

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

深堀 聰子

(国立教育政策研究所)

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ィージビリティ・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」および専門分野「経済学」「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)。

一般的技能	専門分野 ー経済学ー	専門分野 ー工学ー
コロンビア	ベルギー(FI.)	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)

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きを整理する。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。

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なかで、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なかで、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。

そうした情勢のなかで AHELO-FS は構想され、次の 2 つの目的を掲げて着手された。第一に、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。大学教育をとお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めざされた。

第二に、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めざす取り組みである点について、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、成績評価に関連づ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めざされた。

こうした研究目的にむけて、二段階の研究デザインが構想された。第 1 フ

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めざされた。すなわち、「一般的技能」「経済学」「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、および「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もと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フェーズの結果にもと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めざされた。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、および教員調査もあわせて企画された（OECD, 2012, pp.77-95）。

こうしたサンプル・デザインから明らかなように、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もと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き

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)の協議にもとづく重層的な意思決定の仕組みがとられ、そのもとで事務局 (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およ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もとに推進された。

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、実査にもと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4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」の取り組みを分担した米国のCAE (Council for Aid to Education)、「経済学」分野の取り組みを分担した米国のETS (Educational Testing Service)、「工学」分野の取り組みを分担したACER、日本の国立教育政策研究所 (NIER)、およ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」の取り組みを分担したオランダのCHEPS (Centre for Higher Education Policy Studies)および米国の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、教育調査およ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)。

つぎに、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もとに、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、およ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 の望ましいあり方、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活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。この調査研究をとおして、AHELO-FS とい

¹ フィレンツェ大学は、76ヶ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。

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および教育改善努力の一環として位置づけ、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あ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）。

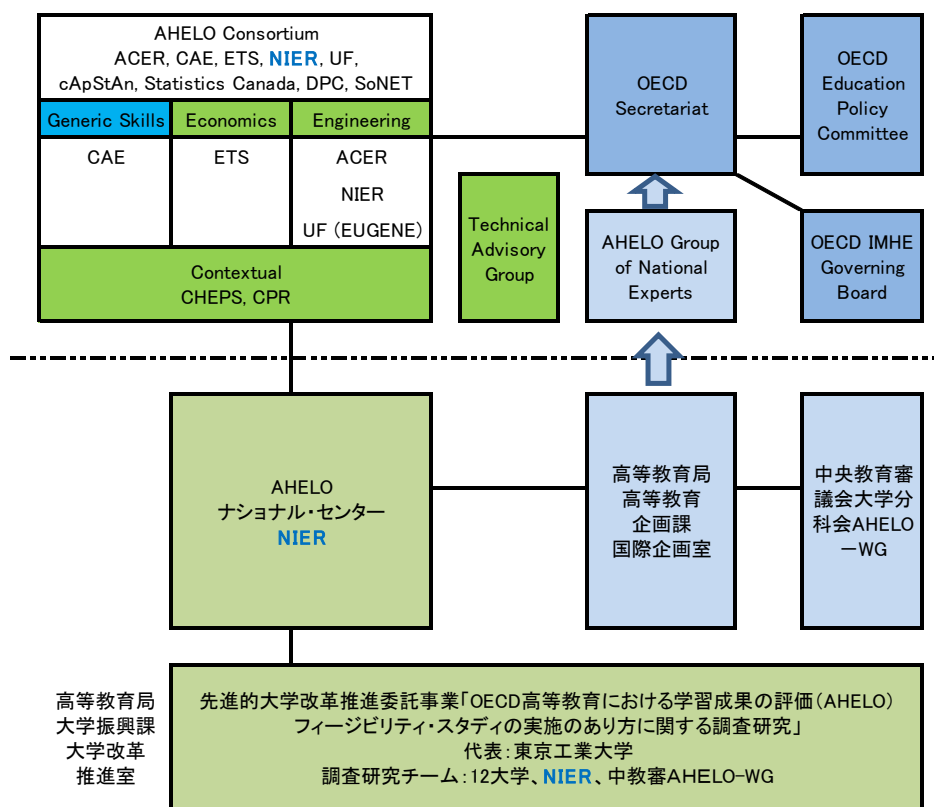


図2 AHELO-FSの組織体制（国際的・国内的）

3.2 工学分野におけるコンピテンス枠組の構築

学習成果アセスメントをとお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もとづいて策定され（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つけさせようとするコンピテンスの枠組を定義し、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り組みを指す。コンピテンス枠組を定義するにあたっては、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あたっては、大学間で共有する抽象性の高い「コンピテンス」の体系のなかから、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、科目の履修をとお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なか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、外的基準にもと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めざすボローニャ・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9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、すでに多くの国々に

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

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り組みも先駆的に進められてきていた。たとえば、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あげられる。そうしたなかで、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が定義された。ここでは、表1に示す通り、工学分野におけるコンピテンス枠組が「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5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にもと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とらえることとした。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)およ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)。

表1 Tuning-AHELO 工学分野におけるコンピテンス枠組

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とる能力。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。

3.3 テストの実施

前述したとおり、AHELO-FS の実査は二段階で実施した。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、小規模の実査にもとづく質的検討をとおして明らかにすることがめざされた。

テストは、問題プールのなかから多肢選択式問題20問、記述式問題1問を組み合わせ、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 フェーズの結果にもと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、大規模の実査にもとづく量的検討をとおして明らかにすることがめざされた。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めざされた。

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もとづいて組み合わせた 18 種類のパターンのなかから、学生ごとに異なる組み合わせを出題し、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%）の参加のもとに、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）および教員調査（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うえ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および採点者のトレーニングにとく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）にもとづいて、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

点ループリックの問題点を整理してコメントした。たとえば、正答とみなされ得る解答の観点が見逃ループリックに含まれていない場合には加筆を要求したり、採点ループ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もとづいて修正した採点ループリック（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プリック（第2版）にもと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（採点ループリック第3版・確定版）。

この際、採点ループ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うえで、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もと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もと、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、採点ループリック（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%にのぼった。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とられた（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・センターの依頼にもと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、過去の調査実績にもと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で項目反応理論にもとづく調査デザインがとられたのは、大学や学生の負担をえきるだけ抑制しながら、一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もとづく分析から、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うるのかについては、AHELO-FSの枠組のなか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、未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。

第一は、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なか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。

第二は、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5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が採用されたが、自校の学生はこのコンピテンス枠組のどの領域で秀でており、どこに課題を抱えているのか。たとえば、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。

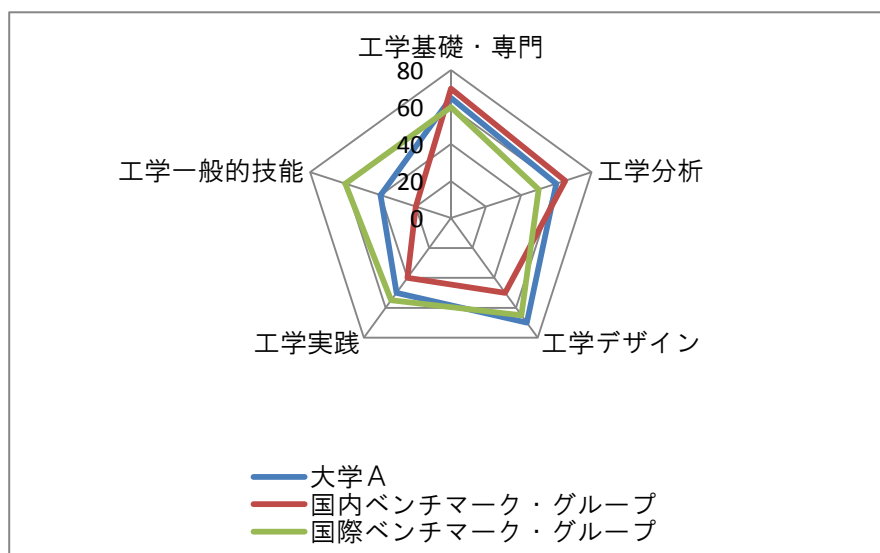


図 3. コンピテンス・プロフィール（架空データ）

第三は、学生の学習実態や大学の教育環境のあ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。大学教育のあ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。

とくに、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、試行的に手掛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、およびテ

テスト結果との関係から「テストの妥当性」を確認したうえで、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もと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%）にのぼり、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%）にのぼり、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、概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うえ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、「学部教育をとおしての成績（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うえ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。

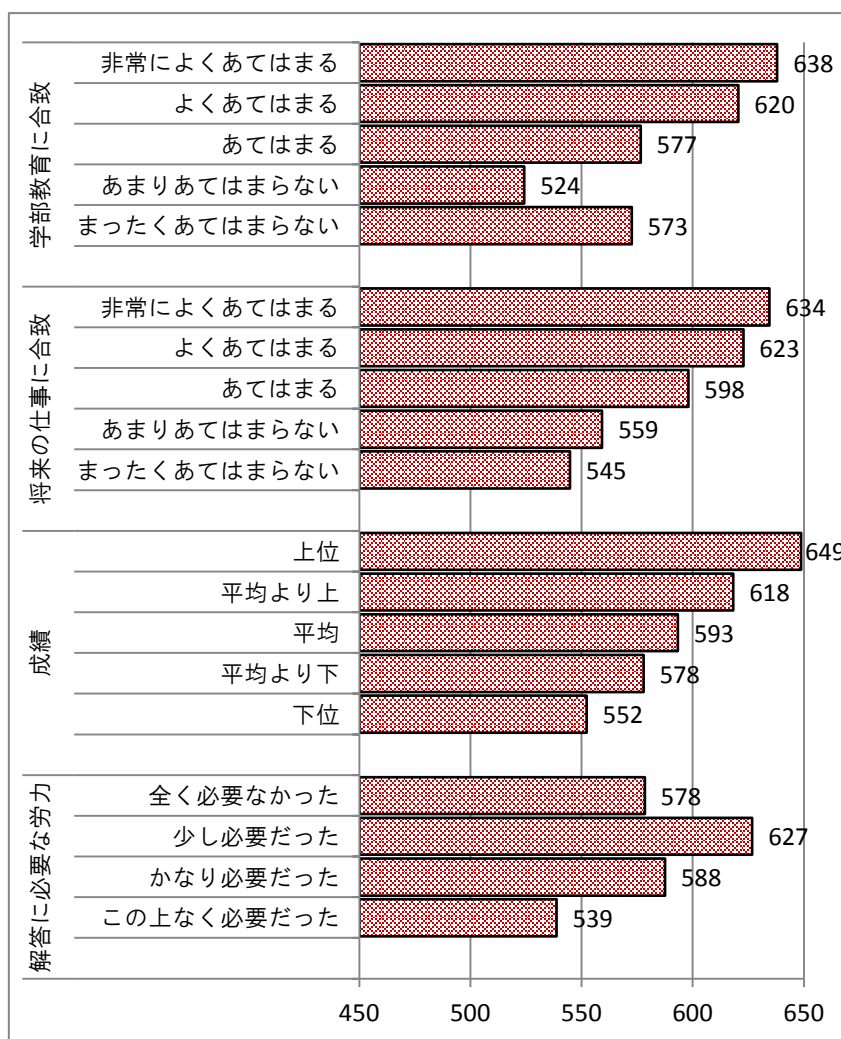


図 4. テストの妥当性：学生の印象とテスト得点

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%)にのぼり、役

立っている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（「非常によかった」9%、「よかった」44%、「普通」41%、「よくなかった」
6%）にとどまったが、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うえで有効であったことを支持している。

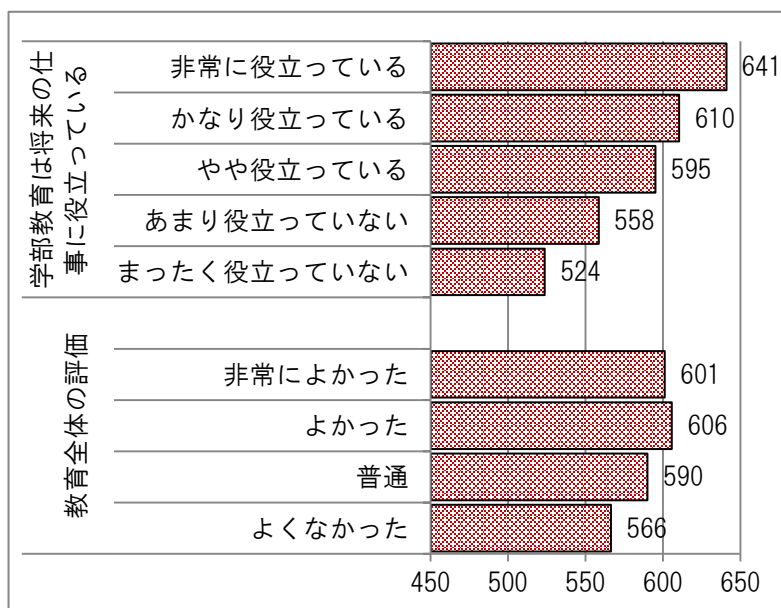


図5. テストの妥当性：学部教育の評価とテスト得点

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
しうるのだろうか。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」）にとどまっ
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。そこでは「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、一日平均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わ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もっとも高かった。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。

「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、とくに一週間に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。

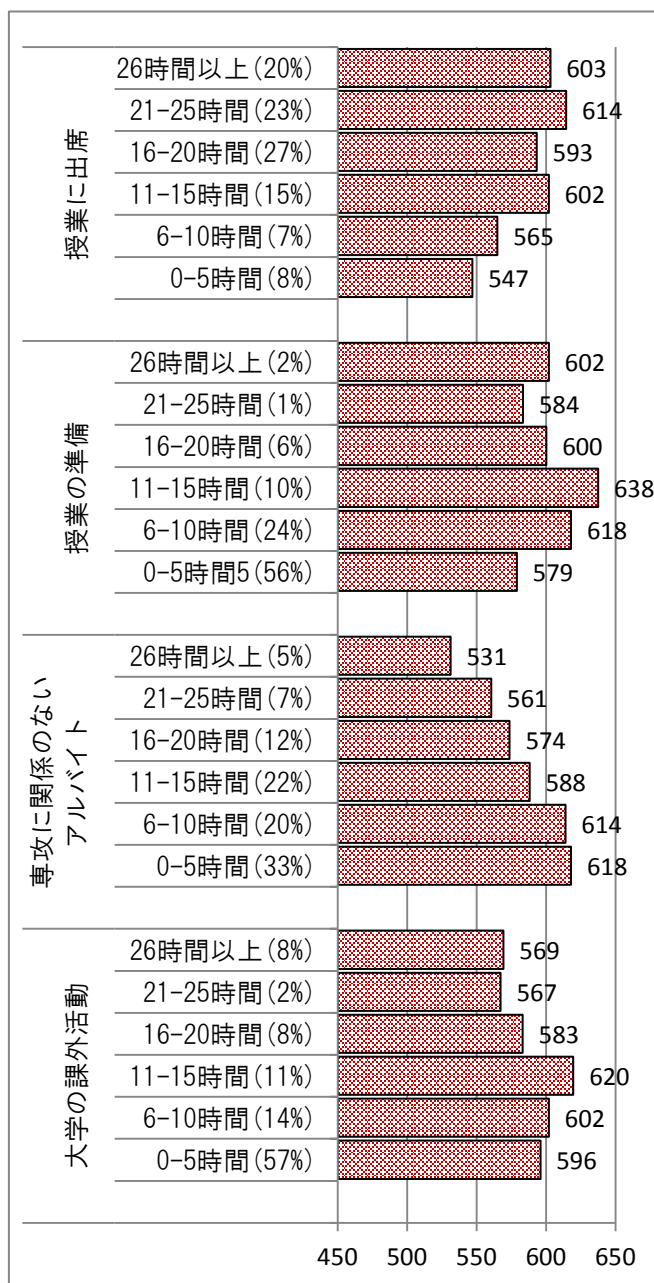


図 6. 時間の使い方とテスト得点

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、テスト得点をもっとも高いのは、「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たてることができる。

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。

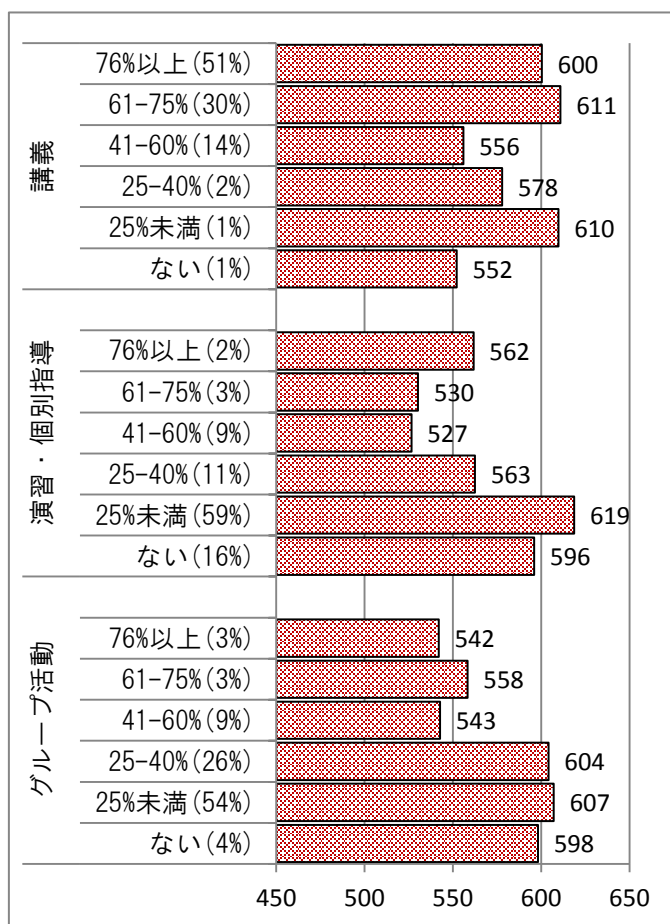


図7. 各授業形態の割合別テスト得点（全体）

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%）、および「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2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、「講義」形式の授業が主流の日本では、未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とら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。

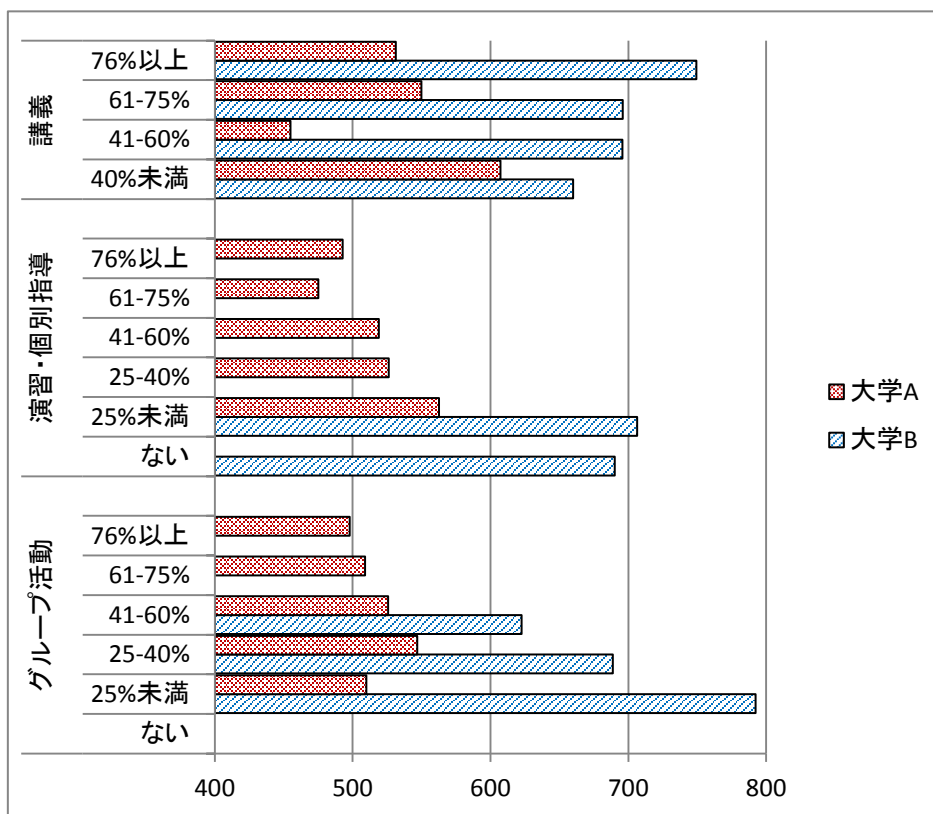


図 8. 各授業形態の割合別テスト得点 (大学A・B)

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あ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がさらに取り組まなければならない課題は、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、この覚書にもとづいて、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。こうした取り組みは、国際的なデータ共有の Protokol 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の取り組みをとおして、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画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。

第一に、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。

第二に、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およ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。

第三に、工学分野のコンピテンス枠組を構成する「工学分析」「工学デザイン」「工学実践」および「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、具体的な取り組みとして教育プログラムに広く反映されている状況とは未だいえない。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り組みを展開することで、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は国際的取り組みであることから、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.ieaagreements.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>)
- ◆ 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)