

Development of a Mechanical Engineering Test Item Bank to promote learning outcomes-based education in Japanese and Indonesian higher education institutions

Jeffrey S. Cross, Estiyanti Ekawati, Satoko Fukahori,
Shinnosuke Obi, Yugo Saito, Nathanael P. Tandian, and Farid Triawan*

doi: [http://dx.doi.org/10.18543/tjhe-5\(1\)-2017pp41-73](http://dx.doi.org/10.18543/tjhe-5(1)-2017pp41-73)

Abstract: Following on the 2008-2012 OECD Assessment of Higher Education Learning Outcomes (AHELO) feasibility study of civil engineering, in Japan a mechanical engineering learning outcomes assessment working group was established within the National Institute of Education Research (NIER), which became the Tuning National Center for Japan. The purpose of the project is to develop among engineering faculty members, common understandings of engineering learning outcomes, through the collaborative process of test item development, scoring, and sharing of results. By substantiating abstract level learning outcomes into concrete level learning outcomes that are attainable and assessable, and through measuring and comparing the students' achievement of learning outcomes, it is anticipated that faculty members will be able to draw practical implications for educational improvement at the program and course levels. The development of a mechanical engineering test item bank began with test

* Jeffrey S. Cross (cross.j.aa@m.titech.ac.jp), PhD in chemical engineering, is a Professor in the Tokyo Tech's School of Environment and Society.

Estiyanti Ekawati (esti@spm.itb.ac.id), PhD in engineering, is an Associate Professor at the Institut Teknologi Bandung, Indonesia.

Satoko Fukahori (fukahori@nier.go.jp), Research Director of the Department for Higher Education Research, National Institute for Educational Policy Research (NIER), Ministry of Education, Culture, Sports, Science and Technology, Japan.

Shinnosuke Obi (obsn@mech.keio.ac.jp), Dr.-Ing, is a professor in the Department of Mechanical Engineering, Keio University, Japan.

Yugo Saito (ugo.saito@gmail.com) is a program-specific assistant professor at Kyoto University's Center for the Promotion of Excellence in Higher Education, Japan.

Nathanael P. Tandian (n4th4n.t4nd14n@gmail.com, n4th4n@termo.itb.ac.id), PhD in Mechanical Engineering, is a faculty member in the Mechanical Engineering Department, Institut Teknologi Bandung - ITB, Indonesia.

Farid Triawan (triawan.f.aa@m.titech.ac.jp, farid.triawan@gmail.com), Doctor of Engineering in mechanical engineering, is a specially appointed Associate Professor (Lecturer) in the Tokyo Tech School of Environment and Society, Japan.

More information about the authors is available at the end of this article (after the bibliography list).

item development workshops, which led to a series of trial tests, and then to a large scale test implementation in 2016 of 348 first semester master's students in 9 institutions in Japan, using both multiple choice questions designed to measure the mastery of basic and engineering sciences, and a constructive response task designed to measure "how well students can think like an engineer." The same set of test items were translated from Japanese into English and Indonesian, and used to measure achievement of learning outcomes at Indonesia's Institut Teknologi Bandung (ITB) on 37 rising fourth year undergraduate students. This paper highlights how learning outcomes assessment can effectively facilitate learning outcomes-based education, by documenting the experience of Japanese and Indonesian mechanical engineering faculty members engaged in the NIER Test Item Bank project.

Keywords: Learning Outcomes; Quality Assurance; Educational Improvement; Mechanical Engineering; Japan; Indonesia; Test Item Bank.

I. Introduction

The assessment of higher education learning outcomes is an issue of increasing importance to universities, as they strive to respond to contemporary social demands on higher education to educate knowledgeable and creative problem solvers who are actively engaged in society. As we become increasingly aware of the complexity of problems that we face every day as the result of technological developments and globalization, we cannot but become increasingly concerned about how well we are preparing our future generations for an increasingly complex world.

Such heightened social expectation has propelled the expansion of higher education, leading to steep increases in public and private expenditure on higher education. As a result, universities can no longer exist as elite institutions, where the value of higher education is unchallenged, but must demonstrate their effectiveness not only in terms of their contribution to the academia, but also in terms of what their graduates know, understand, and can do upon graduation for the sake of society.

The focus on higher education learning outcomes have led to the development of learning outcomes frameworks that describe the knowledge and skills graduates are expected to acquire as the result of their enrollment in a given educational program. Examples of learning outcomes frameworks for disciplinary areas include the UK Subject Benchmark Statements,¹ the

¹ Quality Assurance Agency (QAA), "The UK Quality Code for Higher Education Subject Benchmark Statements," <http://www.qaa.ac.uk/assuring-standards-and-quality/the-quality-code/subject-benchmark-statements>.

Tuning Reference Points for the Design and Delivery of Degree Programs,² the USA Social Science Council Measuring College Learning Project's Essential Competencies,³ and the Science Council of Japan's Benchmark Statements for Curriculum Design for Subject-Specific Quality Assurance in University Education (*Bunyabetsu-Sanshokijun*),⁴ etc. Examples of learning outcomes frameworks for generic learning outcomes that describe the knowledge and skills graduates are expected to acquire regardless of their major include the Association of American Colleges and Universities' LEAP Essential Learning Outcomes,⁵ the Lumina Foundation's Degree Qualifications Profile,⁶ the Japanese Ministry of Education's Graduate Attributes (*Gakushiryoku*),⁷ etc.

Exposure to these learning outcomes frameworks has led to a general understanding and familiarity among faculty members to the concept of learning outcomes-based education. However, based on a survey and in-depth interviews at universities in the EU and USA, Birtwistle et al.⁸ have pointed out that “the discourse about the shift of paradigm is taking place to various degrees, amongst management and to a lesser extent with academic staff, but much less amongst students” and that there is a disturbing “disconnect between the rhetoric, political ambitions and reality.”

² Tuning Educational Structures in Europe, “Subject Areas,” <http://www.unideusto.org/tuningeu/subject-areas.html>.

³ Richard Arum, Josipa Roksa, and Amanda Cook, *Improving Quality in American Higher Education: Learning Outcomes and Assessments for the 21st Century* (San Francisco, CA: Jossey-Bass, 2016).

⁴ Science Council of Japan, “Benchmark Statements for Curriculum Design for Subject-Specific Quality Assurance in University Education,” (Tokyo: Science Council of Japan, 2012); original title in Japanese 「大学教育の分野別質保証のための教育課程編成上の参照基準について」 (pronounced as “Daigaku-kyoiku-no-bunyabetsu-shituhosho-no-tame-no-kyoiku-katei-henseijono-sanshokijun-ni-tsuite”), <http://www.scj.go.jp/ja/member/iinkai/daigakuhosyo/daigakuhosyo.html>.

⁵ Association of American Colleges and Universities, “LEAP Essential Learning Outcomes,” <https://www.aacu.org/leap/essential-learning-outcomes> (accessed 5 September, 2017).

⁶ Lumina Foundation, “Degree Qualifications Profile,” (Indianapolis, IN: Lumina Foundation, 2014) <https://www.luminafoundation.org/files/resources/dqp.pdf>.

⁷ Ministry of Education, Culture, Sports, Science and Technology (MEXT) Central Council for Education, “Towards the Construction of Bachelor Degree Programs,” (Tokyo: MEXT, 2008), original title in Japanese 「学士課程教育の構築に向けて（答申）」 (pronounced as follows: “Gakushikatei-kyoiku-no-kochiku-ni-mukete (Toshin)”), http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm.

⁸ Tim Birtwistle, Courtney Brown, and Robert Wagenaar, “A long way to go...A study on the implementation of the learning-outcomes based approach in the EU and USA,” *Tuning Journal for Higher Education* 3, no. 2 (2016): 429-463, doi: 10.18543/tjhe-3(2)-2016pp429-463.

Japan is not an exception. According to a survey conducted by the National Institution for Academic Degrees and University Evaluation,⁹ very few universities and academic societies responded that their members/faculty were referencing the above mentioned Benchmark Statements proposed by the Science Council of Japan. Hirota¹⁰ attributes this systemic obstinacy to three factors. There is a general tendency among academics to reject quality assurance initiatives firstly because they devalue indicators of quality introduced by external actors as partial, spurious, and irrelevant for measuring academic quality. Secondly, academics reject quality assurance initiatives because they see it as “a threat to academic freedom and autonomy.” They see it as mechanism for strengthening the power of the government and institutional managerial teams, which invade faculty rights to determine what to teach and how to assess student learning. Thirdly, academics currently lack the expertise to collaborate effectively in systemic educational reform. Due to the over-fragmentation of the academic disciplines, faculty members do not necessarily have holistic understandings of their disciplines or their unique roles within the disciplines and in relation to other disciplines. Faculty members are also in general unfamiliar with educational theories and methods. This present condition of the academia has inhibited constructive discussion among faculty about learning outcomes and the pedagogical approaches applicable to learning outcomes-based education.

While such observations are to certain degrees persuasive, and sociologically interesting, they provide little practical implication to those seeking to provide a learning outcomes-based education. How can we facilitate learning outcomes-based education? What are the effective conditions for mobilizing faculty to engage in discussion about learning outcomes, and in the end, embrace learning outcomes based education? In particular, what is the role of learning outcomes assessment, and how can it be utilized to improve the quality of education?

We focus on learning outcomes assessment because it forces us to reach concrete level agreement about what we expect our students to know, understand, and be able to do, which is ultimately necessary to substantiate

⁹ National Institute for Academic Degrees and University Evaluation (NIAD-UE), “A Research on the Quality Assurance of Disciplinary Education in Japanese Universities,” (Tokyo: NIAD-UE, 2016), original title in Japanese 「我が国における大学教育の分野別質保証の在り方に関する調査研究報告書」 (pronounced as: “Wagakuni-niokeru-daigaku-kyoiku-no-bunyabetsu-sanshokijun-no-arikata-ni-kansuru-chosa-kenkyu”), http://www.niad.ac.jp/n_shuppan/project/_icsFiles/afieldfile/2016/07/14/no09_nr16-0714.pdf.

¹⁰ Teruyuki Hirota, “Why Front Line Academics Reject Reform-Towards Utilizing Subject Specific Benchmark Statements,” *University Evaluation Studies* 15 (2016): 37-46.

learning outcomes-based education. At the course level, faculty members are expected to clarify attainable and assessable learning outcomes to be achieved by the students, and then to plan and implement educational practices that will most effectively facilitate students' learning. This "backward design," as articulated by Wiggins and McTighe¹¹ applies also at the program level, where faculty teams are expected to clarify the learning outcomes to be achieved by completing the total set of courses that constitute the program. Learning outcomes defined at the program level are by design more abstract than at the course level, because they must be applicable to different courses that focus on different sets of disciplinary knowledge and skills. Nevertheless, program level learning outcomes must be grounded on concrete level understandings of how they will be articulated within relevant courses, if they are to become trusted indicators of quality. Learning outcomes assessment, when based on a learning outcomes framework can make explicit the connection between course/concrete level learning outcomes and program/abstract level learning outcomes. This substantive understanding of how a course can contribute to the learning goals of a program is of critical importance to those seeking to provide a quality learning outcomes-based education.

The purpose of this paper is to highlight how learning outcomes assessment can effectively facilitate learning outcomes-based education, by documenting the experience of Japanese and Indonesian mechanical engineering faculty members who have been engaged in a collaborative project aimed at developing a "test item bank" at the National Institute for Educational Policy Research (NIER),¹² the Tuning National Center for Japan.

The experience of engineers provides rich implications to other disciplines for two reasons. First, the engineers have been top runners in developing a shared understanding of disciplinary learning outcomes, endorsed by a global network of program accreditation agencies; i.e. the International Engineering Alliance (IEA)'s Graduate Attributes¹³, and the European Network for the Accreditation of Engineering Education

¹¹ Grant Wiggins, and Jay McTighe, *Understanding by Design*, expanded 2th ed. (Alexander, Virginia: Association for Supervision & Curriculum, 2005).

¹² National Institute for Educational Policy Research (NIER), "Tuning Test Item Bank Project," <http://www.nier.go.jp/tuning/centre/kikai.html>.

¹³ International Engineering Alliance, "Graduate Attributes and Professional Competencies," <http://www.ieagreemts.org/assets/Uploads/Documents/Policy/Graduate-Attributes-and-Professional-Competencies.pdf>.

(ENAAEE)'s EUR-ACE.¹⁴ Graduates of engineering programs that have been accredited by agencies within the network are considered to have acquired the learning outcomes defined in the frameworks, fulfilling the foundational requirement for those who wish to become a Professional Engineer in signatory countries. Much is to be learnt from what the engineers have already accomplished, and the challenged they are currently facing.

Secondly, while engineering is a highly structured discipline with specified sets of knowledge and skills graduates are expected to acquire, it is also a discipline that has traditionally placed strong focus on the application of knowledge and skills, and on higher order analytical and problem solving skills. Additionally, engineering programs are increasingly emphasizing the importance of being able to operate meta-cognitive knowledge necessary to deliberate the role of engineers in society. Such emphases are indeed central to the learning outcomes aspired by all disciplines in contemporary higher education.

This paper consists of five sections. Following this introduction, the paper will outline the background, purpose, and the methodology of the NIER Test Item Bank. In the third section, the paper will describe test item development, implementation, and preliminary analysis of test results in Japan and Indonesia. The fourth section will discuss implications drawn from international benchmarking. Finally, the fifth section will discuss the limitations and future directions of the NIER Test Item Bank project.

II. The NIER Test Item Bank – background, purpose and methodology

II.1. Background and purpose – the Japanese perspective

The NIER Test Item Bank is a spin-off project of the OECD-AHELO Feasibility Study, which was an international study conducted between 2009 and 2012. The study aimed to test whether or not it was practically and scientifically feasible to develop and conduct a test that assesses what university graduates know and can do upon completion of their bachelor

¹⁴ European Network for the Accreditation of Engineering Education, “EUR-ACE Framework Standards and Guidelines (EAFSG),” (European Network for Engineering Accreditation, 2015) <http://www.enaee.eu/wp-assets-enaee/uploads/2017/11/EAFSG-Doc-Full-status-8-Sept-15-on-web-fm.pdf>.

degree programs. The study was conducted in the strands of generic skills, economics, and engineering, involving 17 countries, 248 institutions, and over 23,000 students. Japan participated in the engineering strand, with the participation of 12 universities and 504 students.

In the engineering group, a learning outcomes framework was developed by an international team of engineering experts by extracting common elements of the IEA Graduate Attributes and ENAEE EUR-ACE frameworks, taking into account the career paths of engineering students and the expectations of stakeholders.¹⁵ The abstract level learning outcomes defined in the framework were then articulated into concrete level measurable learning outcomes within the context (scenario) of each test item. After the test items and scoring guides were verified based on pilot implementations, they were implemented in large scale, and collectively scored by engineering experts.¹⁶

The OECD concluded the Feasibility Study by stating that it was in fact feasible to develop an international assessment of higher education learning outcomes.¹⁷ The Japanese team, however decided to continue the project, focusing on the value of the endeavor. First, the team deliberated that the collaborative development of test items and joint scoring proved to be invaluable in generating a consensus among faculty members regarding what students are expected to know, understand, and be able to do, and in substantiating abstract level learning outcomes into attainable and assessable learning outcomes. Secondly, the team articulated that once faculty members reached consensus on the logic of scoring within a given scenario, they were able to apply the same logic to different scenarios, reaching high agreement in scoring results at accelerating speed. The implication that can be drawn from this experience is that occasional faculty participation in external learning outcomes assessment exercises may provide long lasting insights that will guide the design and delivery of courses and degree programs. A learning outcomes assessment exercise

¹⁵ Organization of Economic Cooperation and Development (OECD), “A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering,” *OECD Education Working Papers* 60 (2011), <http://dx.doi.org/10.1787/5kghtchn8mbn-en>.

¹⁶ Karine Tremblay, Dianne Lalancette, and Deborah Roseveare, “OECD-AHELO Feasibility Study Report Volume 1 - Design and Implementation” (Paris: OECD, 2012), <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>.

¹⁷ Organization of Economic Cooperation and Development (OECD), “OECD-AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences” (Paris: OECD, 2013), <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf>.

can be an effective faculty development model for encouraging the implementation of the learning outcomes-based approach.^{18,19,20}

It is important to mention that in Japan, engineering program accreditation has not been adopted uniformly. Although there is an engineering accreditation body, the Japan Accreditation Board of Engineering Education (JABEE)²¹ which is a signatory of the Washington Accord, engineering program accreditation is voluntary, and many of Japan's top research university's engineering programmes are in fact un-accredited. This is because graduating from an accredited engineering program is not necessarily required to find employment at Japan's leading multinational manufacturing corporations. However, countries from which Japan accepts international students, such as Malaysia do require their engineering students to graduate from accredited programs if they were to return and practice engineering as their profession. By not being accredited, Japanese engineering programs are under-serving their internationally mobile students, which is an issue that is becoming increasingly put to attention at high-level meetings.

In many engineering fields in Japan, there has been little incentive for graduates to pursue additional certificates such as a Professional Engineer. This is because most large manufacturing corporations offer lifetime employment opportunities. Once a graduate is hired, there are very few opportunities in which they are required to formally demonstrate their professional qualifications. Another reason may be the fact that major corporations²² have traditionally had their own effective in-house training

¹⁸ National Institute for Educational Policy Research (NIER), "The National Institute for Educational Policy Research International Symposium for Educational Reform 2013: The TUNING-AHELO Global Quality Assurance through Sharing Competence Frameworks and Degree Level Specifications" (Tokyo: NIER, 2013), http://www.nier.go.jp/06_jigyou/symposium/i_sympto25/.

¹⁹ Satoko Fukahori, ed., "An Analytical Report of the Results of AHELO Feasibility Study" (Tokyo: National Institute for Educational Policy Research, 2014); original title in Japanese: 「AHELO調査結果の分析に関する研究会（研究成果報告書）」 (pronounced as follows: "AHELO Chosa-kekka-no-bunseki-ni-kansuru-kenkyukai (Kenkyu-seika-hokokusho)"), <http://www.nier.go.jp/koutou/ahelo/2014/report.pdf>.

²⁰ Satoko Fukahori, "Competence-based Quality Assurance of University Education - Lessons Learnt from the OECD-AHELO Feasibility Study," in *Educational Studies in Japan: International Yearbook* Vol. 8 (Japanese Educational Research Association, 2014), 49-61.

²¹ Japan Accreditation Board of Engineering Education, <http://www.jabee.org/english/> (accessed 5 September, 2017).

²² National Research Council, *Engineering Tasks for the New Century: Japanese and U.S. Perspectives* (Washington, DC: The National Academies Press, 1999), <https://doi.org/10.17226/9624>.

programmes for their new employees, which allowed them to be relatively unconcerned about what students actually learnt in their specialization of study at universities. Under these circumstances, it is not the case in Japan that all soon to be engineering graduates undergo standardized engineering examination. While the Japanese university entrance examination serves as the entry level screening device for admitting students of the highest academic aptitude to top-tier universities, there is no formal mechanism at the exit level for measuring and comparing the achievement of learning outcomes of engineering graduates across institutions.

It is also important to note that in Japan, unlike in the United States and European universities, the offering of faculty development programs are somewhat limited in scope. Many engineering faculty members actually learn how to teach from having been taught by their advisors, and in many cases follow this same approach when teaching their own students in their research laboratories and courses. There have been very few opportunities for faculty members from different institutions to engage in discussion about educational improvement.

However, globalization is forcing Japanese universities to change. For example, from the 1990s, there has been a growing consensus within the ASEAN region that engineers from a given country should be able to practice their profession in member countries. Furthermore, ASEAN countries and Japan, because of their proximity and interlinked economies related to manufacturing, share the common goal of developing highly skilled and innovative human resources in all fields of engineering. This has led to the strengthening of cooperation among engineering programs in leading universities in ASEAN countries and Japan.

The OECD-AHELO Feasibility Study coincided with growing recognition among engineering leadership regarding the necessity to embrace the learning outcomes-based approach in the face of globalization. Within this context, the NIER Test Item Bank project was launched in 2014 in the field of mechanical engineering, with the full support of the Japanese OECD-AHELO Feasibility Study team. The purpose of the project is to develop among engineering faculty members, common understandings of engineering learning outcomes, through the collaborative process of test item development, scoring, and sharing of results. By substantiating program/abstract level learning outcomes into attainable and assessable course/concrete level learning outcomes, and through measuring and comparing the students' achievement of learning outcomes, faculty members will be able to draw practical implications for educational improvement at the program and course levels. Only when we are able to convince faculty members of its

effectiveness, will we be able to witness a systemic shift towards the adoption of the learning outcomes-based approach.

II.2. *Motivation for Test Item Bank participation at ITB in Indonesia*

Institut Teknologi Bandung (ITB) is one of the leading technical universities in Indonesia, especially in the fields of sciences and engineering. ITB plays a significant role in shaping the policies of higher education in Indonesia, and has become a benchmark for other universities in Indonesia for improving their curriculum.

Since 2013, ITB stated its commitment to developing the “Outcome Based Education (OBE)” for all its undergraduate programs. The programs responded to this policy by developing and implementing OBE in various stages, depending on their readiness. By 2017, international accreditation agencies, such as ABET, ASIIN, JABEE and KAAB had recognized the implementation of OBE in 25 ITB programs. These include the 2016 ASIIN accreditation for three undergraduate programs within the Faculty of Mechanical and Aerospace Engineering. These are the Mechanical Engineering Program (MEP-ITB), the Aeronautics and Astronautics Program and the Material Engineering Program.

The development and implementation of OBE is an ongoing process. Therefore, the Faculty of Mechanical and Aerospace Engineering is continuously developing their methods to assess the achievements of learning outcomes, to evaluate and to determine the required improvements in curriculum and relevant academic processes. Before 2012, these programs utilized the annual tracer studies to assess the achievements of learning outcomes based on their alumni performances in the workforce. Henceforth, the Mechanical engineering program also utilizes its own Comprehensive Tests for the same purpose. Since the Comprehensive Test directly evaluates the students’ understandings of the scientific and engineering aspects of the respective programs, it is critically important to find a way to improve its accuracy and effectiveness.

The participation of MEP-ITB in NIER’s Test Item Bank Project in 2015 was a valuable opportunity to assess the students’ achievement of the program’s learning outcomes in an international setting. The programs considered the project as a “calibrator/benchmark” for the Comprehensive Test at MEP-ITB. Results of the implementation of the large-scale implementation are summarized below. A preliminary comparison of the NIER Test Item Bank implementation results in Japan and at ITB are also discussed.

II.3. Methodology

The basic methodology for the NIER Test Item Bank is based on the OECD-AHELO Feasibility Study model, as discussed earlier.²³

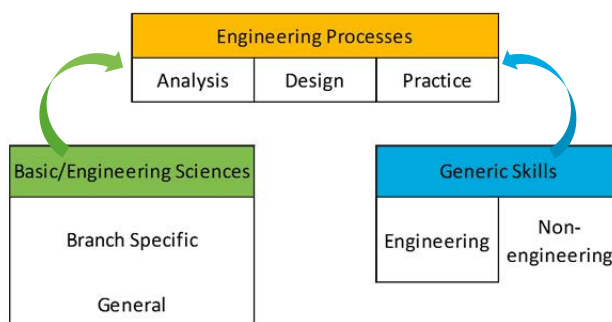


Figure 1

The Engineering Learning Outcomes Framework

The Engineering Learning Outcomes Framework as shown in Figure 1 and Table 1. consists of learning outcomes categorized into the following five areas: basic and engineering sciences, engineering generic skills, engineering analysis, engineering design, and engineering practice. As stated earlier, these learning outcomes were conceptualized based on the existing IEA Graduate Attributes and ENAEE EUR-ACE frameworks. Additionally, in order to reflect the organization of content areas taught in Japanese universities, basic and engineering sciences was categorized into the following seven areas: fundamental mathematics, fundamental physics, materials, motions, energy, informatics, and mechanical process. Multiple-choice questions (MCQ) were developed mainly to measure the mastery of basic and engineering sciences, whereas constructive response tasks (CRT) were developed to mainly measure “how well students can think like an engineer,” or how well students can operate engineering generic skills, engineering analysis, engineering design, and engineering practice.

²³ Group of National Experts on the AHELO Feasibility Study. “Engineering Assessment Framework,” (AHELO Feasibility Study 8th meeting of the AHELO GNE Paris, 18-19 November 2011), [http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne\(2011\)19/ANN5/FINAL&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne(2011)19/ANN5/FINAL&doclanguage=en).

Table 1
The Engineering Learning Outcomes Framework

Engineering Generic Skills	
EGS1	The ability to function effectively as an individual and as a member of a team.
EGS2	The ability to use diverse methods to communicate effectively with the engineering community and with society at large.
EGS3	The ability to recognise the need for and engage in independent life-long learning.
EGS4	The ability to demonstrate awareness of the wider multidisciplinary context of engineering.
Basic and Engineering Sciences	
BES1	The ability to demonstrate knowledge and understanding of the scientific and mathematical principles underlying their branch of engineering.
BES2	The basics of mathematics include differential and integral calculus, linear algebra, and numerical methods.
BES3	The ability to demonstrate a systematic understanding of the key aspects and concepts of their branch of engineering.
BES3	The ability to demonstrate comprehensive knowledge of their branch of engineering including emerging issues; high-level programming, solid and fluid mechanics; material science and strength of materials; thermal science; thermodynamics and heat transfer; operation of common machines; pumps, ventilators, turbines, and engines.
Engineering Analysis	
EA1	The ability to apply their knowledge and understanding to identify, formulate and solve engineering problems using established methods.
EA2	The ability to apply knowledge and understanding to analyse engineering products, processes and methods.
EA3	The ability to select and apply relevant analytic and modelling methods.
EA4	The ability to conduct searches of literature, and to use data bases and other sources of information.
EA5	The ability to design and conduct appropriate experiments, interpret the data and draw conclusions.
EA6	The ability to analyse mass and energy balances, and efficiency of systems; hydraulic and pneumatic systems; machine elements.
Engineering Design	
ED1	The ability to apply their knowledge and understanding to develop designs to meet defined and specified requirements.
ED2	The ability to demonstrate an understanding of design methodologies, and an ability to use them.
ED3	The ability to carry out the design of elements of machines and mechanical systems using computer-aided design tools.
Engineering Practice	
EP1	The ability to select and use appropriate equipment, tools and methods.
EP2	The ability to combine theory and practice to solve engineering problems.
EP3	The ability to demonstrate understanding of applicable techniques and methods, and their limitations.
EP4	The ability to demonstrate understanding of the non-technical implications of engineering practice.
EP5	The ability to demonstrate workshop and laboratory skills.
EP6	The ability to demonstrate understanding of the health, safety and legal issues and responsibilities of engineering practice, the impact of engineering solutions in a societal and environmental context, and commit to professional ethics, responsibilities and norms of engineering practice.
EP7	The ability to demonstrate knowledge of project management and business practices, such as risk and change management, and be aware of their limitations.
EP8	The ability to select and use control and production systems.

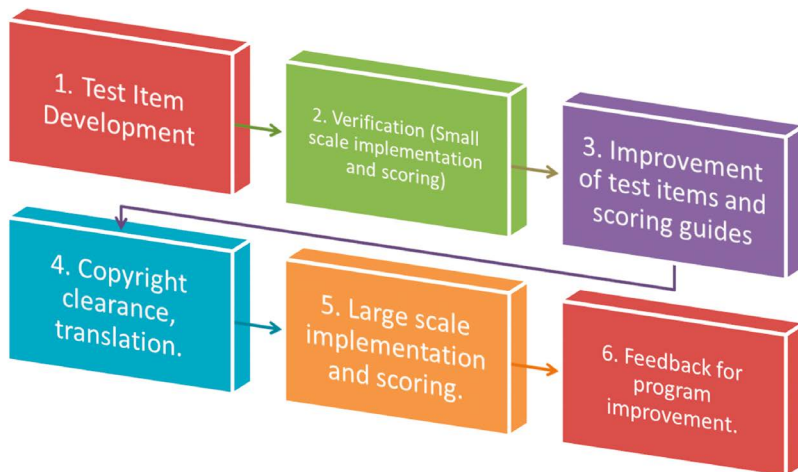


Figure 2

The Test Item Bank Development Process with Feedback for Improvement

The test item development follows the six-step Plan-Do-Check-Act (PDCA) approach shown in Figure 2. First, the test items and scoring guides proposed by members of the team were reviewed, discussed, and revised in face-to-face workshops. Second, the test items were tested in small-scale trials, to verify whether or not the test items made sense to the students, that it was tapping on the knowledge and skills intended to be measured, and that the tasks were not too difficult or too easy. The third step involved the improvement of test items and scoring guides, based on the results of the small-scale trial assessments. As the fourth step after verification, permissions were obtained to use photos, diagrams, charts, etc. protected by copyright licenses. Then, the test items were translated into English. The fifth step involved the large-scale implementation and scoring of the tests at universities in Japan and abroad. As the sixth step, feedback reports consisting of information meant to facilitate educational improvement were compiled and sent to the project team, participating universities, and participating students.

The operational model of the NIER Test Item Bank is shown in Figure 3. There is a central committee consisting of a total of 54 members at 22 institutions, and a secretariat at NIER. Members work in three regional hubs to propose test items, which are reviewed, discussed, and revised in the central committee held several times a year, and then brought back to the hubs for trial and full-scale implementation. This hub and spoke type of

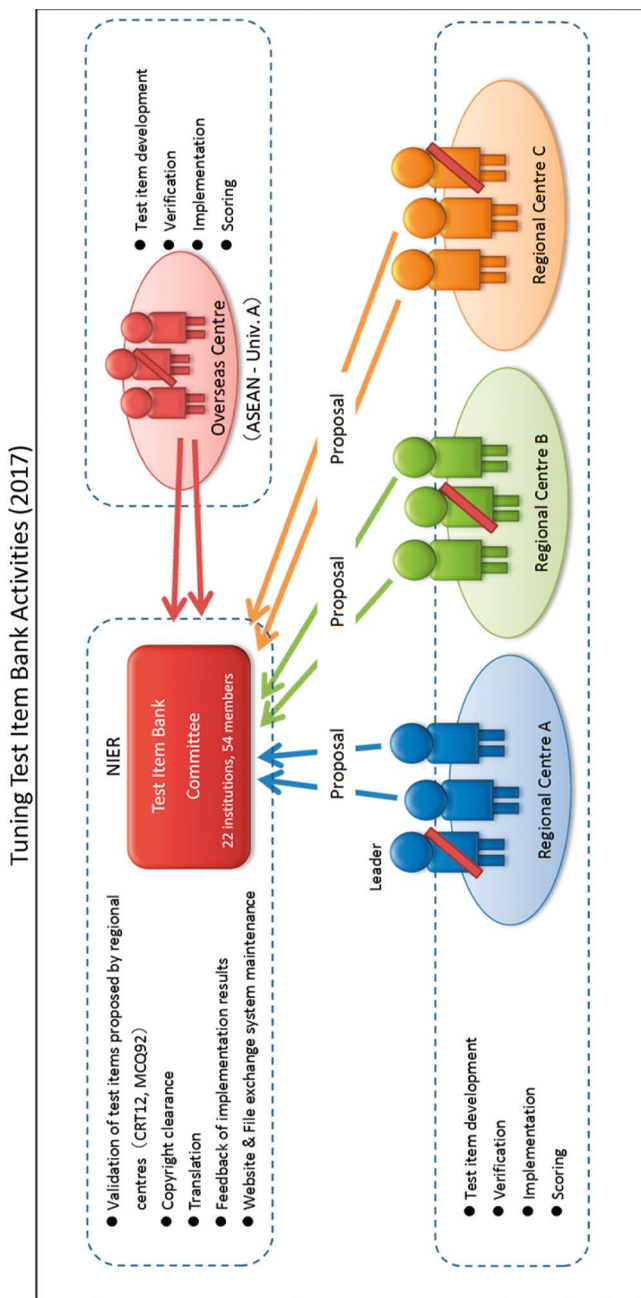


Figure 3
The Operational Model of the NIER Mechanical Engineering Test Item Bank within Japan and Indonesia

model is also applied in the case of ITB in Indonesia. Figure 4 shows Japanese and Indonesian members engaged in test item development in a central committee meeting in September 2014 in Japan.

It should be noted that the core concept behind the test item bank is that faculty members who produce the test items will be able to use items from the test item bank for assessment in their universities, under the agreement that they will share the assessment results with the secretariat, who will compile a test item profile indicating the level of difficulty, the distribution of test scores, etc. As of March 2017, 94 MCQs and 12 CRTs are being shared among members using the “Direct Cloud Box” system. Participation in the test item bank is currently free of charge.



Figure 4

The Test Item Development Faculty Workshop
in Hayama, Kanagawa, Japan, September 2014

III. Test Item development, implementation, and analysis of results

III.1. Test item development – a sample item

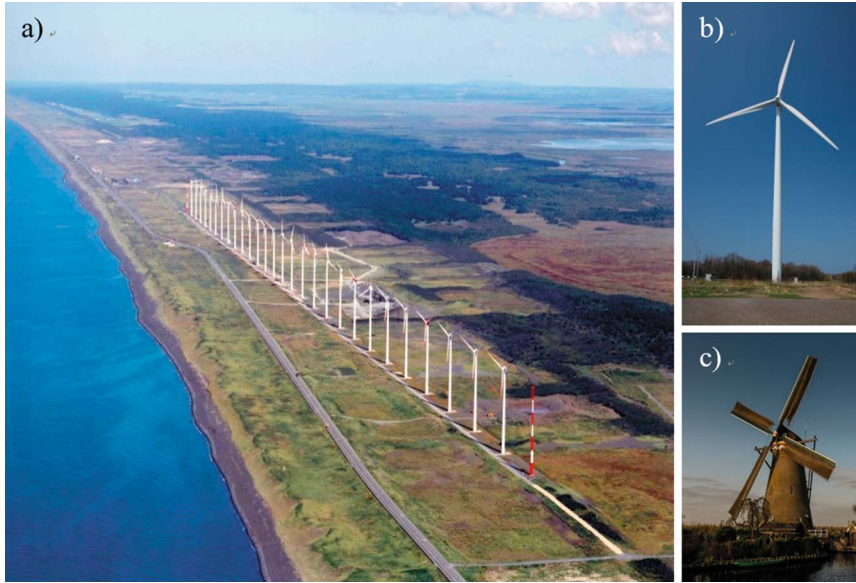


Figure 5

Wind Power Generation Sample Item showing a) wind farm,²⁴
b) wind turbine,²⁵ and c) wind mill²⁶

Figure 5 from a sample constructive response task featuring “wind power generation” is shown in Figure 5 and the scoring guide for the first task is shown in Table 2 showing a wind farm, wind turbine for electrical power generation, and windmill for pumping water. In this example, photographs of a wind farm which generates electricity using wind turbines as well as the characteristics of a windmill and wind turbine are shown. There are pronounced differences between the turbine and mill related to their design, construction and functionally. Based on mechanical engineering education, it is expected that students would be able to:

²⁴ Otonrui Wind Farm, photograph courtesy of Horonobe City, Japan.

²⁵ <http://sozai-free.com/sozai/01541.html>.

²⁶ Martijn Roos, <http://free.photos.gatag.net/2014/11/07/040000.html>.

Table 2
The Scoring Guide for Wind Power Generation shown in Fig. 5

<p>Learning outcomes to be assessed: The ability to analyze and to examine the function and efficiency of machines by applying basic knowledge of mechanical engineering by explanation of the locational condition of a wind farm.</p>	<p>Underlying competences: BES2: The ability to demonstrate a systematic understanding of the key aspects and concepts of their branch of engineering. EA2: The ability to apply knowledge and understanding to analyze engineering products, processes and methods. EA6: The ability to analyze mass and energy balances, and efficiency of systems.</p>
<p>Viewpoints: Lists two features out of three below or equivalent, and explains the reasons for each of them appropriately.</p> <p>(a) The wind farm is located on flat land along a seashore and hence there is no obstacle to block the wind from flowing around the wind turbines.</p> <ul style="list-style-type: none"> • The wind kinetic energy can be utilized effectively with little loss because the wind directly blows against the wind turbines to a maximum degree. • The wind turbine blades rotate freely because the wind flows around the stationary tower and against the turbines. <p>(b) Many wind turbines are installed in one location.</p> <ul style="list-style-type: none"> • All wind turbines can be manufactured to the same design requirements because the local environment for all turbines is basically the same. This reduces the manufacturing and design costs required in designing and producing the turbines. • The cost for installation and maintenance of wind turbines is reduced because many turbines are located adjacent to each other. • The cost for installation and maintenance of accompanying facilities to recover the electric energy generated by all turbines is reduced because such facilities can be also installed on-site. <p>(c) No building or structure is located around the wind farm.</p> <ul style="list-style-type: none"> • A wind turbine can be designed specifically for the wind conditions at the location because there is no limitation on size of the wind turbine. This increases the efficiency in generating the electric energy. • There is no possibility to cause damage to the neighboring buildings or structures in case of accidents such as the collapse of wind turbine column. 	

1. By examining the locational condition or site of a wind farm for wind electrical power generation, explain two reasons why the picture shows a good location for a wind farm (Basic and engineering Sciences, Engineering Practice)
2. By comparing the shapes of the blades for a traditional windmill and a wind turbine, explain from a mechanical engineering point of view two features of blades that characterize wind turbines for wind power generation (Basic and engineering Sciences, Engineering Practice)
3. By examining the “number of blades” of a wind turbine for wind power generation, identify from a mechanical engineering point of view, three advantages of having many or few blades, and to explain why large scale wind turbines used for wind power generation often have three propeller type blades (Basic and engineering Sciences, Engineering Practice).
4. Explain the “responsibilities of a mechanical or power engineer” in designing wind turbines for wind power generation, under the following condition.

Only one year after the wind turbine was constructed, it became apparent that the wind velocity could be larger than assumed in the initial design. In order to prevent the collapse of wind turbine column due to strong wind, identify three potential countermeasures that are technically possible for different levels of risk. For each countermeasure, evaluate their advantages and/or disadvantages from a broad perspective, including the non-technical aspects (Engineering design, Engineering practice).

Following the six-step PDCA test item development and implementation cycle, this item was proposed by members of the Kanto hub, reviewed, discussed and revised by the central committee, pilot tested in multiple universities at each hub including Indonesia, and then revised based on the verification procedure.

III.2. *The 2016 test implementation in Japan and Indonesia*

The following sections describe the implementation procedures in the two countries, during the 2016 large-scale implementation, which involved 10 MCQs and one CRT (six tasks based on one scenario) featuring “mechanical process (machine tools).” The testing time was 30 minutes for the MCQs and 50 minutes for the CRT. Given that this test item is still kept confidential,

details of the item will not be discussed in this paper. The learning outcomes assessed through this implementation are shown in Table 3.

Table 3
Learning Outcomes Assessed by the Test Items

	MCQ	CRT1	CRT2	CRT3	CRT4	CRT5	CRT6
Basic and Engineering Sciences	O	O		O	O	O	
Engineering Generic Skills							
Engineering Analysis		O	O	O	O	O	O
Engineering Design							O
Engineering Practice			O		o		o

The implementation guide compiled by the secretariat defined in detail the testing procedure to be followed at every university.

Each participating student received a test kit (envelope), which included the MCQ booklet, the CRT booklet, contextual survey (questionnaire), confidentiality agreement, and answer sheets. Each student was given an identification number so that test results and questionnaire responses could be matched and analyzed.

The first 5 minutes was spent for explaining the purpose of the project and for distributing the test kit. The coordinator at each university was requested to read the prescribed instructional text, and the students were asked to sign the non-disclosure confidentiality agreement. The next 30 minutes was spent for responding to the MCQs, followed by 50 minutes spent for the CRT.

After the test, students were asked to spend 10 minutes responding to a questionnaire consisting of a series of questions about educational experiences, which may influence students' performance on the test. The testing session was closed by a 10-minute discussion, focusing on the student perspective on the validity of the test items. The entire procedure took 105 minutes with no breaks taken in between. Students in Japan and Indonesia were actually familiar with this style of testing through their experience in the university entrance examination or the Comprehensive Test.

The scoring procedure was also defined by the secretariat. The MCQs were machine scored whereas the CRT was hand scored by the faculty members of participating institutions. In each university, scores were asked to calibrate their understanding of the scoring guides in advance. However, this procedure was not sufficiently defined, which may have led to several

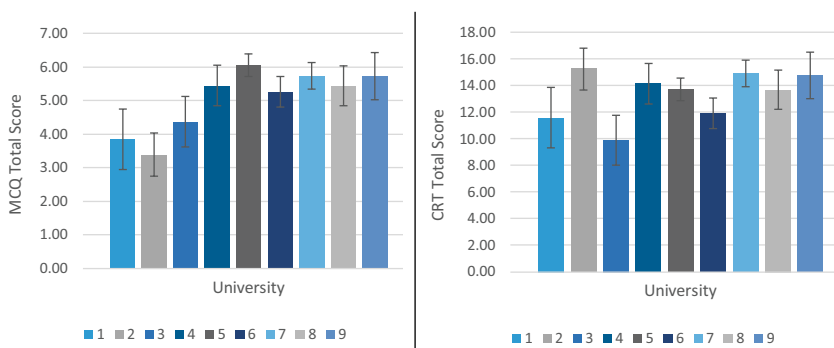
incidents of low inter-rater reliability scores, which is an issue which will be discussed in the later section. Then, in Japan, each student response was scored by two faculty members from the students' university and one faculty member from another university. At ITB, three ITB professors scored each student's response. The average of the three scores was taken to represent the score of the student. This procedure was undertaken to cancel out any individual level and institutional level bias.

Feedback reports were compiled by the secretariat and delivered to the project team, participating universities, and individual participating students.

III.2.1. Testing at Japanese universities

First semester master students who had just completed their bachelor degree programs in Mechanical Engineering in March 2016 were invited to participate. A total of 348 students in nine Japanese universities participated in the testing which took place between June and August, 2016. The number of participating students varied with institution from 22 to 98, with the average number of 39.

The test was administered only in Japanese. Students were not given any incentives for participation, but were informed that they would receive feedback reports on the test results.

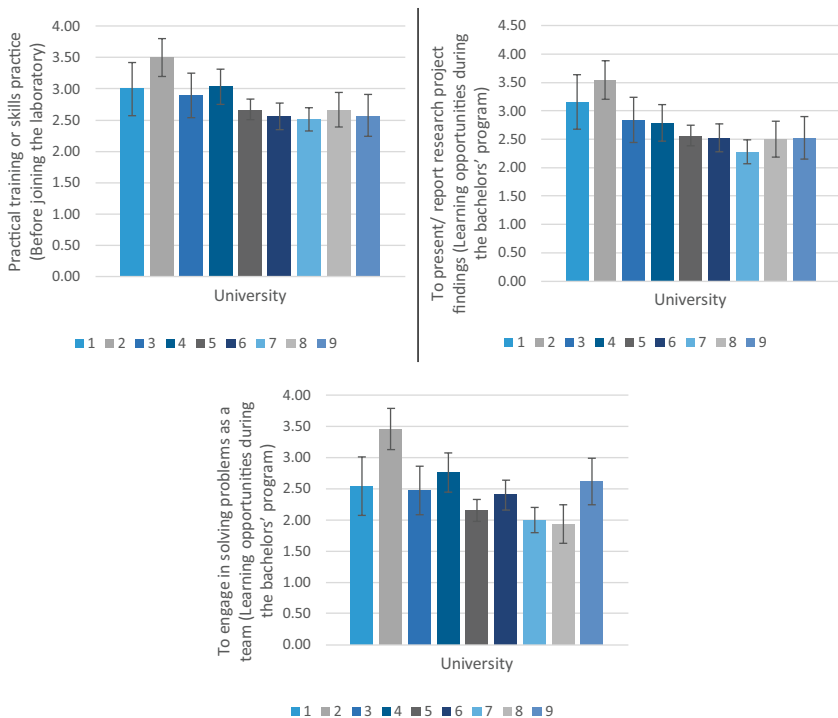


Note: Error bars represent the 95% confidence interval.

Figure 6
MCQ and CRT Results for Japanese students

Total MCQ and CRT scores at the student level revealed to be only weakly correlated ($r=0.173, p<.01$), indicating that students with higher scores on MCQs do tend to have higher scores on CRTs, but not necessarily all the time. This implies that MCQs and CRTs were measuring different types of learning outcomes that may depend to different degrees on student aptitude, and that may be fostered to different degrees through educational experiences.

Figure 6 shows the total MCQ and CRT scores at the institutional level, or the average total scores of Japanese students at nine universities. Note that students in institution 2 did not perform well on the MCQs but performed exceedingly well on the CRTs. This result brings our attention to the unique educational experiences these students have experienced that may have fostered the ability to “think like an engineer.”



Note: Error bars represent the 95% confidence interval.

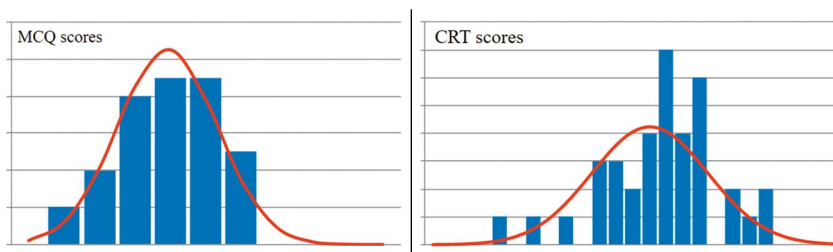
Figure 7
The Educational Experiences of Japanese Students

Figure 7 shows the students' educational experiences reported by the students in the questionnaire. Students in institution 2 responded more affirmatively that they committed themselves to “practical training or skills practice” before joining the laboratory to undertake research projects, and that they had opportunities “to present/ report research project findings” and “to engage in solving problems as a team” after joining the laboratory (Items with mean-difference effect size $d > 1.0$).

Results such as these have encouraged discussion on providing educational experiences that can be expected to enhance student performance, which is a necessary condition for educational reform to take action.

III.2.2. Testing in Indonesia at ITB

The test at ITB was conducted on September 3, 2016, at 13:00–15:00 (local time). Thirty-seven mechanical engineering students who had completed their third year (junior or sixth semester) participated in the test. Because the topic covered in the CRT was very specific to Mechanical Engineering fields, i.e. “machine tools” topic, aeronautics engineering student were not invited to participate. The test implementation was conducted only in English. All participants received a USB flash disk as a token of appreciation for their participation.



Note: Error bars represent the 95% confidence interval.

Figure 8

Distribution of MCQ and CRT Total Scores for ITB Students

The histogram of the MCQ scores for ITB students shown in Figure 8 fits well to its normalized curve. The deviation between the MCQ score histogram to its normalized curve is about 5.5%. This indicates that the number of students participated in the test was large enough resulting in the distribution

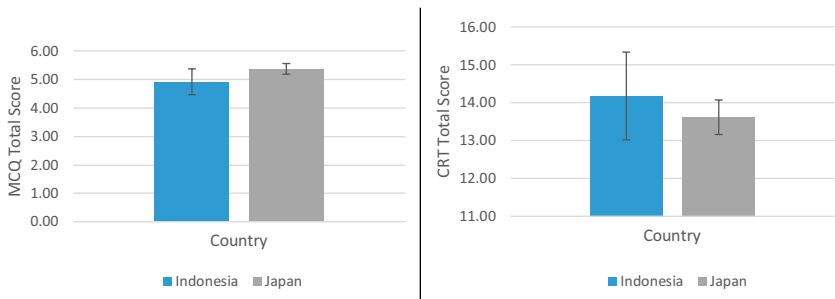
closer to a normal distribution. Moreover, Figure 8 shows the distribution of CRT total scores and its associated normalized curve, in which the deviation of CRT scores to its normalized curve is 8%.

The correlation coefficient between the ITB MCQ and CRT total scores was small at 0.0009. In other words, we may consider that the MCQ test assess completely different aspect of students’ academic competencies in comparison to the CRT test items.

IV. Implications drawn from international benchmarking

This international collaboration in learning outcomes assessment is meaningful, most directly because it allowed for verification of the international validity of test items and the underlying learning outcomes. The implementation at ITB of test items developed by the Japanese team has allowed for engaging in an extremely stimulating and enriching international conversation about what learning outcomes that should be measured and how to do so.

International collaboration is meaningful also because it allowed for recognition of the learning outcome which may be taken for granted, in relation to how we organize our educational practices. It provided information that allows faculty time to stop and think about the strengths and weaknesses of our educational programs.



Note: Error bars represent the 95% confidence interval.

Figure 9

Total Average MCQ and CRT Scores for Japanese and ITB students

Figure 9 shows the total average MCQ and CRT scores for Japanese and ITB students. While Japanese students score slightly higher on MCQs

(Welch's t test $t(49.74) = -1.84, d = -0.26, p = 0.07$) and ITB students slightly higher on CRTs (Welch's t test $t(47.83) = 0.89, d = 0.13, p = 0.38$) the differences are small, cautioning us from over-interpreting these results. A closer examination on the other hand, of how the students in the two countries responded to each test item stirs attention to similarities and differences in educational experiences, curriculum content coverage, sequence, and emphasis.

There were also striking differences in the questionnaire results. When focusing on activities that the students reported to have committed themselves to before joining the laboratory to undertake research projects, significantly more ITB students indicated "foreign language," "general education subjects" and "co-curricular engineering projects," while significantly more Japanese students indicated "paid part time job." After joining the laboratory, significantly more Japanese students responded that they committed themselves to "writing their graduation thesis." With regard to learning opportunities, significantly more ITB students responded that they had opportunities "to engage in solving problems as a team," "to engage in solving real life engineering problems," and "to engage in solving problems that require knowledge beyond engineering (society, economy, politics, etc.)" (Items with mean-difference effect size $d > 0.5$).

Although these are preliminary analysis of the test results, they illuminate the possibilities of a faculty developed learning outcomes assessment in facilitating learning outcomes-based education. Because faculty members have deep understandings of what learning outcomes the assessment tools intend to measure, they are alerted by the reality of that they reveal, and search for clues on how to make improvements. By combining carefully designed MCQs, CRTs, and questionnaires, assessment can become a powerful tool for faculty-led educational improvement.

V. Discussion – limitations and future directions

The 2016 large-scale implementation of the test items confirmed that the NIER test items designed to measure mechanical engineering learning outcomes were valid in Japan and Indonesia, and that the MCQs and the CRT complemented each other by assessing different aspects of engineering learning outcomes. The implementation also revealed critical problems that will need to be addressed in order to achieve robust results that can be utilized for the improvement of learning outcomes-based education.

First, because measuring cognitive processes, or assessing how well a student can “think” is difficult, the scope of learning outcomes that the “machine tools” item was able to address with multiple tasks, as show in Table 3 was limited to Basic Engineering Science, Engineering Analysis, and Engineering Practice. The item was unable to measure Engineering Generic Skills, and was able to measure Engineering Design with only one task. This limitation did not allow for identifying or benchmarking achievement levels of learning outcomes in the five areas, which is the kind of information necessary to highlight the strengths and weaknesses of a program based on the learning outcomes-based approach. Ideally, feedback reports should include information on how well students in a program on average scored in the five learning outcomes areas, benchmarked against their local and international peers, as shown in Figure 10. Hence, one future direction of test item development would be to develop sophisticated and effective approaches to measuring all areas of learning outcomes, and to more intentionally and explicitly map the tasks so that a fuller coverage of the learning outcomes can be achieved.

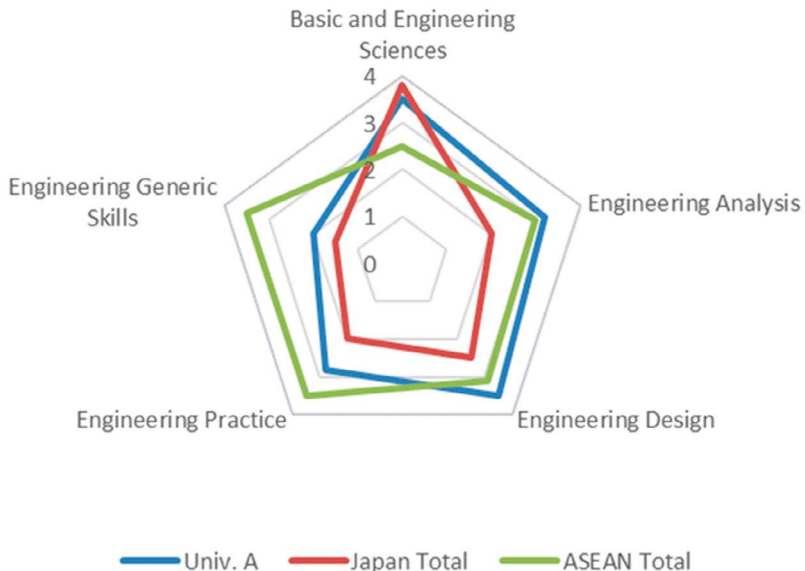


Figure 10
 Hypothetical Comparison of the Average Achievement Level
 of University A, Japan Total, and ASEAN Total

Sophistication of measurement approaches include disaggregating the multiple learning outcomes measured in each task. The correlation coefficients between the five tasks in the “machine tools” item indicate that there are relatively high and statistically significant correlation between all pairs of tasks, implying that the tasks were measuring the same learning outcome or outcomes.

This result led to the realization of an issue in the design of our scoring guide, in which one score was assigned for each task, despite the fact that multiple learning outcomes were being assessed with each task. For example, Task 6 (CRT6) was designed to measure Engineering Analysis, Engineering Design, and Engineering Practice, while the scoring guide was written to assess the level of “integrated” performance supported by the three learning outcomes (one score). A better assessment would have been to write three separate scoring guides to measure the achievement of the three learning outcomes (three scores). Hence, another future direction of test item development would be to develop tasks that more directly target the learning outcomes to be measured, and to develop separate scoring guides for each learning outcome to be measured in each task. This requires more systematic categorization and clearer specification of the learning outcomes, modeled on theories of learning objectives such as those proposed Anderson and Krathwhol.²⁷

Secondly, calibration among scorers (raters) proved to be more difficult than anticipated. Although scorers were asked to calibrate their understanding of the scoring guides in advance, the calibration procedure was not sufficiently defined, which may have led to several incidents of low inter-rater reliability scores. The inter-rater reliability score varied greatly by institution and task, ranging from .29 to .98 (average .79).

Note that the average score of the three scorers were taken to represent the score of the student, in the attempt to cancel out any individual level and institutional level bias. We believe that his approach is basically effective. However, measures must be taken to prevent large discrepancy in the scores, as low inter-rate reliability will challenge to the reliability of the overall result. Hence, one future direction of test implementation is to define in detail the calibration procedure, so that scorers will indeed be able to score the test based on common understandings of the scoring guides.

Thirdly, because knowledge and understanding of Basic and Engineering Sciences encompass an extremely vast area, we must reserve from conceptualizing as if Basic and Engineering Sciences is a single robust construct. The correlation coefficients analysis between the ten MCQs

²⁷ Lorin W. Anderson, and David Krathwhol, *A Taxonomy for Learning, Teaching, and Assessing* (New York, NY: Addison Wesley Longman, Inc., 2001).

indicate that the correlation between pairs of questions are weak overall, and some are negative, indicating that the questions may be measuring different constructs. This observation is also supported by results from a categorical factor analysis. Further investigation is necessary to understand the behavior of MCQs, and in the meanwhile, caution is required to aggregate or construct composite scores. Hence, another future direction of test implementation is to, in order to better understand the nature and structure of the Basic and Engineering Sciences learning outcomes, include multiple items from the same content areas (fundamental mathematics, fundamental physics, materials, motions, energy, informatics, and mechanical process), which will inevitably increase the number of MCQs to be administered.

While the 2016 large-scale implementation of the NIER Test Item Bank tests posed many challenges, it has unmistakably led the engineering team to reach a deeper and more substantive understanding of the learning outcomes that we are trying to help our students achieve, as well as a more critical understanding of the complexities of assessment. Preliminary analysis of the test results has invited faculty to engage in discussion about the students' curricular and extra-curricular educational experiences that may affect the achievement of learning outcomes. The experience of Japanese and Indonesian engineering faculty members engaged in the NIER Test Item Bank project is indeed proof of how learning outcomes assessment can effectively facilitate learning outcomes-based education.

In the 2017 large-scale implementation, scheduled for this winter will involve 15 MCQs and one CRT, the "new machine tools" item, revised based on lessons learnt from the 2016 implementation. We anticipate to further deepen our understanding of engineering learning outcomes and to upgrade our expertise in test item development.

As the test item bank continues to grow, and more quality MCQs and CRTs that allow meaningful analysis and feedback are made available to a wider community of engineering faculty not only in Japan and Indonesia but throughout East Asia and beyond, we aspire to make the "shift of paradigm" amongst faculty and students, the critical actors of a learning outcomes-based education.

Bibliography

Arum, Richard, Josipa Roksa, and Amanda Cook. *Improving Quality in American Higher Education: Learning Outcomes and Assessments for the 21st Century*. San Francisco: Jossey-Bass, 2016.

- Anderson, Lorin W., and David Krathwhol. *A Taxonomy for Learning, Teaching, and Assessing*. New York: Addison Wesley Longman, Inc., 2001.
- Association of American Colleges and Universities. "LEAP Essential Learning Outcomes." <https://www.aacu.org/leap/essential-learning-outcomes>.
- Birtwistle, Tim, Courtney Brown, and Robert Wagenaar. "A long way to go...A study on the implementation of the learning-outcomes based approach in the EU and USA." *Tuning Journal for Higher Education* 3, no. 2 (2016): 429-463. doi: 10.18543/the-3(2)-2016pp429-463.
- European Network for the Accreditation of Engineering Education. "EUR-ACE Framework Standards and Guidelines (EAFSG)." <http://www.enaee.eu/wp-assets-enaee/uploads/2012/02/EAFSG-Doc-Full-status-8-Sept-15-on-web-fm-Denis.pdf>.
- Group of National Experts on the AHELO Feasibility Study. "Engineering Assessment Framework." (AHELO Feasibility Study 8th meeting of the AHELO GNE Paris, 18-19 November 2011). [http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne\(2011\)19/ANN5/FINAL&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/imhe/ahelo/gne(2011)19/ANN5/FINAL&doclanguage=en)
- Fukahori, Satoko, ed. "An Analytical Report of the Results of AHELO Feasibility Study" (Tokyo: National Institute for Educational Policy Research), original title in Japanese: 「AHELO調査結果の分析に関する研究会（研究成果報告書）」 (pronounced as follows: "AHELO Chosa-kekka-no-bunseki-ni-kansuru-kenkyukai (Kenkyu-seika-hokokusho)"). <http://www.nier.go.jp/koutou/ahelo/2014/report.pdf>.
- Fukahori, Satoko. "Competence-based Quality Assurance of University Education - Lessons Learnt from the OECD-AHELO Feasibility Study." *Educational Studies in Japan: International Yearbook* 8 (2014):9-61.
- Hirota, Teruyuki. "Why Front Line Academics Reject Reform – Towards Utilizing Subject Specific Benchmark Statements." *University Evaluation Studies* 15 (2016): 37-46.
- International Engineering Alliance. "Graduate Attributes and Professional Competencies." <http://www.ieagreements.org/assets/Uploads/Documents/Policy/Graduate-Attributes-and-Professional-Competencies.pdf>.
- Japan Accreditation Board of Engineering Education. <http://www.jabee.org/english/>.
- Lumina Foundation. "Degree Qualifications Profile." <https://www.luminafoundation.org/files/resources/dqp.pdf>.
- Ministry of Education, Culture, Sports, Science and Technology (MEXT) Central Council for Education. "Towards the Construction of Bachelor Degree Programs" (Tokyo: MEXT, 2008); original title in Japanese: 「学士課程教育の構築に向けて（答申）」 (pronounced as follows: "Gakushikatei-kyoiku-nokochiku-ni-mukete (Toshin)"). http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm.
- National Institute for Academic Degrees and University Evaluation (NIAD-UE). "A Research on the Quality Assurance of Disciplinary Education in Japanese Universities" (Tokyo: NIAD-UE, 2016); original title in Japanese 「我が国における大学教育の分野別質保証の在り方に関する調査研究報告書」

- (pronounced as: “Wagakuni-niokeru-daigaku-kyoiku-no-bunyabetsu-sanshokijun-no-arikata-ni-kansuru-chosa-kenkyu”). http://www.niad.ac.jp/n_shuppan/project/_icsFiles/afiedfile/2016/07/14/no09_nr16-0714.pdf.
- National Institute for Educational Policy Research (NIER). “Tuning Test Item Bank Project.” <http://www.nier.go.jp/tuning/centre/kikai.html>.
- . “The National Institute for Educational Policy Research International Symposium for Educational Reform 2013: The TUNING-AHELO Global Quality Assurance through Sharing Competence Frameworks and Degree Level Specifications” (Tokyo: NIER, 2013). http://www.nier.go.jp/06_jigyousymposium/i_sympo25/.
- National Research Council. *Engineering Tasks for the New Century: Japanese and U.S. Perspectives* (Washington, DC: The National Academies Press, 1999), <https://doi.org/10.17226/9624>.
- Organization of Economic Cooperation and Development (OECD). “A Tuning-AHELO Conceptual Framework of Expected Desired/Learning Outcomes in Engineering.” *OECD Education Working Papers*, 60 (2011), OECD Publishing, Paris. <http://dx.doi.org/10.1787/5kghtchn8mbn-en>.
- . “OECD-AHELO Feasibility Study Report Volume 2 – Data Analysis and National Experiences.” (Paris: OECD, 2013). <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf>.
- Quality Assurance Agency (QAA). “The UK Quality Code for Higher Education Subject Benchmark Statements.” <http://www.qaa.ac.uk/assuring-standards-and-quality/the-quality-code/subject-benchmark-statements>.
- Science Council of Japan. “Benchmark Statements for Curriculum Design for Subject-Specific Quality Assurance in University Education.” (Tokyo: Science Council of Japan, 2012); original title in Japanese 「大学教育の分野別質保証のための教育課程編成上の参照基準について」 (pronounced as “Daigaku-kyoiku-no-bunyabetsu-shituhosho-no-tame-no-kyoiku-katei-henseijono-sanshokijun-nitsuite”). <http://www.scj.go.jp/ja/member/iinkai/daigakuhosyo/daigakuhosyo.html>.
- Tremblay, Karine, Dianne Lalancette, and Deborah Roseveare. “OECD-AHELO Feasibility Study Report Volume 1 - Design and Implementation.” (Paris: OECD, 2012). <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>.
- Tuning Educational Structures in Europe. “Subject Areas.” <http://www.unideusto.org/tuningeu/subject-areas.html>
- Wiggins, Grant, and Jay McTighe. *Understanding by Design*, expanded 2th ed. (Alexander, Virginia: Association for Supervision & Curriculum, 2005).

About the authors

JEFFREY S. CROSS (cross.j.aa@m.titech.ac.jp) received a Ph.D. in chemical engineering from Iowa State University in 1992. He has worked in Japan at Fujitsu Lab/Fujitsu Ltd., National Institute for Research in Inorganics Materials (now

NIMS), and at the Tokyo Institute of Technology (Tokyo Tech) for over 24 years. Jeffrey is a Professor in the Tokyo Tech's School of Environment and Society, Dept. of Transdisciplinary Science and Engineering and a graduate coordinate for the Energy Science and Engineering Major. He teaches online courses on academic writing and on education technology, researches biofuels, educational technology and analyzes Japan's energy policy. He is the founder and general manager the Tokyo Tech Online Education Development Office, which develops MOOCs on edX. He was an observer in the Japan Team that participated in the OECD-AHELO project on civil engineering competence assessment. He currently serves as the secretariat for the UK-Japan Engineering Education League.

ESTIYANTI EKAWATI (esti@spm.itb.ac.id) received a Ph.D. in engineering from Murdoch University, Australia in 2004. She is an Associate Professor at the Institut Teknologi Bandung, Indonesia, Faculty of Industrial Technology. She teaches courses on engineering designs, probability and statistics, dynamic modeling, controls and optimizations. She is a member of IEEE Control Systems Society and IEEE Education Society. She currently serves as the secretary of Institut Teknologi Bandung Quality Assurance Unit.

SATOKO FUKAHORI (fukahori@nier.go.jp) is the Research Director of the Department for Higher Education Research, National Institute for Educational Policy Research (NIER), the research arm of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. Her research topics include quality assurance of higher education, program design, assessment of higher education learning outcomes, etc. She was the principal national coordinator for the 2008-2012 design and implementation of the Organization for Economic Cooperation and Development (OECD) Assessment of Higher Education Learning Outcomes (AHELO), to which Japan participated in the engineering strand. From 2014, Dr. Fukahori has been leading the Japanese follow-up program "The Global Quality Assurance of Higher Education through the Development of a Tuning Test Item Bank," working with a team of engineers seeking to develop a shared understanding of expected learning outcomes in the field of mechanical engineering. She is also the Director of the Tuning Japan National Center, the Japanese contact point for Tuning activities. Dr. Fukahori completed her B.A. (1991) and M.A. (1993) degrees in Comparative Education at Kyoto University, and received her Ph.D. (2000) at Columbia University Graduate School of Arts and Sciences in Sociology of Education.

SHINNOSUKE OBI (obsn@mech.keio.ac.jp) received Dr.-Ing. from the University of Erlangen-Nuremberg, Germany, in 1991. He is a professor in the Department of Mechanical Engineering, Keio University, Japan, and Fellow of the Japan Society of Mechanical Engineers, member of Japan Society of Fluid Mechanics, Heat Transfer Society of Japan, Gas Turbine Society of Japan, and the Japan Society of Civil Engineers. From October 2017 he is appointed to the Dean of the International Center of Keio University where he supervises the operation of various international programs offered to the students of Keio University and its partner institutions in the world. He serves also as a member of various

committees at the Ministry of Education, Culture, Sports, Science and Technology of Japan, related to scholarship programs for international students, and human resource development project at JICA, Japan International Cooperation Agency, through AUN/SEED-Net Program.

YUGO SAITO (ugo.saito@gmail.com) is a program-specific assistant professor at Kyoto University's Center for the Promotion of Excellence in Higher Education. He uses statistical methods to explore assessing learning outcomes in higher education.

NATHANAEL P. TANDIAN (n4th4n.t4nd14n@gmail.com, n4th4n@termo.itb.ac.id) completed his studies in Institut Teknologi Bandung - ITB, Indonesia (bachelor in 1983), University of Minnesota, Twin Cities, USA (MSME in 1991 and Ph.D. in 1994); all are in Mechanical Engineering. He has been being a faculty member in the Mechanical Engineering Department, ITB since 1984 and teaches thermodynamics, mechatronics, and applied mathematics. He is also a member of the Energy Conversion Research Group and the Center for New and Renewable Energy in ITB, with research interest in the area of refrigeration, heat transfer, and computational fluid dynamics. He actively serves as a member of the ad hoc committee for curriculum, and a member of the quality assurance unit in the Mechanical, Aeronautic and Aerospace Engineering Faculty, ITB. Currently he is also a fellow of the Institution of Engineers Indonesia.

FARID TRIAWAN (triawan.f.aa@m.titech.ac.jp, farid.triawan@gmail.com) earned his bachelor degree in mechanical engineering from Bandung Institute of Technology, Indonesia in 2005, and then received his master (in 2009) and doctoral degrees (in 2012) in mechanical engineering both from Tokyo Institute of Technology (Tokyo Tech), Japan. He worked at Torishima Pump Mfg. Co., Ltd. in Japan for 3.5 years as a mechanical engineer and was assigned as the leader of materials research group of R&D Department. Currently, Farid is a specially appointed Associate Professor (Lecturer) in the Tokyo Tech School of Environment and Society, Dept. of Transdisciplinary Science and Engineering. He teaches several courses related to materials mechanics and structural engineering, and has been actively managing the newly opened International Undergraduate Program (so-called GSEP) at Tokyo Tech. He is also a member of Japanese Society of Mechanical Engineers (JSME).

Acknowledgements

The authors are deeply indebted to their institutions and the participating faculty members below for their leadership, and the development, implementation and scoring of the test items. Finally, the authors would like to express our sincere gratitude to the quality assurance unit of ITB that has supported the Tuning project within ITB.

Member Name		Affiliation	Leadership Position
Kikuo	Kishimoto	Tokyo Institute of Technology	Representative
Nobuyuki	Iwatsuki	Tokyo Institute of Technology	Test Item Development Leader
Takashi	Nakamura	Hokkaido University	East Japan Hub Leader
Hiroo	Yugami	Tohoku University	East Japan Hub Leader
Hirotsugu	Inoue	Tokyo Institute of Technology	Kanto Hub Leader
Osamu	Nakabeppu	Meiji University	Kanto Hub Leader
Seiichi	Hata	Nagoya University	West Japan Hub Leader
Satoshi	Watanabe	Kyushu University	West Japan Hub Leader
Hideki	Mima	The University of Tokyo	Copyright Clearance Committee
Ichsan S.	Putra	Institut Teknologi Bandung	
Sigit P.	Santosa	Institut Teknologi Bandung	
Indria	Herman	Institut Teknologi Bandung	
Masao	Watanabe	Hokkaido University	
Nozomu	Hashimoto	Hokkaido University	
Itsuro	Kajiwara	Hokkaido University	
Yuko	Suto	Tohoku University	
Takeshi	Okuyama	Tohoku University	
Yasuhisa	Hirata	Tohoku University	
Koichi	Hashimoto	Tohoku University	
Masayoshi	Mizutani	Tohoku University	
Kazuhiko	Kudo	Tokyo Denki University	
Ichiro	Tanaka	Tokyo Denki University	
Hiroshi	Ooue	Tokyo City University	

Member Name		Affiliation	Leadership Position
Naoki	Ono	Shibaura Institute of Technology	
Tetsuyuki	Hongo	National Institute of Technology	
Hiroshi	Yamaura	Tokyo Institute of Technology	
Makoto	Ichitubo	Nagaoka University of Technology	
Hirohumi	Daiguji	The University of Tokyo	
Naohiko	Sugita	The University of Tokyo	
Norikazu	Suzuki	Nagoya University	
Akihiro	Saso	Nagoya University	
Kenji	Tahara	Kyushu University	
Kosaku	Kurata	Kyushu University	
Shinnichi	Hino	Kyushu University	
Hidemi	Toh	Kanazawa Institute of Technology	
Yoshitaka	Morimoto	Kanazawa Institute of Technology	
Keiichi	Sato	Kanazawa Institute of Technology	
Yasuhiro	Takaya	Osaka University	
Toshitsugu	Tanaka	Osaka University	
Akiko	Sakaguti	National Institute for Educational Policy Research	Secretariat

Development of a Mechanical Engineering Test Item Bank to promote learning outcomes-based education in Japanese and Indonesian higher education institutions

Jeffrey S. Cross, Estiyanti Ekawati, Satoko Fukahori,
Shinnosuke Obi, Yugo Saito, Nathanael P. Tandian, and Farid Triawan

Copyright

Copyright for this article is retained by the Publisher. It is an Open Access material that is free for download, distribution, and or reuse in any medium only for non-commercial purposes; provided any applicable legislation is respected, the original work is properly cited, and any changes to the original are clearly indicated.