Australia was logistically this can be done. We had the systems, the information technology but also the student record data, and the interest from academics to do this, so logistically it can be done but I guess the big question is what is the cost? For Australia, the costs were relatively large and the government did contribute a lot to the OCED but also to running the programme nationally. But it's not just money that's involved, it's commitment of time, and resources within institutions which we found in Australia to be the most important part in encouraging student participation, and encouraging more than just people going and sitting and taking the test and spitting out data. What we found was the institutions that put in more commitment got more out of it, in terms of discussion about what this is all about. And ultimately what this is about is about making sure that the young people who go through our universities come out at the other end with skills and competencies that will equip them in the work that they would like to do in the future. And sometimes that gets lost I think.

So what we found in Australia while I think there are limited, there are significant limitations in the data, we had low responses from students and we were trying to play down the emphasis on the data, what we did get was the beginning of conversations about learning outcomes, about what it is that students should be learning, the framework development and the Tuning process has been very good in terms of thinking about the competencies. And then the AHELO process has been very good in terms of thinking about well how do you assess those, how do you measure whether the students who are coming out actually have these competencies. And that's been very good and I mean the idea of thinking like an engineer, was sort of a catchphrase that we used all the way through. And it helped spark the discussion not only between academics, deans and their faculty but also between students. In the more engaged universities, there were very good discussions between students about, oh hang on, what we do in class is quite different from what this test is asking us to do. And Professor Kishimoto mentioned this before that especially the constructed responses and they were saying, we want more of this project-based, project management problem based approaches in terms of assessment, to be able to understand what we might face out there in the real world. And prove perhaps that we can think through these issues and provide you know solutions that others think are relevant. So that was a very important part for Australia I think, in terms of what we gained from it, it's the conversations, and it's the starting to think about that.

Just briefly, on the future, I think the others have mentioned, or question, there's been some questions about the use of the OECD as a country-based model for disseminating this kind of information and I think there are good questions about how this must be done in the future whether the OECD is the right conduit for producing, for managing such a system.

At the moment, I guess they're the ones who are putting their hands up trying to find the money and you are taking the initiative. But that's not to say that that initiative will be followed through and we'll wait and see what happens there.

One thing's for sure and I know Robert has said the same thing is this, almost certainly, is going to happen. We have, and I'm sure Japan is interested, Australia is fixated with university rankings, but all the metrics that you see in the university rankings are research based metrics. OECD and AHELO doesn't want to be a ranking and I personally don't think any of it should be but the point is that the way the comparisons and benchmarks are being made at the moment are all based on research, research metrics, bibliometrics stuff. The teaching and learning side of things, which is one of the essential core roles of the university, is not benchmarked and universities who are doing a good job, cannot necessarily prove that using a measure. So I think it will go ahead.

In terms of other areas we're doing some work with Canada, with Japan, talking with Germany about mechanical engineering, a more collaborative exercise that we're thinking about. It's much more bottom up than top-down. And also in Australia we've been developing quite a successful collaboration in medicine, where we're looking at, we've built a framework in assessing medical students' outcomes. It's Australian based at the moment but there are similar initiatives going on elsewhere which could match up. So I think medicine is another one of those ones which is, its context dependent on a lot of things, but there are some areas that are common.

■ **Satoko Fukahori** : Thank you very much. I would like to provide an answer to the second question. In regard to cost, MEXT has contributed money to OECD, and also to the five year research project with the engineering experts to examine optimal approaches to participating in the AHELO initiative. NIER has covered significant costs associated with translation, adaptation, and verification of instruments. Honorariums were provided to students for their participation. So, the monetary cost was not small. Apart from the monetary cost, I must emphasize that the five-year commitment by engineering and education experts have been enormous, and I would like to extend my appreciation to all involved for their generous contributions.

So, what were the benefits derived from those costs? I have been involved in the initiative for five years as a coordinator, and as Prof. Kishimoto mentioned in his presentation, we did indeed learn a lot from the experience. We were able to confirm that the internationally developed competence framework was acceptable in Japan and that the Japanese situation was relatively well aligned with those competences. We were able to experience the process of defining concrete learning outcomes from abstract competences, and substantiating them to assessment items. Through scoring, we experienced building consensus on the scope and level of expected learning outcomes for bachelor students. This experience had rich implications in itself. I'm hoping that the people who have been involved in this initiative feel the same way, and that this experience will have some irreversible impacts on them. And I am hoping that the professors will share what they have learned from this experience with their colleagues. What I regret is that we have not been able to provide the resources necessary for professors to persuade their colleagues. Because of confidentiality issues, only a part of the test items and scoring rubrics developed has been publicized. If AHELO were to have an impact on the quality of education, we need to share the test items and scoring rubrics with faculty, who need the information to be able to share competence frameworks and agree on the scope and levels of bachelor learning outcomes. We need to publish these materials to be widely utilized for faculty development. I would have liked to provide participating universities with outlooks for extensive feedback on benchmarked competence profiles, so that professors could discuss more seriously with their colleagues about the benefits of participating in these kinds of initiatives. That is something we need to work on, but if we are able to do that, I think initiatives like AHELO would be of great benefit to universities. For any university education reform to have an impact, it is the professors that need to be convinced and take action. The role of MEXT, or the Ministry of Education, is to provide a platform for that to happen through issuing initiatives and competitive grants. The role of public research centers like NIER would be to provide coordination and expertise. In other words, we need an action scheme, leadership, and coordination to enable faculty, universities, networks of faculties or universities, and academic societies to take primary roles in their reform efforts.



■ **Motohisa Kaneko** : Are there any people who would like to add to the previous comments? Please.

■ **Peter Ewell** : I think that a theme throughout all of our responses is that data of this kind is best used to start

a conversation about improvement. That proposition was not intentionally put on the table as part of the AHELO feasibility study. It's happening now quite nicely, though almost by accident. We should continue those conversations and it's important to recognise that the data necessary to start a good conversation do not need to be as technically sound as the data needed to make an estimate of a country performance. The cost of producing the kind of data that could be used, as in PISA, to make a country estimate is very, very high. The level of precision needed to start a good conversation in a discipline is a good deal less. Starting with regional consortia is a very promising way to proceed, I believe. The difficulty is that the OECD has never run a process like that. The OECD's experience is with large scale testing that produces a point estimate for a country. I therefore think that their institutional culture was to try to do the same thing with AHELO in producing a formative assessment. And it didn't really work for that purpose.

■ **Motohisa Kaneko** : Dr. Fukahori and Dr. Ewell both considered whether or not this in itself was a success. Putting that aside, I think we can rightly say that this is the starting point of various discussions. But this framework itself went through an international organization, and there's a certain inconvenience in the usage of that which poses a constraint. At the moment, we're faced with difficulty even in using the existing data. Bearing in mind that point, what is the least extent to which we can use what we already have and do you plan to do anything further with what you already have in terms of data? Dr. Edwards mentioned that there are initiatives underway in Australia, and Dr. Fukahori also mentioned the initiatives in Japan. But do you at least have any plans to use the collected data to do something in the future? Does anyone like to volunteer?

■ **Satoko Fukahori** : Thank you. If I may, first of all, Ms. Lennon and Dr. Edwards have been invited to the closed meeting of the universities participating in AHELO, which will be held tomorrow. In that meeting, we plan to update ourselves on what's ongoing in each of the countries. Dr. Hamanaka, who is serving as the emcee, will do the data analysis; we are trying to provide as much feedback as we can to the participating universities. By the end of March, NIER's report on the analysis of AHELO will be written and made public. The extent to which we can include data in that paper will be subjected to further study, but we are trying to come up with a document that will be useful to universities when they think about strategies for educational improvement.

■ **Motohisa Kaneko** : Dr. Edwards, you touched upon the plan slightly, but contextual analysis has hardly been done. Any comments on that?

■ **Daniel Edwards** : Yeah, thank you professor Kaneko. We are working with Japan and Canada on a combined data set and we've started to do some analysis looking at the contextual data. The plan, one of our plans, I think we have multiple plans, but one paper that we've been working on at ACER has been around the contextual data, and looking at certain aspects. I mean Satoko did put up some interesting data already about the amount of time spent in work, paid work unrelated to study etc. We've been doing some more of that, and some more looking at what issues make a difference and which issues don't make a difference.

One of the key things and it's not surprising from the result work we've done and the work that Professor Ewell has done, is that when we do a regression analysis and control for a whole lot of variables, the students who say they are quite good in class so they get high marks, are doing better on, perform better on the AHELO assessment. So we know there's some concurrent validity in the AHELO assessment. We also know that, and some of the stuff that I'll talk about tomorrow with the institutions, certainly from the Australian data we can see that up to a certain point, students who are working in paid work that is related to their fields, so they're working as an engineer already, up to a certain number of hours, if they work too much it doesn't work, but up to a certain number of hours, we can see a correlation with that work and their outcome on the AHELO assessment.

So there's that, the other key finding, I think that I've put in my paper is relating to the extent to which students, what they plan to do after graduation. The students who are planning to become an engineer and work in that field did better on AHELO than the students who were not planning on working in that field once they left the university. So this is a start but there is more to be done.

■ **Motohisa Kaneko** : Dr. Ewell?

■ **Peter Ewell** : Daniel and I compared the US results on generic skills and the Australian results on Engineering and, although they are different domains, we found the same general patterns of results. So there is definitely something there. And we both found that the level of self-reported academic ability and the level of self-reported effort can be used as statistical controls because we found that that both were strongly related to outcomes. But at the same time, other variables such as what the student was going to do after graduation and how relevant AHELO questions were to this also showed significant relationships with outcomes after controlling for academic ability and self-reported effort. At the same time, some student-reported curricular experiences were present in our data that even if you regressed out self-reported ability and self-reported effort, also remained significant. As a result, I think that we have some solid foundations to further explore teaching and learning processes as a result of our AHELO results.

■ **Motohisa Kaneko** : Dr. Kishimoto?

■ **Kikuo Kishimoto** : This time around, we not only looked at how much they (the students) were able to score, but we were also able to get the contextual data. What we wanted to do was to look at the correlation between the contextual data and the actual test results; we believe that kind of correlation is very important. When the universities change the way that they educate people, we want to see how the results will be reflected in the change. In order to see that, it would not be enough to do it just one time. Maybe we should wait for several years and take this data again. We need to have this as a continuous process, so we need to continue taking data.

■ **Motohisa Kaneko** : I would like to finish this soon. In Canada, the US, or in the EU, do you see the experience of AHELO linking to a different movement? Or what kind of impact do you think it will have on any activities or movements that are in place or being conducted right now?

■ **Robert Wagenaar** : I can give it a try. Actually not much is happening, at least not globally. But I think there's great awareness among universities that this is important. So the fact that it's accepted as a notion that it's important, also in Europe, although Europe -- as you have noticed has not been very much represented in AHELO, although the OECD is an European organisation from its origin which is amazing. I talked with the engineer organisations for example recently, and they have a strong interest. But I'm not so sure, and I said this before, that the format we are using at the moment is the one which actually will be sustainable. I personally think and others have referred to this so this seems to be common ground that a sustainable basis seems to be a bottom up approach, and not a top down approach. Universities should see the advantages, academics should see the advantages from the very start. They should be involved from the very start in the exercise, as also Peter has mentioned.

The moment that happens, so people see the advantage because they can learn on the spot more or less, from the dialogue, with colleagues, whether it's in a region or worldwide, that we have to see how it develops further then there will be interest. Also students will have an interest, if they can get a personalised let's say outcome with regard to the test they are doing, of course they want to take it, because they can throw it away when it's

not good. If it's good, they can show it to a future employer.

So there are means of strategies to draw attention to this approach. But the fact that we want to compare, because we are comparing as was stated already with regard to research, we also want to compare teaching and learning. It forces us more or less -- at least morally -- to come up with a follow-up. So we have to look for the right framework, and that framework might also be a sort of consortium of universities that actually set this up, which are subsidised by organisations like the OECD -- that's probably asked too much because they don't have money -- governments, the European Commission, organisations like Lumina, or other organisations throughout the world, that have an interest. Because these organisations exist, they exist in Germany, they exist in the United States, they exist in other countries. So that might be a better approach than we have had so far. That doesn't mean that what the OECD has done so far as an initiative has not been of tremendous value, because we would never have found out where we stand today if they wouldn't have done this. So we should be grateful, I think, to all that have been involved, but also to the OCED as an organisation for the work done.

■ **Peter Ewell** : I very much agree with that. We owe the OECD a good deal of gratitude for getting this moving. But whether the OECD is the correct vehicle for sustaining a formative assessment like this, I think is a very open question. Again, methodological purity is not necessarily the goal here. The goal here is to get data that are good enough to support a conversation about improvement. So I think regional consortia of countries is a very good idea. Look at the countries that are not represented here-- countries like Columbia, countries like Egypt, countries like Mexico-- that put a lot of effort into AHELO. These countries got close to 100 percent student cooperation rates, are analysing their data now and, there is a Latin American initiative between Mexico and Columbia under way to share their data. None of this was originally envisioned in the OECD's plan, but it happened. And it happened because the need is there. So my prediction for the future is that we will begin to see small sets of countries getting together to create AHELO-like initiatives of their own, hopefully with support from third party foundations. And all of these should be working from the same kinds of assessment frameworks in the disciplines that they choose to examine, not identical frameworks, but the same kinds of frameworks.

■ **MC Lennon** : Thank you. I think that AHELO is part of a massive movement towards establishing, implementing, and measuring learning outcomes. Not only around the world, but in various levels, both in the government, the system level, in the institutions and also in disciplines as well as programmes. So I think that with the interest of those various levels, the approaches to assessing learning outcomes has the opportunity to be done at those different levels. In the engineering programmes for Canada anyway, we have recently developed graduate attributes, graduate learning outcomes, and the institutions are being mandated to not only implement them but also have them demonstrably assessed by 2016. So they're working very, very hard on trying to figure out these assessment tools amongst themselves, they were very interested in participating in AHELO for that purpose and also looking forward to participating in the conversation about developing the mechanic engineering assessment as well. So we're seeing pockets of innovation come together, but I think that having the opportunity to collaborate internationally, certainly helps forward our conversations and our conceptual thinking about how to do it. So I think that the OCED offers a very vocal and a very prominent platform to have these conversations. But I don't think that it's the only one and I think that there's the opportunity to support smaller innovative activities as well.

■ **Motohisa Kaneko** : The others, do you have anything else to add?

■ **Daniel Edwards** : Just a, I'll try to be, a quick point, but comparing, talking about this bottom up versus top down approach. In the Australian experience, so we've run AHELO with civil engineering and I've been involved

in medical programmes so I've had pretty much the same role in two different programmes. One where I guess government started it through the OECD and we were contracted to do it. And we then said to universities who wants to take part and from there we've had relatively good participation from the institutions but students haven't been in the institutions, have had little bit to do with developing the assessment, but not much. And to be honest, if we tried to run something like this in Australia, we would probably struggle to get this many people. So the commitment isn't there as much.

On the other hand, the medical programme we've done has actually started from institutions contacting our organisation saying, "Hey we want to do something like this. We know that the Australian Council for Education Research has expertise that we don't have. We need someone to facilitate, how about we band together, get a grant and build something" . We started the medical assessment collaboration with two medical schools. We ran a forum about two months later and we had a room full of people like this. And a year and half later we have sixteen medical schools and we didn't have to try, everyone wanted to be part of it, because the way that we've been building the assessment is through workshops with the institutions, the academics themselves are coming, and the clinicians themselves are coming and doing this work. We're developing items that are relevant to their curricula, but also that can be commonly used across schools. And also the institutions feel that they are improving their assessment more generally, because of this involvement. So this is a small example but I think an important one and a successful example of how that bottom up approach can work and now I guess our future aim is to then expand regionally, with this idea but starting small for us on that side of things has been good. And the one last thing that I want to say is about that initiative. In the medical schools we are giving the students a pretty detailed student report, every single student who's offered an opportunity is sitting that test. In engineering, we had to, well we couldn't but we would have dragged them kicking and screaming to get them to sit the test. It was very difficult in most institutions, so again there's completely different outcomes on things that are trying to kind of achieve the same thing, but done differently.

■ **Motohisa Kaneko** : Thank you. It's about time. So let me wrap things up with my personal impression. At the beginning, Prof. Kishimoto said that Japanese educational reform will have to focus on competition. In principle, what can be done at the universities, and what difference do the individual faculty members make in terms of the learning outcomes of the students? They have become specific agenda items for the Japanese education system already. In one way or another, measuring what the students are doing is important, not too ambiguously, too abstractly, but in a more concrete manner. In that sense, AHELO is an ambitious project because it tries to establish an international framework that is measurable. But if we actually embark upon such an effort as exposed this afternoon, many problems may arise, such as the conceptualization of the test and various other issues. But this is not the end, and when I say international, even if not all OECD members can participate, a few countries can volunteer. It would also be meaningful if a few institutions can participate on the domestic level. I think these are some of the lessons learned through our trials and errors. In that sense, I think this conversation was extremely useful as Japan enters into a more concrete phase of measurement and comparison. Again, I apologize if I missed some of the questions that were submitted. However, all the responses were greatly interesting and provocative. So, once again, let us thank the panelists with a round of applause.

Closing Remarks

Tsuyoshi Sugino*
Deputy Director-General, NIER

N.B.

- The * mark indicates that the original language of the speech was Japanese and that the transcript is a tentative translation based on the simultaneous interpretation provided during the symposium.
- The transcripts include changes made after the symposium for purpose of publication.
- The affiliations and professional titles of the speakers are as of December 10, 2013.

Closing Remarks

Deputy Director-General, NIER
Tsuyoshi Sugino



Ladies and gentlemen, thank you very much for your participation. Thank you very much for attending this symposium despite your busy schedules. We come from different universities, private sector organizations, and others. Members of the embassies in Japan have also participated. And we witnessed very interesting presentations. The flow of higher education reform in the last 30 years was shown by Mr. Kimura at the beginning of the symposium. We were able to position this as the base for other symposiums. Prof. Robert Wagenaar presented the latest information regarding specific methods of Tuning, which builds degree programs based on competence frameworks of specialty areas, and the relationship of cross-discipline efforts and learning outcomes studies. Dr. Peter Ewell shared the overall picture of AHELO, explained its results and issues, and described the future of the main study in a way that was very easy to understand. Prof. Kikuo Kishimoto presented the specific efforts made in Japan in the engineering domain of AHELO. It made us acutely aware of the fact that the mere participation in international learning outcomes studies is not enough. There is a need to participate at the development stage of the questions to gain sufficient outcomes. Dr. Satoko Fukahori, senior researcher of NIER, reported the analysis of the study results of the three countries – Canada, Australia and Japan – on which she has worked together with Ms. Mary Catharine Lennon and Dr. Daniel Edwards.

This time it was a feasibility study. However, the participatory work done in these three countries is extremely meaningful. The panel discussions after the keynote lectures and presentations were quite frank, and very valuable exchanges were conducted. I would like to once again show my appreciation to our facilitator, Prof. Motohisa Kaneko, who made the discussion a very fruitful one by effectively organizing the diverse points of discussion. We can say that Tuning-AHELO provided an important insight into Japan's higher education policies and the practices of university education. As Prof. Kimura said, higher education is the weak point of Japan. However, this symposium opened the opportunity to undertake university education improvement reforms with a deeper insight. I would like to thank all the participants for attending this symposium despite your busy schedules. Thank you for your kind attention.

Appendix

Reference Paper

The AHELO experience – implementation, outcomes and learning from an Australian perspective

Daniel Edwards

Principal Research Fellow, Australian Council for Educational Research

AHELO: The experience of Ontario's institutions

Mary Catharine Lennon

Senior Research Analyst, Higher Education Quality Council of Ontario

Presenters' Biographies

The AHELO experience – implementation, outcomes and learning from an Australian perspective

Dr Daniel Edwards

Principal Research Fellow, Australian Council for Educational Research and National Project Manager, Australia for AHELO.

Introduction

This paper provides background, implementation processes and outcomes relating to Australia's participation in the Organisation for Economic Cooperation and Development's (OECD) Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study. The Federal Department of Innovation, Industry, Science, Research and Tertiary Education (DIISRTE) funded the Australian participation in this project via direct contribution to the OECD, and through funding of a National Centre to implement the study in Australian universities.

The AHELO Feasibility Study was a major OECD project. Its objective was to determine a robust approach to measuring learning outcomes in ways that are valid across cultures and languages, and across the diversity of institutional settings and missions. Further detail including aims and final reports relating to the project can be found at <http://www.oecd.org/edu/ahelo>.

The AHELO Feasibility Study involved the testing of students in three strands – Generic skills, Economics and Civil Engineering. For the Feasibility Study, Australia participated in the Civil Engineering strand and it is participation in this strand that is the focus of this chapter. In addition to these assessments, contextual dimension questionnaires were deployed to students, faculty and institutional staff at all participating institutions.

The implementation of the Feasibility was split into two phases. Phase One was undertaken in 2011 and involved engaging institutions, coordinating focus groups in participating universities; undertaking an academic review of draft assessment instruments; marking of student responses; and the collation and communication of feedback from institutions and students. Phase Two took place in 2012 and involved re-recruiting universities, sampling of students and staff, technical assistance in preparation for AHELO assessments, administration of the assessments and questionnaires, scoring of test scripts and reporting of results.

This paper offers insight into the processes, activities, outcomes and dissemination of results from the AHELO Feasibility Study in Australia with particular emphasis on the student-related data. It begins by providing a brief background about the concept of AHELO and the practicalities relating to the project. It then examines the process of implementing the study in Australia, beginning with Phase One and the collection of feedback relating to the draft instruments. The implementation of the AHELO assessment (Phase Two) is then detailed, alongside some insight into participation numbers and some notable but broad outcomes reported. The final section of the paper discusses overall outcomes of the Feasibility Study from an Australian perspective and considers some possibilities for alternative approaches to dissemination and reporting.

AHELO Background

This section provides a background to the AHELO project. It was jointly prepared with members of the AHELO Consortium and provides insight into the aims of the study. It is included here to provide an overall perspective for framing the work detailed in this paper. The overall aims and ambitions of the Feasibility Study are important and are sometimes lost in the detailed discussions of process in this work, so this section offers the chance for reflecting on the 'bigger picture' of what was trying to be achieved.

Is it possible to undertake an international assessment of final-year students' capacity to use, apply and act on the knowledge and reasoning they have gained from their degrees? Is it possible to assess these outcomes in an efficient

and internationally comparable way? Can policymakers, institutional leaders, faculty and students be convinced that the assessment of higher education learning outcomes is an essential checkpoint in the educational process? These questions lie at the core of the OECD's AHELO Feasibility Study.

AHELO involves the development and validation of assessments in three core areas – Generic Skills, Economics and Civil Engineering, as well as the development of contextual instruments to aid with the interpretation of assessment data. The assessments are targeted at students in the final year of bachelor degrees and aim to assess their capacity to apply their skills and knowledge to real-world problems. AHELO is an ambitious project in that it aims to develop new methodologies and technical standards for the assessment of higher education learning outcomes. It is taking place on a global scale, with 17 countries participating in the development and validation of assessments, and engagement from experts, institutions, governments, and key higher education bodies from around the world. The project is run by a consortium of international organisations, led by the Australian Council for Educational Research.

AHELO responds to a critical information gap. Efforts to improve the quality of teaching and to enhance students' learning outcomes are stymied by the absence of reliable information which enables comparative judgments to be made about the capabilities of students in different institutions and in different countries, or about the quality of teaching. In the absence of such data on core higher education activities of learning and teaching, the standing of a higher education institution is based largely on reputation and research performance. AHELO's objective is to create a richer source of information through designing and testing measures which give due weight to teaching practices and learning outcomes.

The AHELO assessments were developed through the collaboration of experts from around the world and subsequently translated and adapted for use in all participating countries. Validation of the assessments with students in participating countries was completed in 2011. The assessments that were developed through for AHELO were implemented across 17 countries in 2012.

The study has had an impact on the way in which discussions about higher education and assessment are carried out, suggesting that AHELO has the potential to reshape the higher education landscape in important ways. First, discussions among some stakeholders appear to have moved on from whether learning outcomes should be measured to how they can be measured. At the same time, the engagement of national systems and institutions (both within and outside the OECD) has increased significantly throughout the life of the Feasibility Study. Taken together, these indicate the existence of a desire for data on the quality of teaching and learning, both to inform improvements and also to demonstrate quality.

Facilitating AHELO in Australia

Australia's participation in AHELO was funded and supported by the Federal Government. The Government facilitated the organisation of a National Centre which included a National Project Manager (NPM), some basic support and stipend for travel. The NPM in Australia was essentially responsible for implementing all aspects of the AHELO Feasibility Study, under the direction of the AHELO Consortium and the OECD Secretariat for the study. Many of the key activities of the NPM and National Centre are detailed in the sections of the paper dealing with activities in Phase One and Two. One aspect of the work not covered as much in these latter sections is the element of communication required through the duration of the project. This aspect is discussed briefly here.

The NPM maintained close and frequent communication with the Federal Government through the Department of Innovation, Industry, Science, Research and Tertiary Education (DIISRTE) throughout the Feasibility Study, providing monthly progress reports during the busy times in the project and having regular teleconferences. The progress reports and discussions helped keep the Department aware of the day-to-day activities of the NPM as well as providing insight into the extent to which the project was being embraced by universities, university leaders, academics and students.

The National Centre also played a critical role in developing relationships and communication with participating institutions. Contact and collaboration with institutions was maintained during the study through teleconferences, a symposium, site visits and email. The key institutional relationship maintained by the NPM was with each Institution Coordinator, who was in charge of activities within their university. Building personal connections with these people was an important and highly valuable facet of the project.

Another key role of the NPM that is not explored as much in the other sections of this paper is that this role provided a channel through which Australian national interests are represented in the implementation of the AHELO Feasibility Study. The Australian Government were interested in ensuring that there was important Australian input into the development and adaptation of relevant test instruments and in the processes and procedures used in data collection.

Implementation of AHELO

This section of the paper provides an overview of the AHELO Feasibility Study as a whole from the perspective of the in-country implementation of the study. Implementation within countries and institutions participating in the Feasibility Study was carried out in two phases. Phase One involved qualitative testing of items through surveys and focus groups and providing feedback to the OECD and AHELO Consortium relating to the instruments constructed for the study. Phase Two involved the implementation of the AHELO assessments in universities and with students under test conditions as well as the completion of staff and institution-level surveys.

During Phase One of the Feasibility Study, qualitative testing of the draft assessments was undertaken in a number of countries. During focus groups and cognitive labs, students provided feedback on the assessments, while a number of senior academics in the area of civil engineering were asked to provide comment and review of the draft AHELO assessments. In summary, the key facets of Phase One were:

- *University engagement and recruitment* – securing the interest and participation of universities.
- *Support for participating universities* – providing detailed information to universities about participation, including, a Manual, discussions and training for Institutional Coordinator and ongoing contact and support.
- *Implementation of focus groups* – carried out at each university and facilitated by the Institutional Coordinator. Involved students taking a sample of the draft assessment and providing feedback through a survey and focus group.
- *Academic Review of AHELO instruments* – by experts in civil engineering education.
- *Collation of test results and feedback from students and academics* – organised by NPM and included marking of test papers, collation and coding of surveys and construction of a national database.
- *Providing feedback on test instruments* – involving analysis of responses and collation of key points for feedback to the AHELO Consortium.

Phase One activities in Australia were undertaken in 2011 and involved ten universities (further detail is contained in the ‘outcomes’ section of the paper).

Phase Two involved the full implementation of the AHELO assessment in universities to cohorts of students online under test conditions. It also involved the collection of survey data from academic staff and Institution Coordinators.

In summary, Phase Two involved the following activities for Australian participation:

- *Institutional recruitment* – re-securing the interest and participation of universities.
- *Development of manuals for institutions* – to assist in implementation of the online testing and in recruiting of students.
- *Institutional Coordinator training and support* – to ensure all participating institutions were familiar with the AHELO test system and AHELO test administration protocol.
- *Institutional sampling* – collecting population data on eligible students and staff members in participating

institutions.

- *Setting up test computers* – providing advice and information on establishing secure online profiles for testing centres.
- *Test administration* – involving the implementation of the AHELO tests to students in participating universities and monitoring the administrations to confirm test sessions were run according to protocol.
- *Survey administration* – sending the staff and institutional survey to identified population and following up to secure maximum response numbers.
- *Scoring of assessment responses* – involving recruitment of a Lead Scorer, training in scoring and facilitating scoring among a group of experts.
- *Supporting data processing and international analyses* – ensuring adherence to AHELO protocol and maximising international comparability of data.
- *Attend international training and dissemination meetings* – maximising the value of lessons drawn in the Australian context as well as benefitting from insights in other countries.

Phase Two was conducted in Australia during 2012, with testing being undertaken between March and May in eight universities (details in ‘outcomes’ section).

Recruiting Institutions

Based on requests from the OECD and AHELO Consortium, Australia aimed to have ten universities participate in the Feasibility Study. As such, early in 2011, the Vice-Chancellor and Dean of Engineering at ten Australian universities were written to and invited to participate. All ten invited institutions agreed to be part of at least Phase One of the study.

Selection of universities involved liaison between the NPM, the Department (DIISRTE) the Australian Council of Engineering Deans (ACED). Universities were selected on the following basis (in order of selection priority):

1. They had actively shown prior interest in AHELO (through informal discussion with the Government).
2. They had previously provided contact details for an Institutional Coordinator to DEEWR.
3. They had an active Civil Engineering program.
4. The ‘sample’ included universities from a range of geographic locations.
5. The ‘sample’ included universities from a range of different ‘groupings’ of institution (i.e. research intensive, ‘new’ universities, regional universities etc.)
6. The ‘sample’ included universities of various sizes.

As each university confirmed participation in the Feasibility Study, they were contacted by the NPM and provided with an overview of the project. Institutional Coordinators were identified and directly contacted by the NPM to provide details on involvement in the project.

Across the two years of the project, students from 11 different Australian universities were involved. These institutions are listed below. Institutions with one asterisk (*) participated only in Phase One, these institutions generally cited resourcing issues as the key impediment to further participation. The one with two asterisks (**) participated only in Phase Two following an invitation from the NPM in early 2012. All other institutions listed were involved in both phases of the Feasibility Study.

- Charles Darwin University
- Curtin University of Technology*
- James Cook University
- Swinburne University of Technology
- The University of Adelaide
- The University of Melbourne

- The University of New South Wales*
- The University of Newcastle*
- University of Technology Sydney
- University of Western Sydney
- RMIT University**

Outcomes

In this section the broad outcomes from students in Australian institutions participating in AHELO are discussed. The discussion is relatively broad and focuses more on the participation levels than it does on the actual ‘results’ from the AHELO assessments.

Qualitative testing – Phase One

Phase One was carried out in ten participating institutions in Australia in April and in August 2011. In total 78 students took part in focus groups in which they were asked to take some sections of the draft assessment and then complete a survey and provide feedback in a focus group setting. The draft assessments included in this session were:

- a Constructed Response task - where students are presented with a scenario or problem and asked a number of questions in relation to this. In most cases these required ‘open answer’ responses (i.e. they were not multiple choice); and
- a set of Multiple Choice questions.

All sessions were run within participating universities by the Institution Coordinator or a designated member of faculty. Sessions lasted approximately one hour and were undertaken by final year Bachelor of Civil Engineering students. Table 1 displays the demographic characteristics of the students who participated in the focus groups for Phase One.

After taking part in the test, students were asked to focus on the following topics during the moderated discussion:

- The task challenged me to think (*challenge*)
- The materials stimulated my interest in the task (*interest*)

Table 1: Focus group demographics

Demographics		(n)	%
Gender	Male	60	76.9
	Female	18	23.1
Age	20	4	5.1
	21	18	23.1
	22	26	33.3
	23	12	15.4
	24	9	11.5
	25	3	3.8
	26-30	3	3.8
	31 and above	2	2.6
Citizenship	Australia	54	69.2
	Other country	23	29.5
Language background	English	52	66.7
	Other language	25	32.1
Total students		78	100

- The task made me apply knowledge and skills in real-world ways (*apply*)
- The task assessed an appropriate range of knowledge and skills (*range*)
- The task was relevant to my program of study (*program*)
- The task was relevant to future professional practice (*future*)

The responses of students were scored and input into a national dataset for Phase One that was provided to the AHELO Consortium. At the national level, the National Centre undertook some analyses of the student responses.

The main outcomes derived from Phase One in Australia was that students rated the test mostly positively on the criteria of challenge, interest, application, range, relevance to program and future professional practice. The Constructed Response tasks were rated slightly higher on these criteria compared to the Multiple Choice questions.

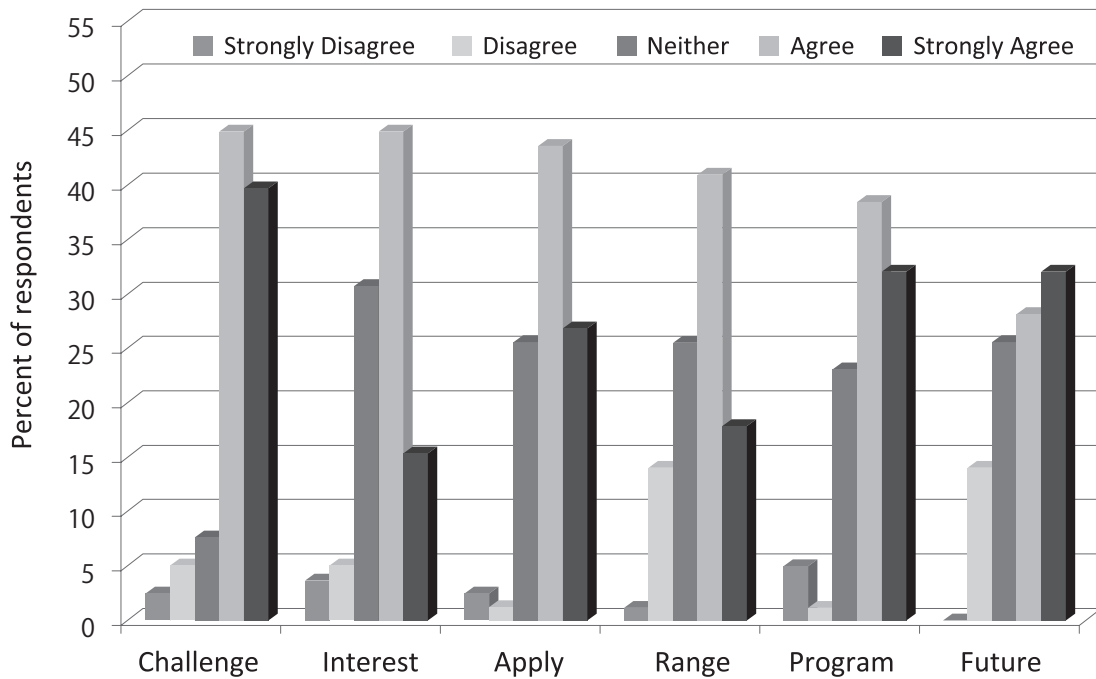


Figure 1 : Feedback on Constructed Response Tasks

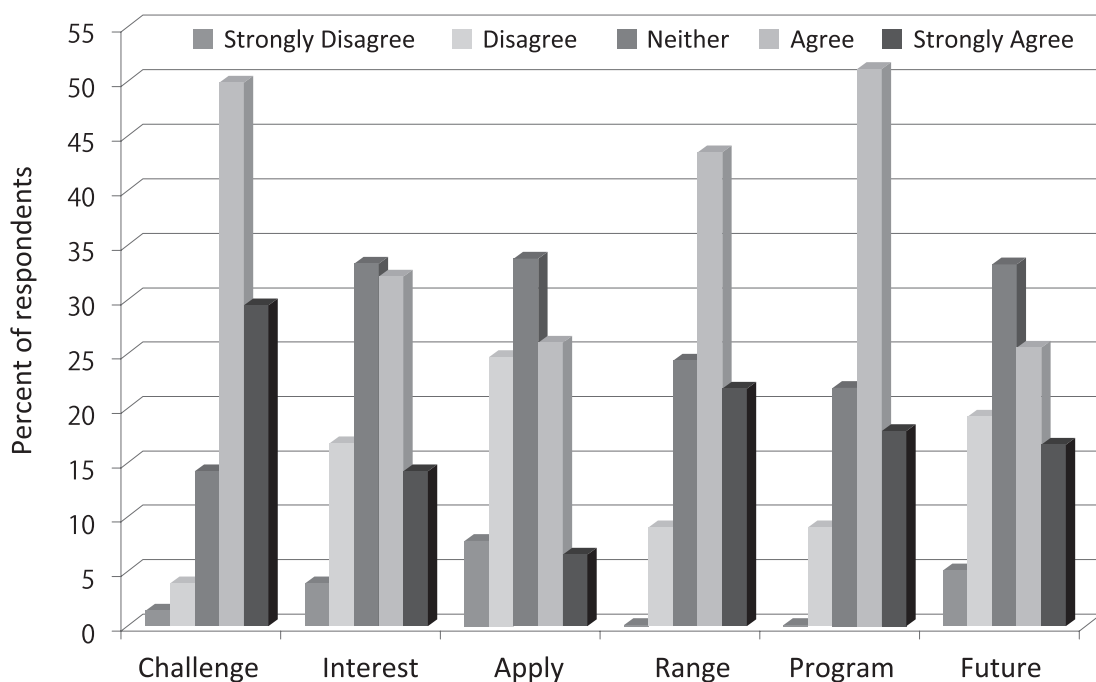


Figure 2 : Feedback on Multiple Choice Tasks

An overview of the Australian student responses to survey questions about the AHELO assessment instruments is provided in Figure 1 and Figure 2 below. They show that students were highly challenged by the Constructed Response tasks and were also generally positive about the other specific areas explored through the survey – with relevance to program and relevance to their future work being of particular interest (Figure 1). Levels of challenge and relevance to the degree program for these students was also relatively good for the Multiple Choice items, while the detail in Figure 2 suggests interests, application and the future relevance of these items were given lower scores by participating students.

In addition to rating the test through a questionnaire, the student focus groups also entailed a moderated survey format. In brief, the major responses of these moderated discussions were:

- When asked about the Constructed Response tasks, in general students found these made sense and the instructions were adequate. Overwhelmingly students found the task interesting, because of the diagrams, real-world problems, and because the test was different to how we students are tested usually.
- Student opinion about the Multiple Choice tasks was generally that these were more like the tasks they were familiar with at university. Students found the tasks to be more recall oriented and while relevant to their degree seemed less relevant in the long term for their future career.

A review of the draft assessment instruments was also undertaken by three academics in Civil Engineering from participating Australian universities. The outcomes of the academic review can be summarised as follows:

- In general, the feedback was similar to that of students.
- There was some suggestion that the content was focused on the early stages of degree, rather than final years.
- There was more specific item-based feedback relating to technical explanations, language used etc., suggested for change in order to make some items clearer.

Based on the results from the student focus groups and academic review both in Australia and in other participating countries, a number of changes were made by the AHELO Consortium to the instruments used in the Civil Engineering strand. A major outcome at this point in the study was that more time was allocated for the test than originally planned. One Constructed Response task was completely removed from the test since it was too difficult. Other Constructed Response tasks were amended.

Quantitative testing – Phase Two

Phase Two of AHELO was conducted in 8 Australian universities in 2012. Students participating in the Feasibility Study undertook the AHELO Civil Engineering assessment, which involved a 90 minute test including one Constructed Response task and a module of Multiple Choice Questions. Students also completed a context questionnaire at the end of the test session which collected data relating to demographics and other background characteristics of students.

Sampling and student recruitment

A major task in preparing for test administration was the collection of sampling frames from each participating university in accordance with the AHELO protocol. To ensure international comparability of the AHELO data and to facilitate processing of the data collected, all participating institutions were required to provide lists of all eligible students and staff members, in a pre-defined table format. Internationally, this format enabled the validation of target populations and where necessary the selection of a sample. In Australia, Civil Engineering faculty sizes were sufficiently small to allow for a census rather than a sample. Eligibility for being included in the test population was defined as students enrolled in civil engineering programmes in their final year before Bachelor completion.

The provision of sampling frames to the NPM was a significant effort on the part of the institutional coordinators, since it required up-to-date information on the enrolment status of students and a range of demographic

characteristics. Institutions were invited to add any background indicators to their sampling frames they deemed enriching to the data, and some provided information such as the programme students were enrolled in, country of birth of the student, or student aptitude measures such as grade point average. Other institutions were more constrained by regulations of data sharing, and were only able to provide essentials such as gender.

Test administration took place during April and May 2012 in all institutions. For the administration of the student assessment and questionnaire, most institutions organised a number of sessions, while other smaller institutions organised a single session. In some institutions, the institutional coordinator supervised all test administration sessions, whereas others hired test administration assistants financially supported through the NPM budget. During test sessions, the NPM was available for support regarding the AHELO test system and any other queries regarding the AHELO protocol.

A major challenge for the Australian implementation of AHELO was motivating students to participate in the study. This was a problem also highlighted by some other countries during the international NPM meeting in October 2011. Given the variety of institutions participating in Australia, the NPM took a flexible approach to incentivising participation, allowing institution coordinators to propose localised incentive ideas that were agreed to and funded by the NPM. Table 2 details the kinds of incentives offered by institutions to participating students. It shows a range of different options and approaches offered. In some institutions, full student population numbers were small enough to offer a voucher to each participant, in others it was necessary to have a draw of prizes for participants. Besides monetary incentives, students were motivated to participate by explaining to them the relevance of the test to their future work experience, and the value of their participation to the success of the study. One university chose not to offer any monetary or 'prize' incentives for students and instead focus solely on the experience of the assessment as being an incentive for participants.

All institutions were provided with participation certificates to provide to students who sat the test. These certificates were developed by the AHELO Consortium and the OECD and included the OECD logo.

Table 2 : Student participation incentives

Institution	Incentive
Uni A	\$AU100 voucher for each participant
Uni B	ipad draw. Lunch and discussion of Outcomes Assessments for participants
Uni C	\$AU50 voucher for participating students
Uni D	ipad draw. Voucher draw in each room (8 different sessions) chance in each room to win \$AU200 or a \$AU100 voucher.
Uni E	2 ipads drawn, 5 \$AU50 vouchers drawn, a lunch
Uni F	\$AU1000 prize draw, and offering student society \$AU1000 for a high overall participation rate (75% - not achieved).
Uni G	none
Uni H	Made test a 'requirement' for a core subject.

Table 3 provides the participation numbers and participation rates (based on the numbers of eligible participants identified through the selection of a sampling frame) for both students and staff for each university. In total 187 students from Australian institutions participated in the AHELO assessment. Across the institutions involved, one in five final year civil engineering students participated in the assessment. However, this average does not adequately represent the notable variation in participation rates of students between institutions. Five of the eight institutions had participation rates below 25 per cent, while one institution (Uni H) had almost full participation from its cohort.

Numerous issues affected the participation rates of students in Australia. These included students having limited 'on-campus' time in their course due to final year internships, a large proportion of students in part-time work and therefore little or no free time, and short notice of the testing window.

Table 3 : Participation numbers and participation rates of students, Australian institutions for AHLEO

Institution	Number of participants	Participation rate* (%)
Uni A	3	30.0
Uni B	27	12.3
Uni C	4	12.5
Uni D	29	19.3
Uni E	27	29.0
Uni F	25	14.2
Uni G	19	12.8
Uni H	53	98.1
Total	187	21.1

*The participation rate is based on the number of participating students as a proportion of the full student population identified in the sample frames constructed by institutions for the study.

However, as the table shows, the Uni H was able to secure a large participation rate, proving that it is possible within Australian higher education to achieve wider engagement of the student cohort in this kind of exercise. Given the success of this institution in securing this level of engagement from students, further information of their approach is detailed in the box below.

Successful engagement model

The approach taken by the Uni H offers an insight into an ideal implementation model. The approach involved strong and dispersed leadership within the institution, insightful planning, and the merging of the assessment into a dialogue about learning outcomes between students and staff.

The implementation of AHELO at this institution was planned in advance, beginning in November the year prior to testing – four months before testing – following an information session facilitated by the National Project Manager. Leadership in the project came at three levels within the institution – from the Chancellery, where a member of the Pro-Vice Chancellor’s office was directly involved, from the Head of the School of Engineering and from a senior faculty member who had ultimate responsibility for implementation. During planning for the testing, the team within the institution identified a core unit (a project/thesis unit) in its final year civil engineering program in which the theme of learning outcomes and graduate capabilities was appropriate. The unit was planned in such a way that AHELO became an integral part of the work during the semester. Students were asked to undertake the assessment as part of the unit of study and following the assessment, were involved in discussions to reflect on their assessment experience and the relationship between their coursework, the skills they expected to employ in the workforce following graduation and professional responsibility related to assurance of educational and practice standards.

Participant characteristics

The characteristics of the student participants in AHELO are displayed in Table 4. Notable student demographics include the dominance of male students in engineering, and the high proportion of students with a main language other than English. The average age of students participating was 23.2. Most students were studying full-time and on campus. A basic comparison of the sample frame with the secured participation group suggests that by gender the sample is representative, although it appears that the participant age may be slightly older on average than the full population.

Table 4 : Student participant characteristics

Characteristic	Category	Participant population (% of total)
Gender	Male	80.1
	Female	19.9
Age	20 and below	5.3
	21	20.5
	22	30.4
	23	19.3
	24	9.4
	25-30	11
	Above 30	4
Enrolment type	Mainly part time	4.1
	Mainly full time	95.9
Enrolment mode	Entirely on-campus	91.2
	Mix of external/distance and on-campus	8.8
Country of birth	Australia	53.8
	China	14.6
	India	2.9
	Malaysia	5.8
	Other	22.8

Assessment outcomes

Discussion relating to the data outcomes from AHELO participation by Australia is limited in this paper due to the fact that this data is relatively sensitive and the sample numbers and representativeness of students for most of the universities involved in Australia was below the level that would be required for the collection of reliable data. Instead, some relatively broad level findings are shown here using the full Australian data set. Two issues are discussed – relationship between student grades and outcome on AHELO, and the link between work ‘practice’ and outcomes in AHELO. For broad reference, in examining the scores below, the international mean for AHELO is 500 with a standard deviation of 100.

Figure 3 shows the relationship between student outcomes on the AHELO assessment and their grades during university study – self-reported by students. It clearly shows that the Australian students who identified themselves as being among the top in their class achieved a much higher average score in AHELO than other students. This goes some way to suggesting there is some concurrent validity present in the AHELO instruments.

An important issue in Australian Higher Education is the link between study and work and the extent to which study is preparing young people for their future careers. The outcomes from Phase One displayed earlier suggested that at least in terms of the content of the tests, students saw substantial relevance with future work in the Constructed Response tasks. Interestingly, the AHELO data for Australia shows that those students who are engaged in paid work relating to their degree in their final year of university, and those who are planning to work in the industry related to their study when they graduate are much more likely to have higher scores than students not currently and not planning on working in their field in the future.

Figure 4 shows that students who worked at least six hours – but less than 31 hours – per week in a paid job that was related to their study on average had a higher outcome on the AHELO assessment than those not working in a paid job in a related field or those working more than 30 hours in such employment. This is an interesting finding in

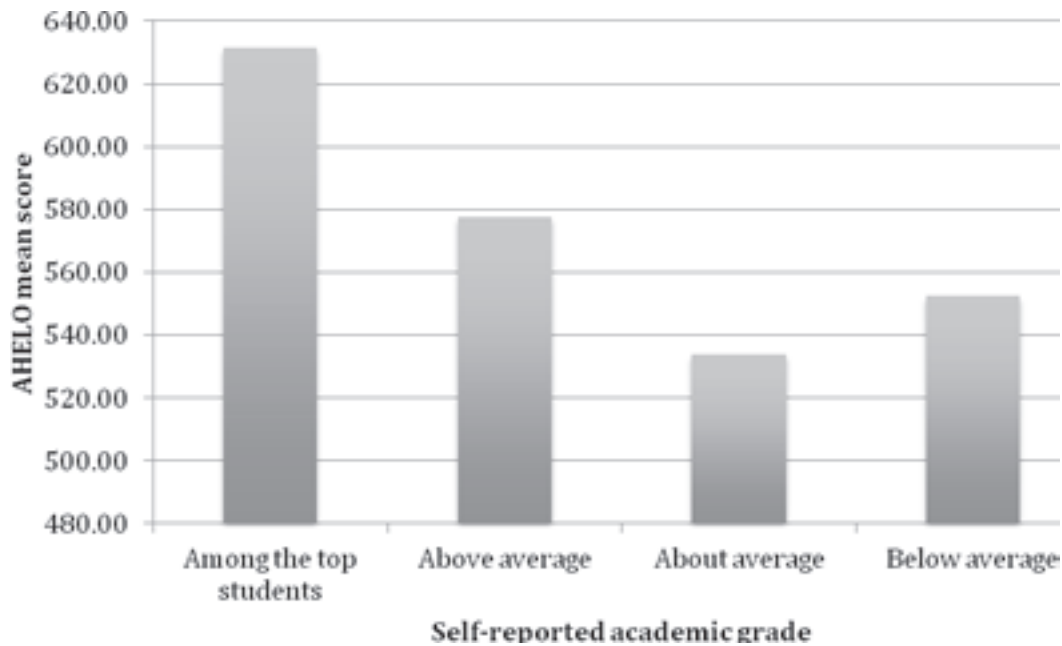


Figure 3 : AHELO mean score by self-reported outcome in university studies, Australian participants

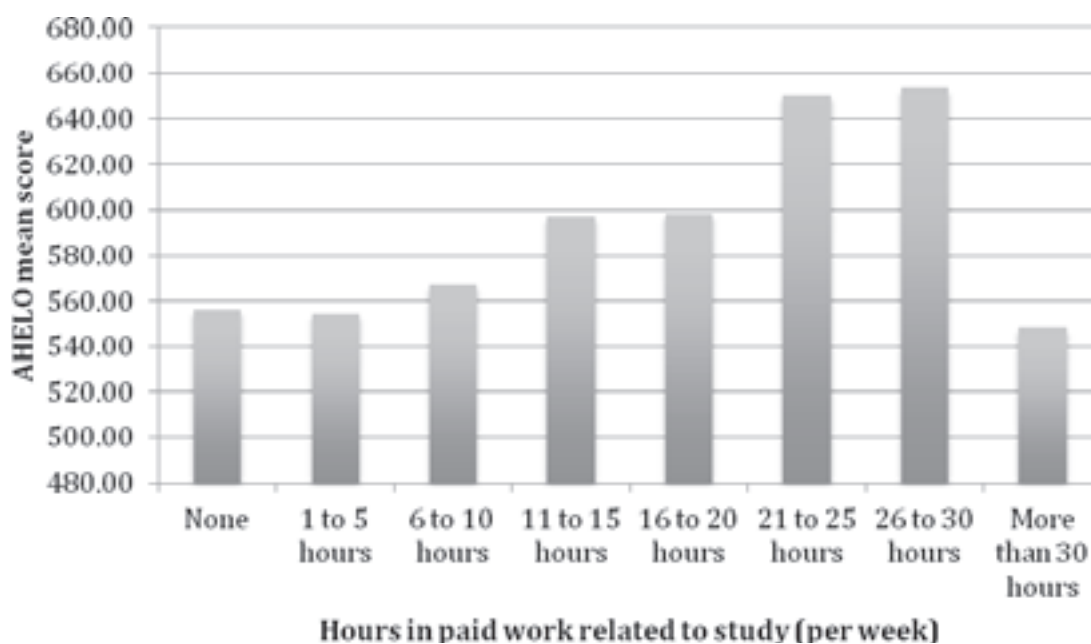


Figure 4 : AHELO mean score by number of paid hours worked in job related to study per week, Australian participants

the context of considering the extent to skills gathered through working in the field might help influence the extent to which students were able to successfully answer questions in the AHELO assessment. It also suggests that for those students with a very heavy work burden during their final year of study (paid work of more than 30 hours), this paid work may be having an adverse impact on their ability to perform at university.

The other work-study related issue that shows some interesting data among the Australian participants in AHELO relates to the post-graduation ambitions of students. The data in Figure 5 indicate strongly that those students who had plans to pursue a career in Civil Engineering straight after completing their degree were clearly more likely to have a higher score on the AHELO assessment. It is likely that these are more motivated students at this time in their degree and perhaps more focussed than those students with no plan to pursue work in the field. Interestingly, those students intending on pursuing further study in the area had a higher mean score than those not planning on working in the field, but lower than those intending on entering the Civil Engineering workforce.

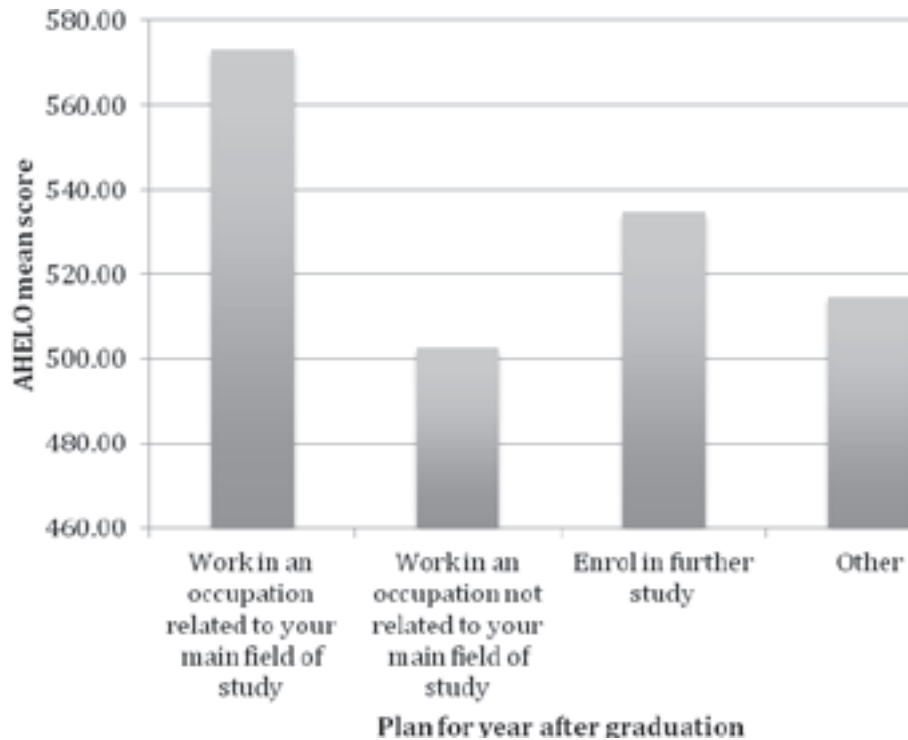


Figure 5 : AHELO mean score by pathway planned by student following graduation, Australian participants

Conclusion and ideas for the future

The Australian participation in the AHELO feasibility study has resulted in a number of valuable lessons. Firstly, it has shown that the Australian sector is equipped to participate in this type of international study. There was interest from stakeholders in participating, and administrative systems in institutions allowed relatively straightforward production of the required student and staff data. The process of a National Centre overseeing a network of institutional coordinators functioned well, and coordinators report finding this type of cooperation valuable.

Secondly, participation in the process of developing and trialling an internationally applicable engineering test has proven to be insightful both to participating institutions and students. AHELO has proven to be a truly groundbreaking test in the sense that students reported not having been exposed previously to the type of applied, integrated problems mirroring their future work. The cooperation between engineering experts from multiple countries both in meetings and via online communication has provided an excellent opportunity to strengthen international bonds. In addition the cooperation between participating institutions opened up opportunities for future cooperation within engineering faculties.

In practical terms, AHELO has shown that students in Australia are not easily motivated to participate in a voluntary test or questionnaire. Future iterations of AHELO or similar studies would do well to integrate as much as possible these types of assessment into the curriculum, using implementation models such as that used by Uni H. Though participation rates in the feasibility study for Australia were disappointing, the process of implementation has built substantial knowledge on the processes and systems needed for engagement among students and institutions in future studies of this kind.

AHELO has provided some small insights into the current state of engineering bachelor education at selected universities in Australia. While this data does not allow any strong conclusions about the skill level of Australian students as it compares to that of students in other countries (and therefore reduces the outcomes analysis of this paper), it allows a glimpse into what could be possible in future iterations of the study. Importantly, involvement in the AHELO Feasibility Study has provided Australia with valuable lessons and models for future implementation of

such assessments in the future.

For Australia, there are some worthwhile considerations for future international participation that have become apparent through the feasibility study. One is that in the testing window used in this phase, Australian students were technically one semester behind those from institutions in other countries (except Japan). The test window fell in Semester 1 of the final year of the Australian students, but in the final semester of study for those in most other participating countries. For accurate international comparisons, future iterations of the study should be implemented in comparable times during academic years across all countries. A second is that a number of the institutions involved in the Australian participation have substantially large final year internships or research projects in the final year, meaning that students spend significant time in this year off-campus. As such, being able to find a time in which a large cohort is able to participate in a secure assessment is difficult for institutions. Longer term planning for the running of such assessment could help in minimising the impact of these key events in the final year. However, the importance of the flexibility of internships and research projects may make this a challenge to achieve. A final issue recommended for consideration in the future is the production of student-level reports for individuals who participate. It is recognised that this was beyond the scope of the feasibility study, but Australia believes building such capabilities in the future would help to stimulate engagement of students in these types of studies.

Since the conclusion of the Feasibility Study, the Australian NPM has worked with colleagues on identifying useful ways of developing reports for institutions within the constraints of the study (in particular the fact that the assessment items are still secure and confidential). Some innovative work has been undertaken in this regard and an example institution level and national level summary report are presented at the end of this paper. These reports show outcomes at the individual item level, order items by level of difficulty, and provide a reference to the type of capability being tested by each item. The example here is a useful one in consideration of developing such reporting in the future. It is the opinion of the author of this paper that greater momentum following the implementation of the Feasibility Study would have been gained by having clearer and more detailed reporting at the student and institution level.

Acknowledgements

Thanks to colleagues and members of the ACER Consortium for advice and content for this paper and assistance throughout the Feasibility Study Eva van der Brugge, Sarah Richardson, Hamish Coates, Jacob Pearce, Yan Bibby and Xiaoxun Sun.

EXAMPLE OF FORMAT FOR REVISED INSTITUTIONAL REPORT – PATTERNS OF STUDENT PERFORMANCE

Civil Engineering Learning Outcomes Item Code Description

Engineering Generic Skills	
A	use diverse methods to communicate effectively with the engineering community and with society at large
B	demonstrate awareness of the wider multidisciplinary context of engineering
Basic and Engineering Sciences	
C	demonstrate knowledge and understanding of the scientific and mathematical principles underlying civil engineering
D	demonstrate a systematic understanding of the key aspects and concepts of their branch of engineering
E	demonstrate comprehensive knowledge of materials and construction
F	demonstrate comprehensive knowledge of structural engineering
G	demonstrate comprehensive knowledge of geotechnical engineering
H	demonstrate comprehensive knowledge of hydraulic engineering
I	demonstrate comprehensive knowledge of urban and rural planning.
Engineering Analysis	
J	apply knowledge and understanding to identify, formulate and solve engineering problems using established methods
K	apply knowledge and understanding to analyse engineering products, processes and methods
L	select and apply relevant analytic and modelling methods
M	conduct searches of literature, and to use databases and other sources of information
N	design and conduct appropriate experiments, interpret the data and draw conclusions
O	demonstrate workshop and laboratory skills
Engineering Design	
P	apply knowledge and understanding to develop designs to meet defined and specified requirements
Q	demonstrate an understanding of design methodologies, and an ability to use them
Engineering Practice	
R	select and use appropriate materials, equipment and tools
S	combine theory and practice to solve engineering problems
T	demonstrate understanding of applicable techniques and methods, and their limitations
U	demonstrate understanding of the non-technical implications of engineering practice
V	demonstrate understanding of the health, safety and legal issues and responsibilities of engineering practice
W	demonstrate knowledge of project management and business practices

Correct Incorrect
 Not Administered

Set*	Item	Item description code	AHELO Sample		Australian Sample		This Institution		% Correct	
Set 1	ENGMQ3	G	54	81	72	1	1	1	1	0
	ENGMQ1	C	51	47	49	1	1	1	1	1
	ENGMQ2	E	35	36	16	0	0	1	1	0
	ENGMQ5	H	20	18	14	0	0	0	0	1
Set 2	ENGMQ7	C	61	57	57	1	1	1	1	0
	ENGMQ8	F	54	50	36	1	1	0	0	0
	ENGMQ9	F	54	64	59	1	1	1	1	1
	ENGMQ10	G	50	51	27	1	1	0	0	0
Set 3	ENGMQ6	C	47	55	41	1	1	1	1	0
	ENGMQ3	C	55	66	53	1	1	1	1	1
	ENGMQ14	E	33	30	19	0	0	0	0	0
	ENGMQ11	H	28	34	14	1	1	0	0	0
Set 4	ENGMQ12	H	26	24	16	0	0	0	0	0
	ENGMQ15	E	22	25	16	0	0	0	0	0
	ENGMQ16	F	59	57	40	1	1	1	1	0
	ENGMQ20	I	45	57	47	1	1	0	0	0
Set 5	ENGMQ18	F	40	41	44	1	0	1	1	0
	ENGMQ17	G	25	31	20	0	0	0	0	0
	ENGMQ19	H	18	14	16	0	0	0	0	0
	ENGMQ22	F	48	49	26	1	1	1	1	0
Set 6	ENGMQ23	G	38	42	39	0	0	0	0	0
	ENGMQ25	I	32	51	52	1	1	1	1	0
	ENGMQ21	C	27	30	13	1	0	0	0	0
	ENGMQ24	H	24	32	20	0	0	0	0	0
ENGMQ29	F	63	70	61	1	1	1	1	1	
ENGMQ27	F	31	31	27	1	0	0	0	0	
ENGMQ30	H	16	22	9	0	0	0	0	0	
ENGMQ28	I	6	6	14	0	0	0	0	0	

*Each student was administered 5 out of 6 item sets.

EXAMPLE OF FORMAT FOR REVISED NATIONAL SUMMARY REPORT

AHELO Feasibility Study 2012

Australian Institutions

Civil Engineering Learning Outcomes

Constructed Response Tasks

 Partial Credit Item Scores

 Not Administered

Task	Item	Item description code	Average Raw Score									
			AHELO Sample	Australian Sample	Institution 3	Institution 4	Institution 9	Institution 6	Institution 2	Institution 5	Institution 10	Institution 1
Task 1 Dams	ENGCRMT17	B, U, V	0.9	0.9	1.0	0.9	1.4	0.8	1.2	0.9	0.6	
	ENGCRMT16	U, V, W	0.7	1.0	1.0	1.0	1.0	0.9	1.0	1.0	0.9	
	ENGCRMT14	V	0.7	0.9	1.0	1.0	1.2	0.7	1.0	0.9	0.9	
	ENGCRMT11	D, J	0.7	0.9	1.0	0.8	0.8	0.8	0.7	1.0	1.1	
	ENGCRMT15	K, L	0.7	0.9	1.0	1.0	1.0	0.9	1.0	0.9	0.9	
	ENGCRMT12	Q	0.5	0.5	1.0	0.6	1.2	0.2	0.5	0.6	0.5	
Task 2 Storm Barriers	ENGCRMT28	M, W	1.2	2.1	3.0	2.7	1.5	2.7	2.0	1.8	2.1	1.0
	ENGCRMT23	P, R	0.6	0.7	0.0	0.6	0.9	1.1	0.8	0.4	0.6	1.0
	ENGCRMT24	S	0.6	1.3	2.0	1.7	1.4	1.1	1.6	1.5	1.1	1.0
	ENGCRMT25	O, Q	0.4	0.7	1.0	0.7	0.9	1.0	0.4	0.8	0.5	0.3
	ENGCRMT27	M, T	0.4	0.5	1.0	0.3	0.9	0.8	0.0	0.5	0.4	0.3
	ENGCRMT21	K	0.4	0.5	1.0	0.4	0.4	0.4	0.4	0.6	0.4	0.3
	ENGCRMT26	D, T, V	0.3	0.4	1.0	0.7	0.3	0.3	0.4	0.3	0.4	0.7
ENGCRMT22	J, Q	0.2	0.3	1.0	0.3	0.3	0.6	0.4	0.4	0.2	0.7	
Task 3 Concrete and Bridges	ENGCRMT36	A, B, U	1.3	2.2	2.5	2.5	2.1	2.0	1.8	1.7	2.4	
	ENGCRMT34	T	0.7	0.8	0.5	0.9	0.9	0.7	0.8	0.9	0.6	
	ENGCRMT33	K	0.7	0.7	1.0	0.7	0.7	1.0	0.8	0.7	0.6	
	ENGCRMT32	D, T	0.6	0.4	0.0	0.8	0.3	0.3	0.4	0.1	0.4	
	ENGCRMT35	L, O	0.5	0.6	1.0	0.6	0.6	0.8	0.7	0.6	0.4	
	ENGCRMT31	N	0.1	0.1	0.0	0.0	0.0	0.2	0.4	0.1	0.1	

AHELO Feasibility Study 2012
 Australian Institutions
 Civil Engineering Learning Outcomes
 Multiple Choice Items

Item Percentage Correct
 Not Administered

Set	Item	Item description code	% Correct									
			AHELO Sample	Australian Sample	Institution 5	Institution 4	Institution 6	Institution 9	Institution 3	Institution 2	Institution 10	Institution 1
Set 1	ENGMCQ3	G	54	81	95	82	95	94	100	43	72	100
	ENGMCQ1	C	51	47	73	41	40	35	0	57	49	0
	ENGMCQ2	E	35	36	18	45	60	59	50	36	16	100
	ENGMCQ5	H	20	18	23	23	20	12	50	14	14	0
Set 2	ENGMCQ7	C	61	57	65	65	38	50	50	65	57	50
	ENGMCQ8	F	54	50	39	52	62	56	75	71	36	50
	ENGMCQ9	F	54	64	61	78	62	69	50	71	59	0
	ENGMCQ10	G	50	51	78	61	52	63	100	41	27	0
	ENGMCQ6	C	47	55	65	70	67	63	25	41	41	50
Set 3	ENGMCQ13	C	55	66	81	67	64	77	100	73	53	0
	ENGMCQ14	E	33	30	42	26	36	46	33	33	19	0
	ENGMCQ11	H	28	34	62	37	55	31	33	20	14	0
	ENGMCQ12	H	26	24	35	26	27	15	33	27	16	0
	ENGMCQ15	E	22	25	42	30	23	23	0	20	16	33
Set 4	ENGMCQ16	F	59	57	38	76	62	76	67	63	40	100
	ENGMCQ20	I	45	57	67	52	71	65	67	50	47	67
	ENGMCQ18	F	40	41	52	28	38	53	33	25	44	67
	ENGMCQ17	G	25	31	43	60	38	18	0	6	20	67
	ENGMCQ19	H	18	14	19	16	14	6	0	13	16	0
Set 5	ENGMCQ22	F	48	49	63	75	57	53	33	53	26	0
	ENGMCQ23	G	38	42	54	33	52	33	33	27	39	100
	ENGMCQ25	I	32	51	46	58	43	47	0	67	52	67
	ENGMCQ21	C	27	30	42	42	33	33	33	33	13	33
	ENGMCQ24	H	24	32	58	29	33	40	100	13	20	0
Set 6	ENGMCQ29	F	63	70	95	71	50	82	100	77	61	50
	ENGMCQ27	F	31	31	32	42	25	35	33	31	27	0
	ENGMCQ30	H	16	22	32	25	35	18	67	23	9	0
	ENGMCQ28	I	6	6	5	0	0	6	0	0	14	0

AHELO: The experience of Ontario's institutions

Mary Catharine Lennon

Senior Research Analyst, Higher Education Quality Council of Ontario

Introduction

In 2011 Ontario joined the Organisation for Economic Co-operation and Development's (OECD) Assessment of Higher Education Learning Outcomes (AHELO) feasibility study. The Higher Education Quality Council of Ontario (HEQCO) led the project on behalf of the Ministry of Training, Colleges and Universities (MTCU) and in cooperation with the Council of Ministers of Education, Canada (CMEC).

Initiated in 2006, AHELO is a feasibility study to determine if standard generic and discipline-specific tests can be used in different countries to measure what university students know and are able to do. Intending to contribute to the international conversation on establishing better indications of learning quality, the study aimed to develop common learning outcomes and assess student performance at the end of a bachelor's degree (first cycle) in a variety of educational cultures, languages and institutions through standard tests. The feasibility study developed three assessments: one for generic skills, and two for discipline specific skills in economics and civil engineering.

Seventeen countries are represented in this global project, and Canada is one of nine jurisdictions participating in the engineering strand. Nine out of ten Ontario universities with civil engineering programs participated in the study. Nearly 450 Ontario students participated, representing approximately 61% of all civil engineering graduating students.

The following report reviews the experience of Ontario's participation in the feasibility study, focussing primarily on the implementation and administration activities, and the value to institutions. While the institutions did not gain specific insight on their programming, AHELO generated considerable interest in international assessments and comparative understanding, and provided significant experience in the administration of large-scale assessments.

Background and Rationale for the AHELO Feasibility Study

The late 20th century saw significant changes to higher education worldwide.¹ The massification of higher education produced a diverse profile of institutions, programs and students unlike the small elite systems of previous times (OECD, 2012a). This expansion was aided by a number of elements. Higher education is no longer contained by either bricks and mortar or national/jurisdictional boundaries. The proliferation of technology-enhanced learning allows programs to operate whenever and wherever the student chooses, and international student and faculty mobility have opened institutions to the world. Furthermore, the student demand for education has supported the creation of new and alternative providers, such as private institutions or those with specialised programming.

The now complex and global "market" for postsecondary education (PSE) demands new forms of governance, accountability and signalling mechanisms. In line with traditional forms of accountability and governance, measures of inputs (funding, library holdings, etc.) and outputs (retention rate, graduation rate, publications) have been the yardstick of performance both within systems and internationally. International rankings based on these performance indicators play a significant role in weighing the research capacities of institutions.

What these traditional measures fail to capture is the quality of education: the teaching and learning that is the heart of most institutions. Recognition of this gap led to the inclusion of another set of indicators that can be called "proxies". These "proxies" of quality include student evaluations, surveys of student satisfaction and engagement,

¹Information contained in this section is largely based on OECD documentation. For further information, please see Tremblay, Lalancette and Roseveare (2012) and OECD (2013a)

labour market outcomes, etc. These indicators suggest that if the student is satisfied and has fared well beyond PSE, then the institution must have provided high-quality education.

More recently, and considered by some to be a paradigm shift (OECD, 2012a: 35), there has been a focus on learning outcomes as a means to understand, demonstrate and assess educational quality. Defined learning outcomes – clear statements of what a learner knows and can do – along with appropriate assessment measures, provide a transparent means to measure student learning. It is thus possible to gain an objective picture of the quality of teaching and learning, and ultimately the quality of education provided.

Clear indications of teaching and learning quality are beneficial in a number of ways. They support a better understanding of educational value to students, employers, and the public at large. They also enhance institutional and programmatic improvement in coordination, curriculum development and teaching practices. Furthermore, they can provide measures of what has been considered intangible until now in the world of quality assurance and accountability – educational quality. They also provide a transparency of programming, which allows for greater international and comparative understanding for institutions and programs. Hence, learning outcomes demystify education to the benefit of the program, institution, and wider public and international community, but ultimately they put students at the centre of it all. This indicates a significant shift towards student-centred learning.

The purpose of the feasibility study was to provide ‘proof of concept’ (OECD, 2009a: 15) to determine whether it is possible to measure what undergraduates know and can do at the international level, to provide relevant information to higher education institutions (HEI’s), governments, and other stakeholders including students and employers (OECD 2009b: 2)

The primary questions that emerged from the work included:

- Is it possible to have international agreement of expected learning outcomes?
- Is it possible to implement the same test across cultures and languages?
- Are the assessments valid and reliable?

Three primary areas of work made up the feasibility study. Three strands were developed to assess student performance. One sought to assess generic skills, such as capacities in critical thinking, analytical reasoning, problem solving, etc.². The other two assessments concentrated on discipline-specific skills in economics³ and engineering.⁴ Rather than assessing content knowledge, both discipline assessments focused on the application of knowledge (i.e., can a student “think like an engineer”). An additional strand, “value-add”, was intended to explore the contribution of higher education to student learning. It is expected that there will be gains in skills and competencies from when a student enters PSE and when they graduate (see OECD, 2013b).

In addition to these two areas of work, a Context Dimension⁵ was developed to survey students, faculty, institutions and jurisdictions to better understand the educational environment of each jurisdiction and identify factors that might explain differences in observed performance of students.

Concurrent to the development of the assessment frameworks, the OECD invited member countries⁶ and other interested countries to join the feasibility study, share in the development of the assessments and pilot the test to a sample of their institutions. Seventeen jurisdictions agreed to take part in the study, representing six continents and 12 languages. Ontario participated in the Civil Engineering Strand.

²For more information, see Tremblay et al. (2012); p 112-114.

³For more information, see OECD 2011a; Tremblay et al. (2012), p 116-118.

⁴For more information, see OECD 2011b; Tremblay et al. (2012), p 121-123; OECD, 2012a; OECD 2012b

⁵For more information, see Tremblay et al. (2012) p 128-132; OECD, 2012c

⁶For a list of member countries, see <http://www.oecd.org/general/listofocedmembercountries-ratificationoftheconventionontheoecd.htm>

AHELO in Ontario

In July 2011, the Ministry of Training, Colleges and University in Ontario decided to join the study independent of other Canadian provinces⁷. The Higher Education Quality Council of Ontario (HEQCO⁸), an arms-length agency of the Ontario government, was asked to lead the project.

As part of a broad research agenda to understand and demonstrate the value of postsecondary education, Ontario, through HEQCO, has engaged in a number of research projects to explore learning outcomes. HEQCO has been engaged in a multitude of activities in the area, including establishing learning outcomes across sectors of disciplines (the Tuning project; see Lennon, Frank, Humphreys, Lenton, Madsen, Omri & Turner, forthcoming A); measuring generic learning outcomes through piloting the Collegiate Learning Assessment (see Lennon, forthcoming B); as well as supporting the incorporation and demonstration of learning outcomes through a variety of institutional activities. Thus, participating in the international study supported the research base by providing comparative information on systems and programming.

Acting as the National Project Centre, HEQCO was responsible for liaising with the OECD and AHELO Consortium in order to administer and implement the assessment. This entailed acting as national experts to the OECD AHELO conversation, providing a National Project Manager (NPM), implementing and administering the study with participating institutions, and providing analysis.

There were a number of activities involved in setting up and administering the study in Ontario (see Figure 1). The feasibility study, operationalized by the Consortium, provided guidance and protocols for implementation in order to ensure that comparable activities took place in all jurisdictions. The following sections review the primary areas of field work and implementation, documenting both OECD protocols and Ontario's actions.

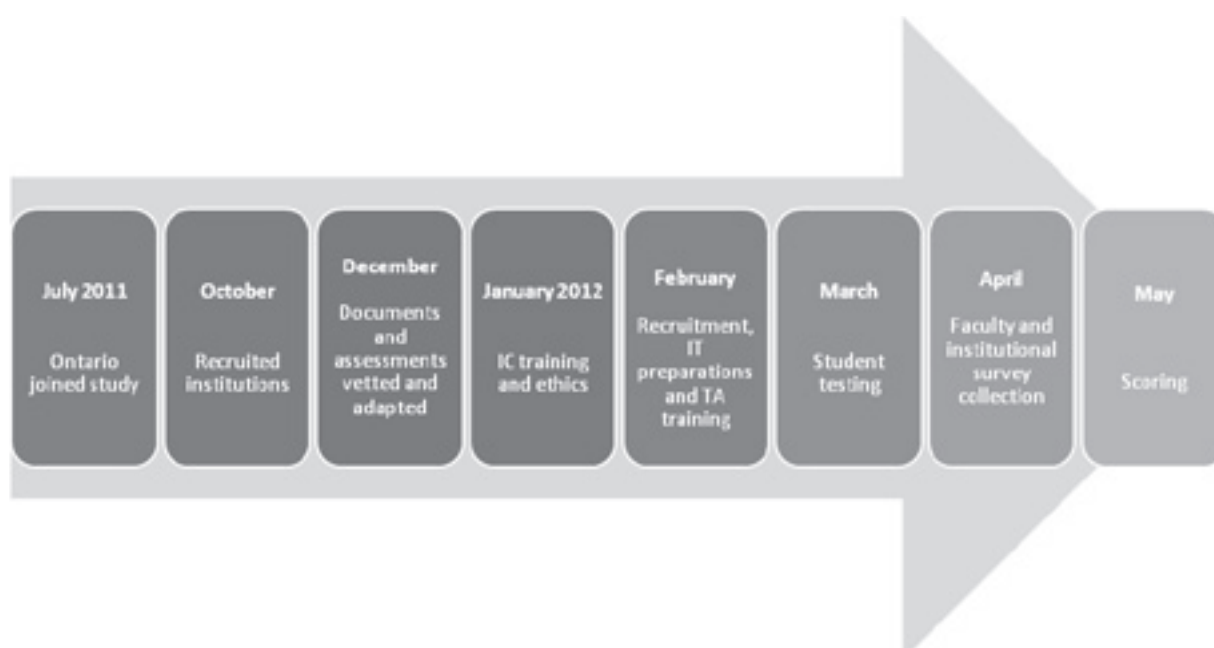


Figure 1: Basic Timeline for Ontario's Implementation Activities

Institution Recruitment and Participation

The OECD recommended that each jurisdiction have a convenience sample of 10 institutions to represent a range of

⁷The Council of Ministers of Education Canada represents higher education at the OECD table and, with agreement from all of the provinces, engages in various OECD projects. Despite general interest, CMEC was unable to gain consensus from the provinces to join the feasibility study.

⁸See www.heqco.ca for more information

types of institutions providing bachelor-level (or first-cycle) civil engineering degrees. It was desirable to have public and private universities/polytechnics/colleges, institutions of different sizes and with different mandates (teaching or research), and those in rural and urban settings, etc. The goal was to ensure that all types of institutions and of students were adequately represented in order to have a valid and reliable understanding of each of the systems.

In Ontario, 10 institutions provide bachelor-level civil engineering programs. Hence, each institution was invited to participate via a letter to the vice-presidents academic and the deans of engineering. To support participation in the project, each institution was offered a nominal sum to cover the costs of administration. Responses from the institutions were immediate and extremely favourable: nine out of 10 institutions agreed to participate, noting their interest in taking part in this international assessment as a way of understanding their own students and program, as well as those internationally.

Ontario Engineering programs were already familiar with notions of learning outcomes and strategies to assess them. As part of the Washington Accord agreement, and for accreditation purposes, engineering programs across Canada are moving to outcomes-based programming, assessment and accreditation for 2016 (EGADS, 2013). Hence there was considerable interest in exploring this large-scale assessment as a tool. Approximately 90% of all of Ontario's civil engineering students were represented in these programs.

The participating institutions were all public institutions offering a broad range of arts and science programs up to the doctoral level. Located primarily in urban areas, they ranged in size from 14,595 to 75,941 full-time equivalent (FTE) students. Participating institutions included:

- Carleton University
- University of Ottawa
- Ryerson University
- University of Waterloo
- University of Windsor
- McMaster University
- Queen's University
- University of Toronto
- Western University

The civil engineering programs were housed in faculties of engineering and were occasionally partnered with environmental engineering. The programs had between 17 and 40 faculty members. The full-time equivalent (FTE) student populations in the programs ranged from 231 to 573. The number of undergraduate degrees awarded annually ranged from 20 to 101 in the 2010 academic year.

Adapting, Vetting and Validating the Documents

A Lead Scorer was identified in each country to provide discipline/content expertise in the development of the assessments. The individual, a recognised expert in his or her field, was required to review all test documentation, adapting the assessments and translations to ensure appropriateness for their jurisdiction, and vet the assessment documents through pilot tests of students and faculty. The lead scorers from each country worked together to determine appropriate grading of test items and were responsible for training an in-country scoring team.

Ontario's Lead Scorer, in collaboration with the National Project Manager, organised various activities to vet the test in the Ontario context, including piloting the test to students⁹ and seeking feedback on the test from faculty members and senior graduate students. The feedback commented on the suitability of the content and difficulty level of the questions and their appropriateness for the Ontario context (including issues of language nuances and technical jargon).

This information was presented to the OECD and incorporated into the assessment. Ontario recommended few significant changes. Indeed there were very few items of contention within the entire international scoring team, and jurisdictional lead scorers came to agreement rather quickly.

⁹Five third-year students wrote the test and provided feedback. Third-year students were selected so that the population of fourth-year students would not be compromised.

Institutional Activities

Each institution participating in the study identified an institutional coordinator. This individual was responsible for administering the study within their institution and liaising with the HEQCO as the National Centre. Playing a vital role in the AHELO study, institutional coordinators were responsible for operationalizing AHELO within their institutions.

Standardising field activities was critical to ensure that any resulting information would be reliable. Regardless of how well-designed the assessment might have been, administration and implementation issues could have impacted the validity and reliability of the data. Thus, an important part of the feasibility study involved securing common field activities.

Operationalizing AHELO Within Institutions

Each institution was in close contact with its National Centre to ensure consistency of implementation across a jurisdiction, and also to receive any support required. In Ontario, Institutional Coordinators (ICs), many of whom were the chairs of civil engineering, were identified within their institutions. The nine individuals attended a one-day training session at HEQCO to provide them with context and background on AHELO, and a roadmap for activities. Following the face-to-face meeting, the ICs and the NMP had weekly teleconferences to provide support and advice on activities.

There were a number of tasks involved in operationalizing the field work, and each institutional coordinator developed a team to assist them in the tasks. Recommended team members included an IT specialist and test administrators. The majority of Ontario institutions followed this recommendation, some with larger teams than others.

One of the ICs' first tasks was to apply for institutional ethics approval to implement the test. Unlike other jurisdictions, Ontario institutions require permission to perform research on human subjects, including students. Based on the short timelines and the need to implement the test quickly, Ontario institutions were required to modify the parameters of the research in order to obtain ethics approval.

For example, whereas AHELO had requested that the ICs have access to students' administrative information (to ensure that the sample was representative of the general population and to better understand the assessment results based on GPA, high school average, etc.), Ontario institutions were not given permission to link the AHELO test results to student files¹⁰. This resulted in less reliable information on the student sample writing the test and reduced institutions' ability to examine the data. For example, had institutions had access to administrative files, they would have been able to examine individual AHELO test scores with other indicators of student ability, such as GPA. Institutional information of this nature would have supported a better understanding of the validity of the AHELO scores in Ontario¹¹.

Implementing Field Work

Student and faculty recruitment

Student recruitment was a critical aspect of the AHELO feasibility study. Each institution was requested to identify a sample 200 students to write the test and aim for a 75% recruitment rate. For those institutions with more than 200 students in the designated area (engineering or economics, or the entire institution in the case of Generic Skills), ICs were required to provide a sampling frame indicating which students were identified as potential participants.¹² This purposeful sampling frame was intended to ensure that a representative sample of the population wrote the test.

¹⁰Only ICs were to have access to the student administrative data. The national centre, consortium and OECD would not.

¹¹Despite the challenges of gaining ethics approval for this feasibility study, any subsequent assessment of this nature would likely be accepted without significant modification given additional time to develop ethics applications and to communicate the goals of the study.

¹²This sampling framework was conducted under the guidance of the Consortium.

In cases where institutions had less than 200 potential test writers, the goal was to have all students participate as a census. All Ontario civil engineering programs had less than 200 final-year students¹³, so all eligible students¹⁴ were identified as potential candidates for the assessment at each institution.

Jurisdictions independently determined their own student recruitment strategies. While some made student participation mandatory, others made it voluntary. Those with voluntary participation did not always incentivize students, though the majority did (OECD, 2013a, p. 169). Student participation in Ontario was entirely voluntary. ICs recruited students in any way they felt was appropriate, as long as it was within the guidelines of their ethical protocol. A range of promotional activities and incentives were offered to students to entice participation, such as posters and information sessions, offers of gift certificates, prize draws and donations to civil engineering class societies. Recognising that it was imperative to have sufficient numbers to conduct any type of analysis, student recruitment was by far the most time-consuming and anxiety-provoking activity for ICs. As student participation was voluntary, all Ontario institutions ended up with a non-random, voluntary sample.

Faculty participation in the short context survey was conducted in a similar way. If a faculty (or institution in the case of generic skills) had more than 40 members, a sampling frame was conducted to ensure that a representative sample was recruited. If there were fewer than 40 faculty members, as was the case in all participating Ontario institutions, a census of all students was attempted. Unable to mandate or entice all faculty members to participate in the survey, Ontario institutions ended up with a voluntary sample of volunteers.

Test administration

The AHELO test system was operational between February and June 2012, and institutions could test their students at any time during this window. Each institution was required to run a test of the computer platform to ensure there were no technical difficulties. The times and dates of test sessions were reported in advance to the NPM (in case of technical or emergency issues), who reported this information in turn to the consortium to ensure that the online system was prepared to handle the influx of activity. Institutional coordinators were responsible for organizing the AHELO test session and ensuring that the test administrator (invigilator) was suitably trained (based on AHELO training guidelines).

In Ontario, the timing of the test window was challenging. The academic calendars schedule a one-week break in mid-February followed by mid-term exams and the close of the academic year in April, preceded by final projects and exams. As a result, a very short period of time remained for test administration. Most institutions ran their tests in early to mid-March. Institutions offered the test at a variety of times and days, often taking into consideration student class schedules.

Despite testing the system, one Ontario institution ran into technical difficulties and students were unable to submit their tests. While the Consortium rectified the technical issue, all but 8% of student responses were lost at that institution.

Scoring

While the multiple choice questions on student assessments were scored automatically using a computerized system, each jurisdiction was responsible for the manual scoring of the constructed responses of its own students. The lead scorers attended two training sessions with fellow lead scorers to finalize the assessment questions and determine appropriate scoring matrices. This activity ensured the consistency of marking around the world. The lead scorer, in collaboration with the NPM, was responsible for recruiting a small team of scorers and training them in both the test system and the scoring matrix.

¹³Participating programs had between 34 and 137 final year-students

¹⁴Students were excluded if they were out of the country or had disabilities requiring special arrangements.

In Ontario, a team of six engineers made up the scoring team. The majority of scorers were ICs interested in reviewing the assessments and the work of Ontario students. Scoring took place over two days in June 2012 at the National Centre. As part of a memorandum of understanding (MOU) with Australia¹⁵, Ontario scored some Australian answers and vice versa. This was conducted as an experiment in inter-rater reliability between scoring jurisdictions and as a point of interest for the scoring teams (who would not otherwise have seen other jurisdictions' student responses).

Successes and Challenges of Administration

A primary objective of the AHELO feasibility study was to understand if it was administratively possible to implement a standard assessment, online, in a common way, to students around the world. Indeed it was proven to be possible: experts and faculty members agreed on the common learning outcomes and the assessment questions, and project management and execution of tasks followed a common protocol around the world. There were minor administrative issues in various countries –some institutions dropped out, others had very low response rates or technical issues – but overall the administration of the international assessment was proven successful. The first volume of the AHELO report lays out a number of international lessons learned from the administration phase of the study (OECD, 2012a, Ch. 6).

There were considerable differences in administrative successes and challenges between participating jurisdictions, as well differences between the three strands. For example, one country recognised that the institutions did not have enough computers for the number of students completing the assessments, and thus required a travelling van to deliver computers to the institutions. Incredibly, Egypt was scheduled to implement the test during the Arab Spring, which created significant – though not insurmountable – challenges. The Generic Skills strand faced challenges in student recruitment.

These implementation issues are critical to understanding the practicality of administering the test – a primary goal of the study. Hence, a great deal of work has been done in collecting and international experiences (see OECD, 2013a; Ch. 8 for a description of each nations' experience).

Within Ontario, challenges and successes were both generic to all large-scale testing and specific to the AHELO context. Generally, student recruitment for low-stakes testing is extremely challenging. It is time consuming in both the advertising (posters, emails, class visits, etc.) and in the organizing of test sessions. Furthermore it can become extremely expensive when students are provided with material incentives, financial or otherwise. Despite the challenges of student recruitment, the institutions were extremely creative in their strategies, and despite being the most time-consuming activity, it was quite rewarding for the ICs to see their students participate.

As one would expect in a feasibility study, there were a few hiccups in Ontario. The need to obtain ethics approval quickly required institutions to modify the assessment framework to reduce the institutional capacity to link the AHELO results to individual student level data. This reduced the potential value of the results for the institutions; at the time of writing, no institution has yet analysed their own institutional data but have expressed interest in examining it when they receive the comparative jurisdictional report.

In follow up interviews and surveys student and faculty participants indicated appreciation in participating in an international exercise, thereby validating their decision to participate in the study. This suggests that rather than being fearful of benchmarks or comparisons, there was true interest, from all levels, in understanding their programmatic characteristics, strengths and weaknesses compared to those of others in jurisdictions around the world.

¹⁵More on the MOU sections below.

Discussion and Conclusions

Analysis of Ontario's Results

Nine jurisdictions and more than 70 institutions participated in the engineering strand of the feasibility study. In Ontario, out of the ten institutions that currently offer a civil engineering program, nine of them participated in the engineering strand of AHELO. Context information was collected from institutions. Across the nine participating institutions in Ontario, there were 155 faculty members and 443 final-year students who participated in the engineering strand of AHELO. Ontario had a response rate of 72% for faculty and 61%¹⁶ for students. While the demographic characteristics, employment status and qualifications of the faculty who participated in AHELO were similar across the institutions in Ontario, there was a noticeable difference in the amount of time faculty reported to spend on teaching and research. Focusing on the student population within the civil engineering program that participated in AHELO, there were also differences in involvement characteristics, such as how students spend their time preparing for class, attending class, or working in paid work related to or unrelated to field of study. Please see Lennon, forthcoming C, for more information on the institutional, faculty, and student characteristics of the participating institutions in Ontario and how they compared to Australia and all nine participating jurisdictions.

What does AHELO contribute to our understanding?

Recall that this feasibility study was not intended to provide any comparative ranking data, but rather to explore the potential for this type of work in both field work and in trialling the assessment tools. Thus, the potential for digging into the data was either a) not agreed to in the framework, or b) impossible due to data limitations discovered later.

Acknowledging that the data gathered must be interpreted with extreme caution, it is impossible to comment on the results and value they provide to stakeholders (i.e. jurisdictions, institutions, faculty members and students), as they are unable to presently use the information.

Thus, the feasibility study revealed that the tools need to be refined in order to say what contribution the information can make. What also became apparent, is that any reworking of the frameworks or assessments further begs the question of who is this information valuable to, in order to tailor it appropriately.

Let us now consider the value of the feasibility results in two ways:

1. What did the various stakeholder groups hope to gain from participating in AHELO?
2. How could the assessments be tailored to suit the needs of the various groups?

Jurisdictions

At the outset of AHELO it was made clear that the purpose was not for international rankings of countries or institutions either internationally or within a jurisdiction. However, the potential for comparing and contrasting across and within countries/jurisdictions was a significant draw. For governments, the ability to know how their institutions and programs are organised and examine the impact on student learning presents valuable information. This has the potential to lead to better comparative policy understanding, and improve the way systems interact with their institutions.

For a variety of reasons, this was not possible in the AHELO feasibility study, but the potential for system-level learning exists. For example, if the information was presented in a way that did not permit for rankings on simple data points, but instead provided comparable information to peer countries or an international average, the potential for international comparisons is possible.

¹⁶For institution 9, out of the 36 students that participated in the assessment, 33 of the responses were not recorded as a result of a technological difficulty with the online system. The response rate for Ontario was determined using the number of students that participated in the assessment rather than the number of actual recorded responses.

Institutional and program level data

The primary goal of the AHELO feasibility study was to provide information to institutions on how their students performed in various capacities compared to others. Similarly, information collected in the context surveys was intended to provide insight into the characteristics of the educational environments in which students work.

The assessments used in AHELO were not sensitive enough to provide institution level information on student capacities by specific competency areas¹⁷. Thus, it was not possible to compare institutions on anything other than the aggregate score of their students. This was disappointing to institutions participating in the feasibility study, however, with a refined tool it would be possible to develop a test that compares competency areas of students in institutions. This information would be very useful in providing an understanding of their strengths and weaknesses. For example, an institution might discover they are very strong in design and practice, but weaker in basic engineering skills. This type of detailed information could provide significant information to program and faculty members when considering curriculum design¹⁸. Understanding of how they compared to other programs in the jurisdiction or internationally was the primary reason Ontario institutions chose to participate in the study. Despite not being able to receive this information, the potential to reshape the tests in order to do it is possible.

The context data collected at the institutional and program level also has the potential to hold valuable information. Ontario institutions also collect administrative information on faculty and students. Thus, much of this information is readily available in other sources, and the information on their own learning environments was not particularly new information to institutions or programs, nor is the within-jurisdiction comparison.

Yet, the point of interest came from seeing the institutions and programs in comparison to others and the international averages. It could be possible to compare institutions and programs around the world to see the commonalities and differences in learning environments and understand if they impact student capacities. Similarly, a pooled analysis of the international trends in program design and environmental characteristics could be provocative in providing information on trends of successful programs or in providing benchmarking through common structural traits.

Thus, while there is value in the institutional information and comparisons currently provided, greater insight would be gained from international contrasts, as well as pooled information on global trends.

Student level data

Student-level data collected in AHELO had the possibility to provide information on observable trends in student demographics and characteristics, and assessment scores.

However, within the feasibility study framework, it was expected that institutions would be provided a comparison of their students' situations to those in local institutions, and that with that information they might be better able to support student success. For example, discovering that older students are less likely to be successful - perhaps due to outside employment - institutions could choose to develop different student support strategies. Providing comparative information might also encourage collaborations between programs grappling with similar issues. While not provided in the current AHELO framework, it would be interesting to identify international trends in habits of student success, thus supporting cooperation and collaboration regardless of jurisdictional boundaries.

The tests were not designed in a way to deliver student level feedback, as the intention was to combine student-level data at the institutional level. However it became clear in discussions that there is value and interest in providing student-level feedback: a way for students to demonstrate their capacities compared to their peers. In order to provide student level feedback the test would need to be significantly altered to ensure each student was writing a precisely equivalent (or same) test.

¹⁷Engineering Design, Practice, and Analysis, and Generic and Basic Engineering skills.

¹⁸It has been suggested that proper assessment of competency levels would require a re-design of the assessment, and would likely necessitate a longer test for students.

One aspect that was discussed at length by participating AHELO jurisdictions, the Consortium, and the OECD was the possibility of providing students with feedback on their assessment scores. The assessment frameworks were not designed in a way to provide reliable scores at the individual level, as it intended to provide institution level feedback. However, if students received their scores relative to others (either in the institution, jurisdiction, or internationally), there could be increased interest from students as it would provide a global and objective documentation of their abilities. This, in turn, could improve student recruitment and student effort in writing the test (creating more reliable information). This was not the intention of the AHELO feasibility study, but was of significant interest to many participating jurisdictions. This is one of many possible outputs of this type of assessment, but to doing so would necessitate a different assessment framework.

Conclusion

The AHELO feasibility study was a successful enterprise in building international relationships, aiding comparative system-level understanding, supporting institutional/programmatic understanding and exploring the potential of international student-level assessments. A great deal was learned from this research, both within jurisdictions and at the international level of the OECD. We know, for example, that it is possible to administer a standard test to students around the world, and it seems there is interest from a variety of stakeholders, particularly the engineering programs, to do so.

Within Ontario, we benefited from participating in the international conversation on the broader value of establishing and measuring learning outcomes. As Ontario is grappling with how to develop a system that incorporates learning outcomes, understanding how other systems are engaging with them has been appreciated. It was particularly interesting to see the tension between how various jurisdictions desire to use international learning outcomes assessments as either system level benchmarks or for institutional and program level improvement.

As a feasibility study, the AHELO work produced many lessons learned but raised even more questions. It was found that there is interest in international assessments from governments, institutions, programs, faculty members and students. It was also determined that it is possible to have agreement on expected learning outcomes and appropriate assessments from around the world. Furthermore, it was found that it is possible to test students around the world in compatible ways. These were the primary questions of the AHELO study.

References

- EGAD (2013) *Engineering Graduate Attribute Development Project*. Retrieved from: <http://egad.engineering.queensu.ca/>
- Lennon, M. C. (2010). *Signalling Abilities and Achievement: Measuring and Reporting on Skill and Competency Development*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming A). *Tuning: Identifying and Measuring Sector-Based Learning Outcomes in Postsecondary Education*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming B). *Piloting the Collegiate Learning Assessment in Ontario: Lessons Learned*. Toronto: Higher Education Quality Council of Ontario.
- Lennon, M.C. (forthcoming C). *The AHELO Feasibility Study in Ontario*. Toronto: Higher Education Quality Council of Ontario.
- OECD (2009a). *AHELO Terms of Reference*. Paris, OECD
- OECD (2009b). *Roadmap for the OECD Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study*. Paris, OECD
- OECD (2011a). *Tuning-AHELO Conceptual Framework of Expected and Desired Learning Outcomes in Economics*. Paris, OECD Publishing.
- OECD (2011b). *A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering*. Paris, OECD Publishing
- OECD (2012a). *Engineering Assessment Framework*. Paris, OECD.
- OECD (2012b) *Final Engineering Assessment Development Report*. Paris, OECD.
- OECD (2012c) *Contextual Dimension Framework*. Paris, OECD.
- OECD (2013a) *AHELO: Feasibility Study Report: Volume 2: Data Analysis and National Experiences*. Paris, OECD.
- OECD (2013b) *AHELO: Feasibility Study Report: Volume 3: Further Insights*. Paris, OECD.
- Tremblay, K. Lalancette, D., and D. Roseveare (2012). *AHELO: Feasibility Study Report: Volume 1: Design and Implementation*, Paris, OECD
- Tuning Association (2009). *A Tuning-AHELO Conceptual Framework of Expected/Desired Learning Outcomes in Engineering*. Netherlands, Tuning Association.

Presenters' Biographies

Robert Wagenaar

Prof. Robert Wagenaar is a historian and at present director of undergraduate and graduate studies at the Faculty of Arts of the University of Groningen in the Netherlands. He is also Joint Director of the Tuning Academy, located in Bilbao (Spain) and Groningen; and a member of the Editorial Board of Tuning Journal for Higher Education (TJHE). He is also coordinator and director of the Erasmus Mundus Master Course of Excellence Euroculture: Europe in the Wider World. Furthermore, he is an external expert on Higher Education for the European Commission and has been involved in main initiatives to harmonise European Higher Education, such as the development of a European Credit Transfer and Accumulation System (ECTS) since 1988 and the Qualifications Framework for the European Higher Education Area and a European Qualifications Framework for LLL. He also chairs the Dutch team of experts for the implementation of the 'Bologna Process' in Dutch Higher Education institutions. Together with Julia Gonzalez (University of Deusto, Bilbao, Spain), Wagenaar elaborated, designed and coordinates the large scale innovative project Tuning Educational Structures in the World. In 2009 the project developed the Tuning-AHELO Conceptual Framework for Expected / Desired Learning Outcomes in Economics and Engineering on invitation of the OECD.

Peter Ewell

Peter T. Ewell is Vice President of the National Center for Higher Education Management Systems (NCHEMS), a research and development center founded to improve the effectiveness of colleges and universities. A member of the staff since 1981, Dr. Ewell's work focuses on assessing institutional effectiveness and the outcomes of college, and involves both research and direct consulting with institutions and state systems on collecting and using assessment information in planning, evaluation, and budgeting. He has directed many projects on this topic, including initiatives funded by the W. K. Kellogg Foundation, the Lumina Foundation for Education, the Ford Foundation, the Consortium for the Advancement of Private Higher Education, and the Pew Charitable Trusts. In addition, he has consulted with over 375 colleges and universities, twenty-four state systems of higher education, and twelve countries on topics including outcomes, program review, enrollment management, and strategic planning.

Dr. Ewell has authored six books and numerous articles on the topic of improving undergraduate instruction through outcomes-based approaches. In addition, he has prepared commissioned papers for many agencies, including the Education Commission of the States, the National Governors' Association, the National Conference of State Legislators, the National Center for Public Policy in Higher Education, and the Organization for Economic Cooperation and Development (OECD). Widely sought as a speaker, in 1985 he gave the keynote address for the first national conference on Outcomes in American Higher Education, and has since spoken widely on this topic at both U.S. and international conferences. In 1996 he designed the curriculum for the outcomes-based Western Governors University (WGU) and in 1998 he led the design team for the National Survey of Student Engagement (NSSE) now used by more than 1200 colleges and universities in the U.S.

A graduate of Haverford College, Dr. Ewell received his Ph.D. in Political Science from Yale University in 1976 and was on the faculty of the University of Chicago.

Mary Catharine Lennon

Mary Catharine Lennon is a Senior Research Analyst at the Higher Education Quality Council of Ontario (HEQCO), an arms-length research agency of the Ontario government, where she is currently leading projects on establishing and measuring learning outcomes.

She has been involved in higher education policy development, advice and research in institutional, provincial, inter-provincial and international educational agencies including the Council of Ministers of Education, Canada, and the Association of Commonwealth Universities.

She also studied comparative and international higher education issues and policies in both her Masters and Doctoral programs where she is currently a PhD Candidate at the University of Toronto.

With this background, the majority of her work extends towards international and comparative system level policy issues including System Design, Accountability, Quality Assurance and Governance.

She is currently responsible for HEQCO's suite of student learning outcomes projects including piloting the Collegiate Learning Assessment, facilitating Tuning projects with three sectors of academic disciplines, and acting as the Canadian National Project Manager for the OECD's AHELO feasibility study.

Mary Catharine is also co-editor of 'Measuring the Value of a Postsecondary Education', a collection of international examples of learning outcome activities.

Daniel Edwards

Daniel Edwards is a Principal Research Fellow in the Higher Education research program at the Australian Council for Educational Research. He was appointed in January 2008 to help develop ACER's higher education research.

Dr Edwards' research encompasses a range of educational issues, with particular emphasis on higher education. He has explored issues relating to demand for higher education places (both amongst students and employers), student achievement, student aspirations and pathways, selection policies for entrance to university, and educational 'choice' theories. He also has experience researching wider social issues regarding social stratification and demographic change.

Dr Edwards' leads a range of projects providing advice and insight into policy and practice across the spectrum of higher education for governments, universities, stakeholder organisations and other bodies. Projects he has led include the National Research Student Survey (NRSS) project, the National Project Management of AHELO in Australia, development of the AHELO Contextual Dimension. Daniel also edits and manages the ACER Joining the Dots, a policy resource for Australian Higher Education.

Dr Edwards is also an Adjunct Research Fellow at the Centre for Population and Urban Research at Monash University.

Tsutomu KIMURA

Advisor to the Ministry of Education, Culture, Sports, Science and Technology

Education:

- Ph.D. (Engineering) Tokyo Institute of Technology, 1968
- M.A. (Civil Engineering) The University of Tokyo, 1964
- B.A. (Civil Engineering) The University of Tokyo, 1961

Professional Experience:

- Professor Emeritus and Specially Appointed Professor, National Institution for Academic Degrees and University Evaluation (NIAD-UE) (Apr. 2009 – present)
- President, Japan Accreditation Board for Engineering Education (Jun. 2009 – present)
- Advisor to the Ministry of Education, Culture, Sports, Science and Technology (Apr. 2009 – present)
- President, NIAD-UE (1998 – 2009)
- Professor Emeritus, Tokyo Institute of Technology (Mar. 1998 – present)
- Fellow, Churchill College, University of Cambridge (1997 – 1998)
- Professor, School of Engineering, Tokyo Institute of Technology (1997 – 1998)
- President, Tokyo Institute of Technology (1993 – 1997)
- Dean, School of Engineering, Tokyo Institute of Technology (1992 – 1993)
- Professor, School of Engineering, Tokyo Institute of Technology (1982 – 1993)
- Researcher, University of Cambridge (1978 – 1979)
- Researcher, University of Strathclyde (1971 – 1973)
- Assistant Professor, School of Engineering, Tokyo Institute of Technology (1968 – 1982)
- Assistant, School of Science and Engineering, Tokyo Institute of Technology (1965 – 1968)
- NIPPO Corporation (1961 – 1965)

Awards:

- The Order of the Sacred Treasure, Gold and Silver Star, 2013
- Japan Academy Prize, 2012
- Most Excellent Order of the British Empire, CBE, 2004
- Japan Society of Civil Engineers Award (thesis prize), 1988
- Japanese Society of Soil Mechanics and Foundation Engineering Award (thesis prize), 1982
- Seiichi Tejima Memorial Research Award (writing prize), 1969
- Japan Society of Civil Engineers Award (thesis encouragement prize), 1966

External Appointments:

- Special Appointment Member, Central Council for Education (2007 – present)
- Member, Science Council of Japan (2005 – present)
- Chairman, Tokyo Metropolitan Board of Education (2004 – present)
- Member, National University Evaluation Committee (2003 – 2004)
- Chairman, Subcommittee on Selection of Persons of Cultural Merits, Council for Cultural Affairs (2002 – 2003)
- Chairman, Independent Administrative Institution Evaluation Committee, Ministry of Land, Infrastructure, Transport and Tourism (2001 – 2010)
- Chairman, Independent Administrative Institution Evaluation Committee, Ministry of Economy, Trade and Industry (2001 – 2011)
- Vice Chairman, Central Council for Education (2001 – 2007)
- Chairman, Council for Selection of Persons of Cultural Merits (1999 – 2000)
- Chairman, Science Education and Industrial Education Council (1997 – 1999)

Selected Papers and Publications:

- Soil Mechanics (Civil Engineering Series, Vol. 8)*, Shokokusya Publishing, 1980
- Diffusion of Soil Stress*, Kajima Institute Publishing, 1978

Kikuo KISHIMOTO

Dr. Kikuo Kishimoto is currently a Professor of the Department of Mechanical Sciences and Engineering and Dean of School of Engineering, Tokyo Institute of Technology. He received his B.S. degree in 1975, M.S. degree in 1977, and Doctor of Engineering degree in 1982 from Tokyo Institute of Technology. He worked as a Research Associate and Associate Professor at Tokyo Institute of Technology from 1977 to 1995. During this period, he was a Visiting Scholar at Cambridge University from 1987 to 1988. He previously served as Vice President for education in 2012.

He has published over 250 journal papers in the areas of applied mechanics, fracture mechanics, reliability of microelectronic devices, and others. He has also published authoritative and widely used reference books in these fields. In 1980 and 2000, he was awarded the JSME Medal for Best Paper for his contributions to dynamic fracture mechanics and interfacial mechanics, respectively. He also received the Best Paper Award from Japan Society of Corrosion Engineering in 1993, Society of Materials Science Award for Academic Contribution in 2006 and JSME Materials and Mechanics Division Achievement Medal in 2007.

He is a fellow of Japan Society of Mechanical Engineers, a fellow of Society of Automotive Engineers of Japan and a fellow of American Society of Mechanical Engineers. He is a member of Science Council of Japan and serving as a chairman of mechanical engineering committee. He has served and chaired on various scientific and technological committees promoting research and education. He is a vice-president of Japan Accreditation Board for Engineering Education (JABEE).

Motohisa KANEKO

Professor, Research Center for University Studies, University of Tsukuba
Professor Emeritus, The University of Tokyo

Education:

- Ph.D. The University of Chicago, 1985
- M.A. The University of Tokyo, 1974
- B.A. The University of Tokyo, 1972

Professional Experience:

- Professor, Research Center for University Studies, University of Tsukuba (Current Position)
- Professor and Research Director, Center for National University Finance and Management
- Dean, Graduate School of Education & Faculty of Education, University of Tokyo
- Professor, Graduate School of Education, University of Tokyo
- Associate Professor, University of Tokyo
- Associate Professor, Research Institute for Higher Education, Hiroshima University
- Visiting Assistant Professor, State University of New York at Albany
- Consultant, World Bank
- Researcher, Institute of Developing Economies, Japan External Trade Organization

External Appointments:

- Member, Central Council for Education
- Member, Science Council of Japan
- President, Japanese Association of Higher Education Research-Advising professor for Peking University, East China Normal University, Fudan University, and University of Science and Technology of China

Selected Papers and Publications:

- Reconstructing College Education in Japan*, Tamagawa University Press, 2013
- Universities' Capacities to Teach*, Chikuma-Shobo, 2007
- P. Altbach & T. Umakoshi (eds.) *Past and Future of Asian Higher Education*, Baltimore: Johns Hopkins University Press, 2004 (co-authored)
- Graduate Education in Engineering in the Age of Globalization*, Report from the Science Council of Japan's engineering education meeting, 2003
- Education, Economy, and Society*, Course material for the Open University of Japan, 1996

Satoko FUKAHORI

Senior Researcher, Department for Higher Education Research, National Institute for Educational Policy Research of Japan

Education:

BA in Comparative Education, Kyoto University (1991).

MA in Comparative Education, Kyoto University, Graduate School of Education. (1993).

Ph.D. in Sociology of Education, Columbia University, Graduate School of Arts and Sciences -Teacher College. (2000).

Academic Appointments:

Assistant Professor, Institute of Social Science, The University of Tokyo.

Associate Professor, Department of Education, Kyoto Women's University and Junior College.

Senior Researcher, Department for Higher Education Research, National Institute for Educational Policy Research.

Major External Appointments:

Expert Group Member for the Government Official Examination, National Personnel Authority.

Advisory Board Member for the Management of the Super Science High School Program at The Tokyo Gagei University Senior High School.

Research Committee Member, The Japan Society of Educational Sociology; Editorial Board Member.

Editorial Board Member, Tuning Journal for Higher Education.

Globalization Project Member, Liberal and General Education Society of Japan.

Research Institute for Higher Education at Hiroshima University, Visiting Scholar.

Major Publications:

Fukahori, S. ed. 2012. *A Study on the Impact of Learning Outcomes Assessment*, National Institute for Educational Policy Research. pp.1-200. (in Japanese)

Gonzales, J. and Wagenaar, R. eds. 2008. *Tuning Educational Structures in Europe – Universities' Contribution to the Bologna Process – An Introduction*. Publicaciones de la Universidad de Deusto. Translated by Fukahori, S. and Takenaka, T. (2012). Akashi Shoten. pp. 1-198.

Fukahori, S. 2011. "The Assessment of Learning Outcomes – Implications from Experiences in Engineering Education" in *The Journal of Hijiya University*. No.4. (Hijiya University Center for the Study of Higher Education). pp.89-101. (in Japanese)

Kishimoto, K. and Fukahori, S. 2010. "The State and Future Directions of Undergraduate Education." in *Science of Machine*. No.62 Vol.1. pp.117-126. (in Japanese).

Nusche, D. Assessment of Learning Outcomes in Higher Education: A Comparative Review of Selected Practices. *OECD Education Working Paper* No.15. Translated by Fukahori, S. (2008). National Institute for Educational Policy Research. pp.1-57.

平成25年度教育改革国際シンポジウム
TUNING-AHELO
コンピテンス枠組の共有と水準規定によるグローバル質保証

平成26年 3月

国立教育政策研究所

〒100-8951 東京都千代田区霞が関3丁目2番2号

平成25年度教育改革国際シンポジウム
TUNING-AHELO
コンピテンス枠組の共有と水準規定によるグローバル質保証

平成26年 3月

国立教育政策研究所

〒100-8951 東京都千代田区霞が関3丁目2番2号



リサイクル適性 

この印刷物は、印刷用の紙へ
リサイクルできます。