The 14th International CDIO Conference in Kanazawa
Sunday 1 July 2018 9:20 – 10:30
Room 23-330/333

MEASURING HOW WELL STUDENTS CAN THINK LIKE AN ENGINEER - A COLLABORATIVE EFFORT AMONG JAPANESE AND INDONESIAN ACADEMICS –

Satoko Fukahori

- National Institute for Educational Policy Research,
- Kyushu University

NIER

Tuning Japan National Centre

ASSESSMENT OF LEARNING OUTCOMES

- What faculty do everyday in their courses and programs.
  - Assignments, tests.
  - Grades, portfolios.
  - Diagnostic, formative, summative.
- Requirements from a quality assurance perspective.
  - Learning outcomes must be relevant not only from the viewpoint of the faculty, but also from the viewpoints of institutional mission and students needs within their unique societal contexts.
  - Learning outcomes should be benchmarked against standards/reference points agreed upon by academic/professional communities.
    - International Engineering Alliance's Graduate Attributes and Professional Competences
    - European Network for the Accreditation of Engineering Education EUR-ACE standards.
  - Development of concrete-level shared understandings of abstract-level learning outcomes, in terms of their scopes and levels.

The focus of this presentation

THE JAPANESE EFFORT

- 2008-2013: Participation in the OECD-AHELO (Assessment of Higher Education Learning Outcomes) Feasibility Study, Engineering Strand.
  - http://www.oecd.org/education/skills-beyond-school/testingstudentanduniversityperformancegloballyoecdsahelo.htm
- 2014-present: Development of a Test Item Bank in Engineering, as part of the National Institute for Educational Policy Research's Tuning Project.
  - http://www.nier.go.jp/tuning/centre.html
- Similar efforts around the globe.
  - Tuning CALOHEE (Measuring and Comparing Achievements of Learning Outcomes in Higher Education in Europe) (EU Erasmus+). https://www.calohee.eu
  - National Institute for Learning Outcomes Assessment's Assignment Library (Lumina Foundation). https://www.assignmentlibrary.org

THE OECD-AHELO FEASIBILITY STUDY

2008–2013

ahelo
Assessment of Higher Education Learning Outcomes

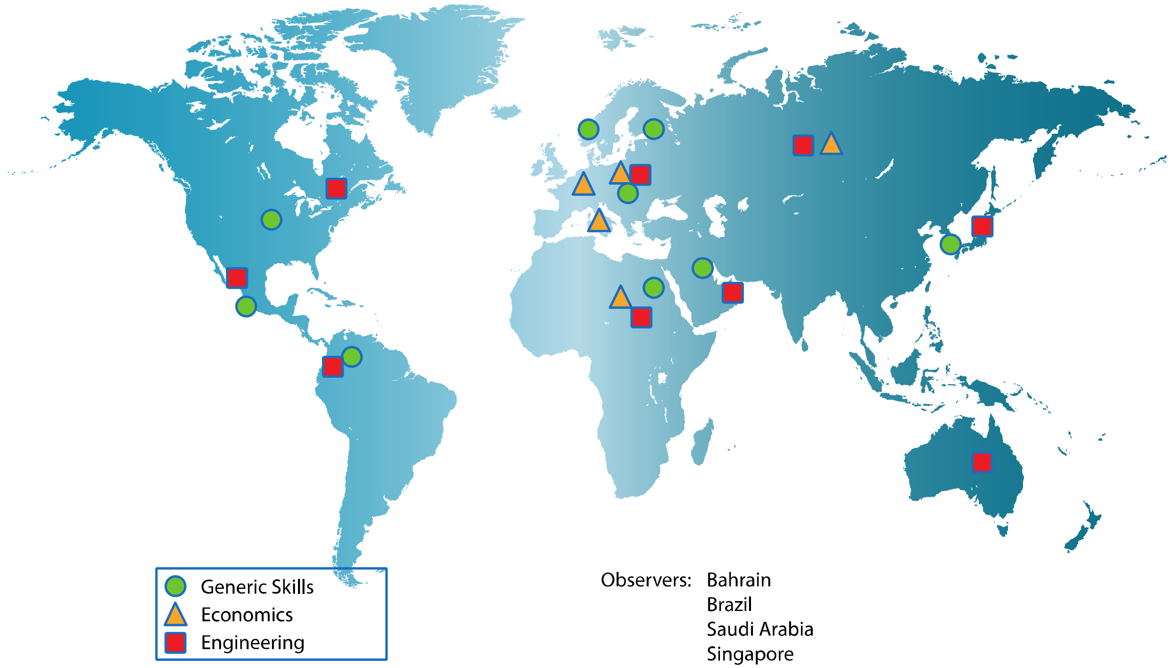
PURPOSE: TO TEST THE SCIENTIFIC AND PRACTICAL FEASIBILITY OF ASSESSING WHAT HIGHER EDUCATION STUDENTS KNOW AND CAN DO AT GRADUATION, ACROSS DIVERSE NATIONAL, CULTURAL, LINGUISTIC AND INSTITUTIONAL CONTEXTS.

Countries and institutions

Seventeen countries, spanning a wide range of languages and cultures, are participating in one or more of the assessments: Abu Dhabi, Australia, Belgium, Colombia, Egypt, Finland, Italy, Japan, Korea, Kuwait, Mexico, the Netherlands, Norway, the Russian Federation, the Slovak Republic, Sweden, and the United States.

There are about 10 higher education institutions participating in each country.

Figure 2. AHELO Country-specific strand of work



http://www.oecd.org/edu/skills-beyond-school/AHELO%20Brochure.pdf

THE ENGINEERING STRAND

Phase 1 2008 to 2011	Small-scale Verification of Instruments (in Japan and Australia): • Instrument developed by Australia (ACER) and Japan (NIER) in consultation with an international panel of engineering experts. • Pencil and Paper test(60 minutes)(1CRT, 20MCI) • Survey about the instrument(60 minutes) • Faculty and Institutional Survey Japan: 10 Universities, 75 students
Phase 2 2011 to 2012	Large-scale Implementation of Instruments (in 9 countries): • Online test(90 minutes)(1CRT, 25MCI) • Contextual Instrument(10 minutes) • Faculty and Institutional Survey online Japan: 12 Universities, 504 students
Final conference 11-12 March 2013 (Paris)	Lessons learnt from the AHELO Feasibility Study and next steps: OECD (2012). Assessment of Higher Education Learning Outcomes Feasibility Study Report Volume 1 - Design and implementation http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf Volume 2 - Data Analysis and National Experiences http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume2.pdf

ENGINEERING LEARNING OUTCOMES

OECD (2011), "A TUNING-AHELO CONCEPTUAL FRAMEWORK OF EXPECTED DESIRED/LEARNING OUTCOMES IN ENGINEERING", OECD EDUCATION WORKING PAPERS, NO. 60, OECD PUBLISHING, PARIS. HTTP://DX.DOI.ORG/10.1787/5KGHTCHN8MBN-EN

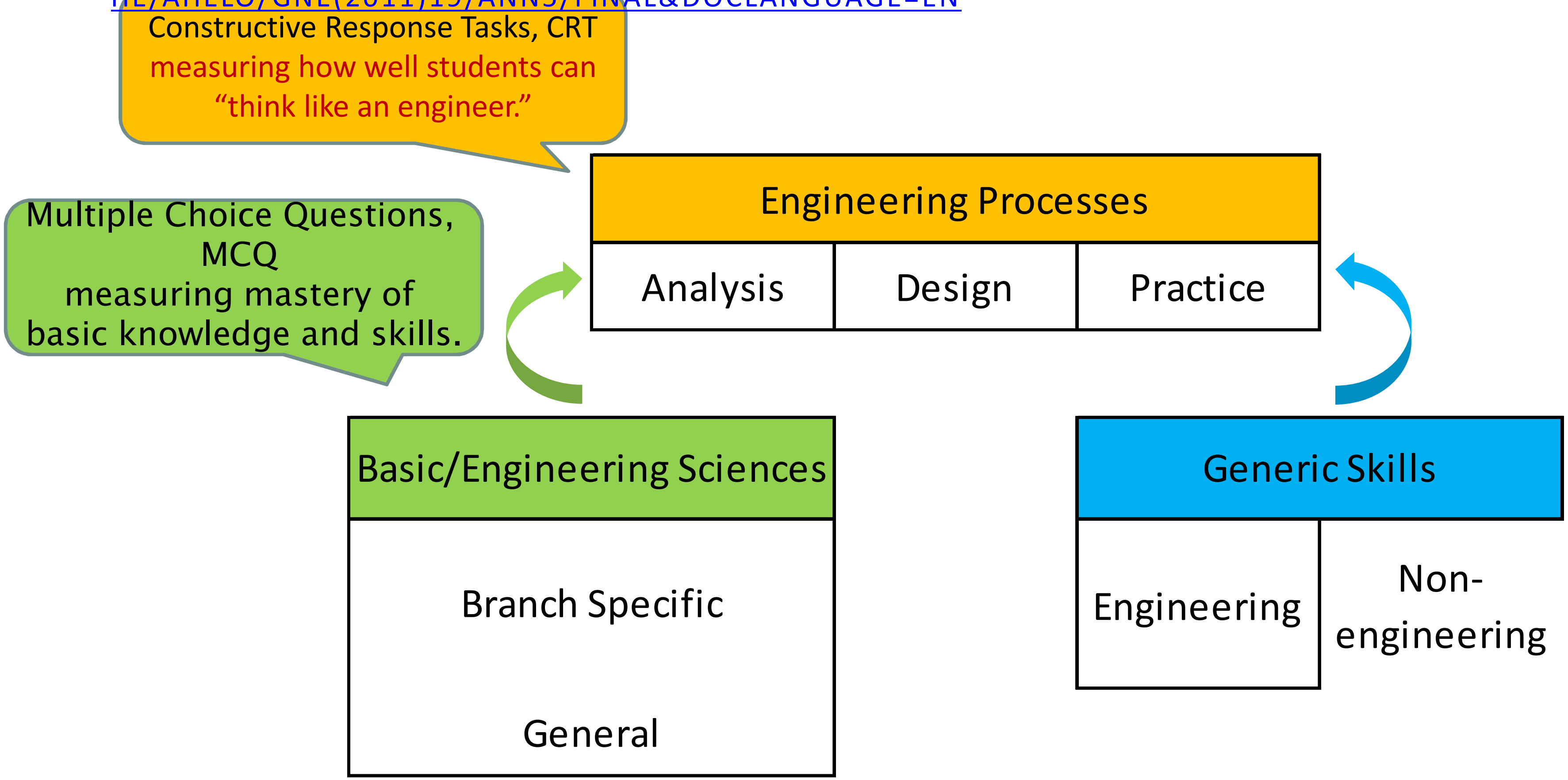
REF: ENAEE: EUR-ACE & IEA - GRADUATE ATTRIBUTES（WASHINGTON ACCORD)

SUFFICIENTLY ABSTRACT FOR DIVERSE INSTITUTIONS AND AUTONOMOUS FACULTY TO ACCEPT AND SHARE.

- Engineering Generic Skills:
  - Effective communication and awareness of the wider civil engineering context.
- Basic and Engineering Sciences:
  - Knowledge and understanding of the scientific and mathematical principles underlying civil engineering – general sciences; materials and construction; structural engineering; geotechnical engineering; hydraulic engineering; and urban and rural planning.
- Engineering Analysis:
  - Using analytical methods to identify, formulate and solve engineering problems.
- Engineering Design:
  - Understanding and application of design methodologies to meet specified requirements.
- Engineering Practice:
  - Practical skills and knowledge required for solving problems, conducting investigations, and designing engineering devices and processes. Addresses non-technical elements of civil engineering practice such as professional ethics, responsibilities and the impact of engineering solutions in a global, economic, societal and environmental context.

ENGINEERING ASSESSMENT FRAMEWORK

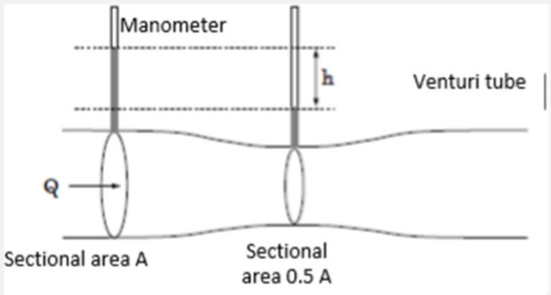
HTTP://SEARCH.OECD.ORG/OFFICIALDOCUMENTS/DISPLAYDOCUMENTPDF/?COTE=EDU/IMHE/AHELO/GNE(2011)19/ANN5/FINAL&DOCLANGUAGE=EN



ILLUSTRATIVE ENGINEERING TEST ITEMS: A MULTIPLE CHOICE ITEM

Two manometers are attached to a Venturi tube as shown below.
Select the correct formula that expresses flow rate *Q* in the pipe channel.

Assume that g represents the gravitational acceleration and that energy loss is negligible.



A. $Q = \sqrt{\frac{2}{3}}A\sqrt{gh}$

B. $Q = \frac{2}{3}A\sqrt{gh}$

C. $Q = \sqrt{2}A\sqrt{gh}$

D. $Q = \frac{\sqrt{2}}{3}A\sqrt{gh}$

Description: Identifies the correct expression for flow rate in a Venturi tube.

Competencies: BES3. (Demonstrates: comprehensive knowledge of their branch of engineering including emerging issues.)

Specialized area: Hydraulic Engineering, including water engineering and management; design of components and systems such as water supply systems and sewer networks.

Correct answer: A

OECD (2012). pp.266-267.

ILLUSTRATIVE ENGINEERING TEST ITEMS:
A CONSTRUCTIVE RESPONSE TASK

The Hoover Dam is a 221-metre high concrete arch-gravity dam in the Black Canyon of the Colorado River in the United States of America.

It was built to provide irrigation water, to control floods and to provide water for a hydroelectric power station at the base of the dam.

OECD (2012). pp.252-264.

Q1. Explain why this is a good dam site for hydroelectric power generation. You should discuss at least two aspects.

Scoring Note 1: The question requires students to explain, therefore responses should both list a feature AND provide an indication of why/how that feature makes the site suitable for the dam.

Scoring Note 2: Note the emphasis of hydroelectric power generation in the question.

(a) Dam height / High potential energy

(b) High flow rate of river (amount of water entering the dam)

(c) Lake capacity

(d) Minimal social impact

(e) Characteristics of rock (could also include hardness and suitable foundation)

(f) Narrow gorge

Examples of answers that can be marked as correct are presented.

ILLUSTRATIVE ENGINEERING TEST ITEMS: A CONSTRUCTIVE RESPONSE TASK (CONTINUED)

Q2. Explain the two main design features that contribute to the structural strength and stability of the Hoover dam.

Scoring Note: The question requires students to explain, therefore responses should both list a feature AND provide an indication of why/how that feature makes the site suitable for the dam.

(a) Arch-shape

(b) Material in canyon

(c) Weight of concret

(d) Tapered shape of concrete wall/low centre of gravity

(e) Spillways and/or tunnels

Q4. Imagine that a new dam is being planned today in a different location. Briefly explain two environmental effects of the dam (which could also be upstream or downstream) that an engineer would need to consider in an environmental impact statement.

Scoring Note: The guide assumes that changes to the river flow are understood. Candidates need to provide indications of the consequences of changes to the flow.

(a) Habitats

(b) Soil and/or siltation/erosion

(c) Ground stability around the storage itself

(d) CO2 Emissions/Greenhouse gase

(e) Aesthetics

(f) Effluent impact

(g) Community impact

REACHING CONCRETE-LEVEL SHARED UNDERSTANDINGS OF ABSTRACT-LEVEL LEARNING OUTCOMES THROUGH SCORER TRAINING

International Scorers Training at OECD, Nov 2011, Mar 2012

Domestic Scorers Training and Scoring Session at Tokyo, Jun 2012

LESSONS LEARNED FROM THE OECD-AHELO FS STUDY

1. The significance and relevance of academic and professional accords for mutual recognition of qualifications and registration in engineering were acknowledged.
2. The exercise of developing, implementing, scoring, and modifying items and scoring rubrics by international/national teams of experts proved to be extremely important to reach concrete-level shared understandings of the abstract-level learning outcomes in terms of their scopes and levels.
3. An international assessment of higher education learning outcomes can become a useful tool for educators to globally benchmark and update their teaching practices. It raises student awareness of their learning, too!
4. Designing constructive response tasks to measure how students can "think like an engineer" requires a thoughtful balance between preciseness and open-endedness.
   - Ongoing efforts to construct and reconstruct items scoring rubrics are necessary.
   - Not yet suitable for high stakes testing.

The Japanese AHELO team acknowl
and decided to continue the wo

ahelo
Assessment of Higher Education Learning Outcomes

THE NIER-TUNING TEST ITEM BANK

2014–present

NIER

Tuning Japan National Centre

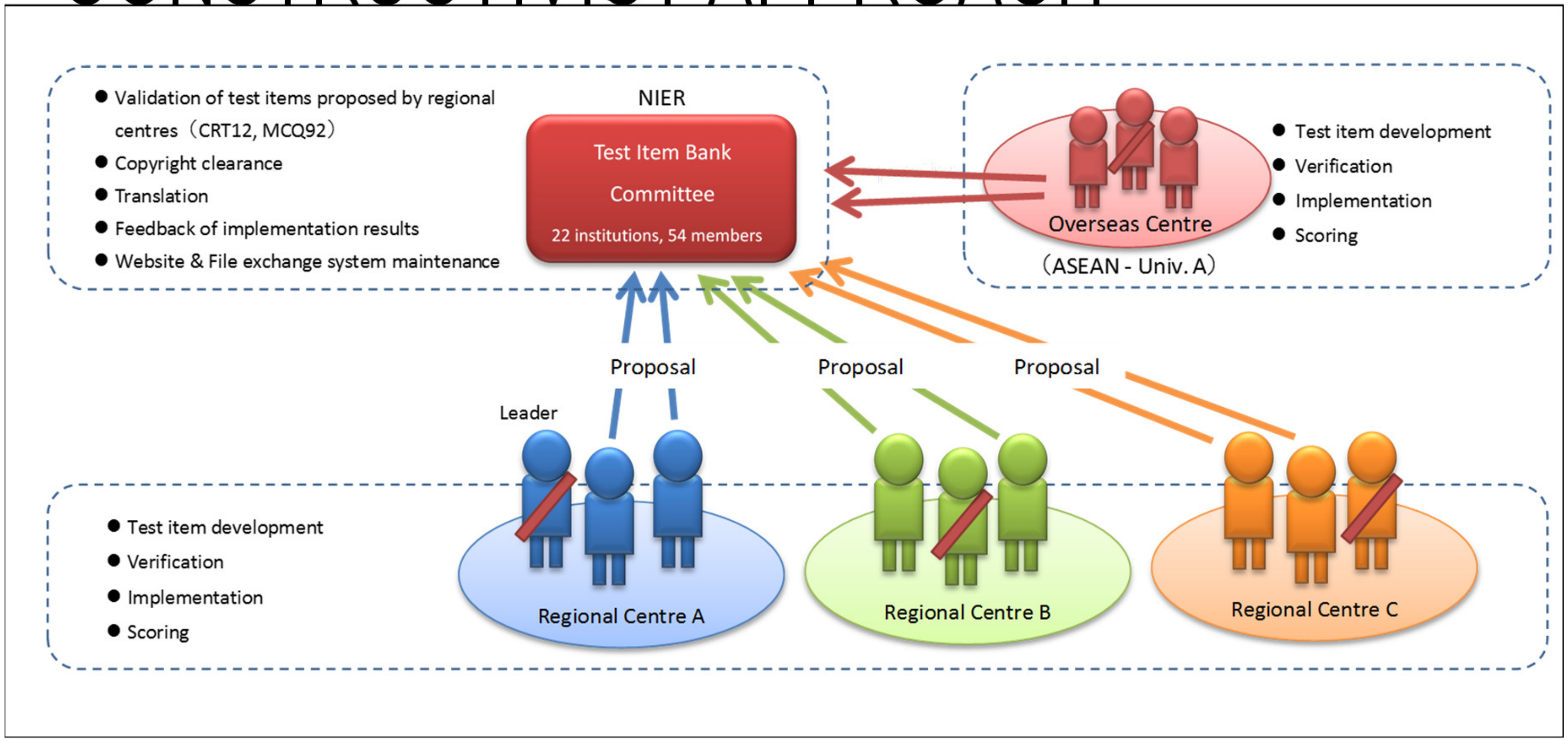
NIER-TUNING TEST ITEM BANK(2014-)

Purpose:

- Developing a shared understanding of expected learning outcomes, through jointly engaging in the process of developing, verifying, implementing, and scoring test items.
  - Professional development at academic societies (Japan Society of Mechanical Engineers, annual meeting, as part of the JABEE accreditation program evaluator training)
  - Faculty improvement at universities.
- Drawing implications for program improvement.
  - Program self review report.
  - Development of new courses & modules.

THE NIER-TUNING TEST ITEM BANK: ACOLLABORATIVE AND CONSTRUCTIVIST APPROACH

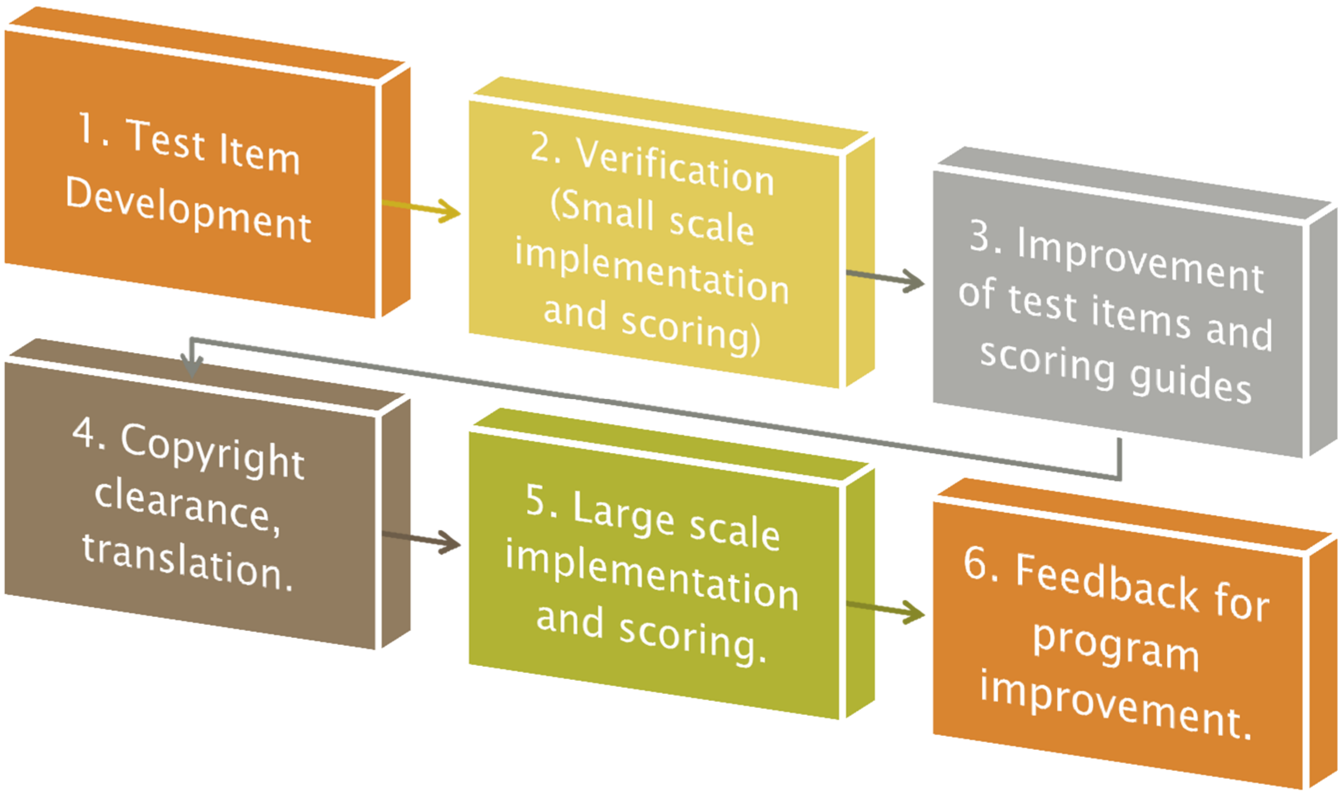
Tuning Test Item Bank Activities (2017)



Host Institutions:
West Japan Hub: Kyushu University & Nagoya University.
Kanto Area Hub: Tokyo Institute of Technology & Meiji University.
East Japan Hub: Tohoku University & Hokkaido University
ASEAN Hub: Institut Teknologi Bandung,

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. The basics of mathematics include differential and integral calculus, linear algebra, and numerical methods.
BES2	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.

TEST ITEM DEVELOPMENT PROCESS



AN EXAMPLE :
WIND ELECTRICAL POWER GENERATION
(HTTP://WWW.NIER.GO.JP/TUNING/CENTRE.HTML)

WIND POWER GENERATION IS THE CONVERSION OF WIND KINETIC ENERGY INTO ELECTRICAL ENERGY OR ELECTRICITY, THROUGH THE USE OF WIND TURBINES....RESPOND TO THE FOLLOWING QUESTIONS WHICH FOCUS ON THE WIND TURBINES USED FOR WIND POWER GENERATION FROM A MECHANICAL ENGINEERING POINT OF VIEW.

Question 1. Examine the locational condition or site of a wind farm for wind power generation.

Figure 2 shows a wind farm for wind power generation. List and explain two reasons
site for winc

Figure 2: An example of a wind farm
Photograph of Otonrui Wind Farm, provided by Horonobe City

Question 2. Examine the "shape of the blades" of wind turbines used for wind power generation.

Compare the shapes of the blades for a traditional windmill and a wind turbine shown in Figures 3a and 3b, respectively. Explain from a mechanical engineering point of view two features of

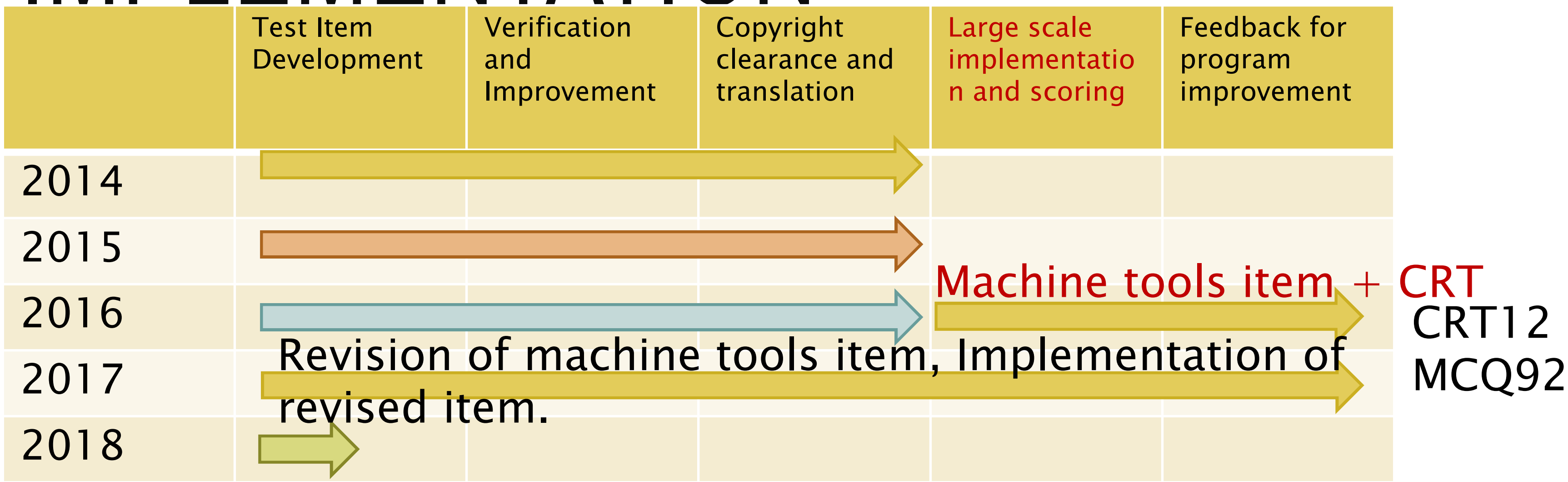
Figure 3a Traditional windmills.
Martijn Roos. www.mroosfotografie.nl
http://free-photos.gatag.net/2014/11/07/040000.html
Figure 3b Wind turbines used for wind power generation.
http://sozai-free.com/sozai/01541.html

SCORING GUIDE FOR Q1

SUFFICIENTLY CONCRETE TO PROVIDE A MEANINGFUL FRAMEWORK FOR INSTITUTIONS/FACULTY TO REFER TO WHEN DESIGNING PROGRAMS AND COURSES.

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.
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.
Viewpoints: Lists two features out of three below or equivalent, and explains the reasons for each of them appropriately. (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. • 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 stationarily tower and against the turbines. (b) Many wind turbines are installed in one location. •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. (c) No building or structure is located around the wind farm. •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.

LESSONS LEARNED FROM THE LARGE SCALE IMPLEMENTATION

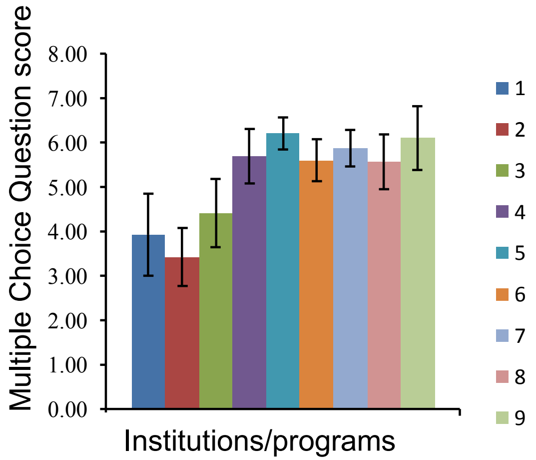


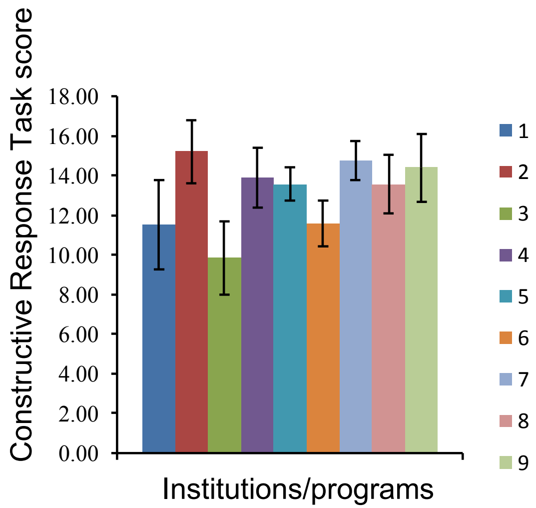
Providing feedback to faculty and students forced us to revisit our test items.
Are our items really assessing the learning outcomes that we meant to assess?
program improvement, the need for sophistication of test items

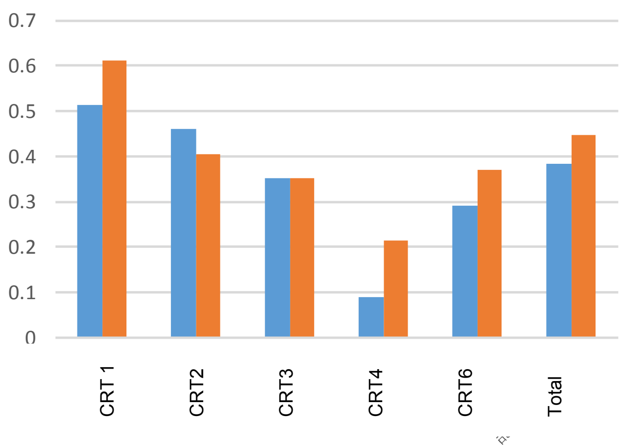
LARGE SCALE IMPLEMENTATION 2016

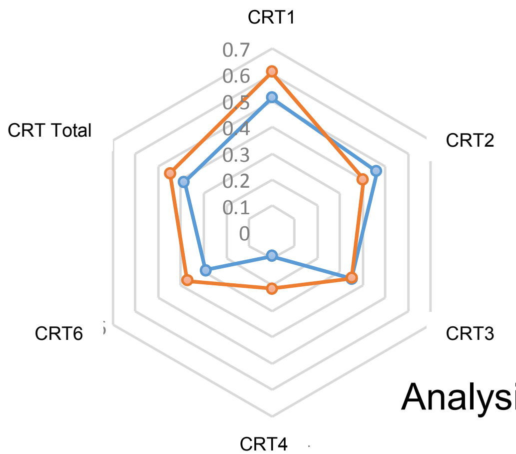
- 10 universities (Japanese and Indonesian) & 385 students
- June-September 2016
- Test items
  - Multiple choice questions (10 items, 30 minutes)
  - Constructive response task ("Machine Tools," 50 minutes)
  - Contextual Survey (10 minutes)
- Feedback reports delivered to the project team, individual participating universities and individual participating students.

RESULTS OF LARGE SCALE IMPLEMENTATION 2016
BENCHMARKING INSTITUTIONAL PERFORMANCE
HIGHLIGHTING THE STRENGTHS AND WEAKNESSES









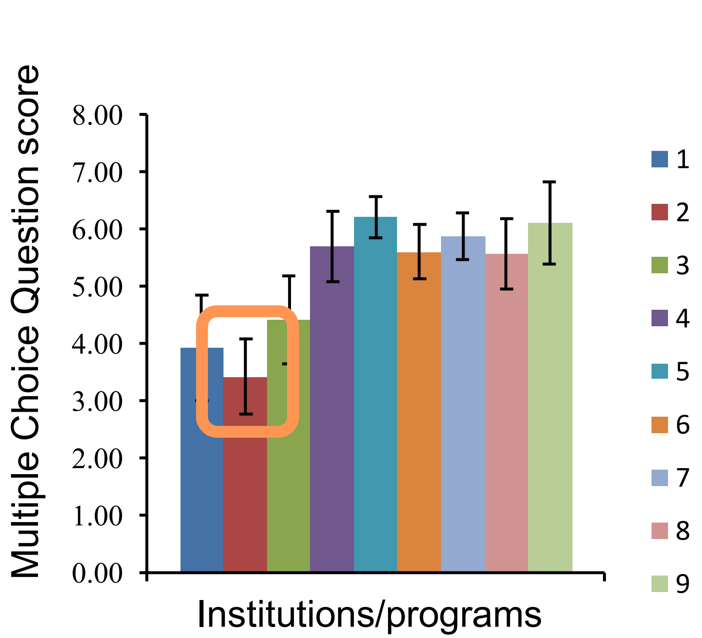
Analysis by Yugo Saito

RESULTS OF LARGE SCALE IMPLEMENTATION 2016

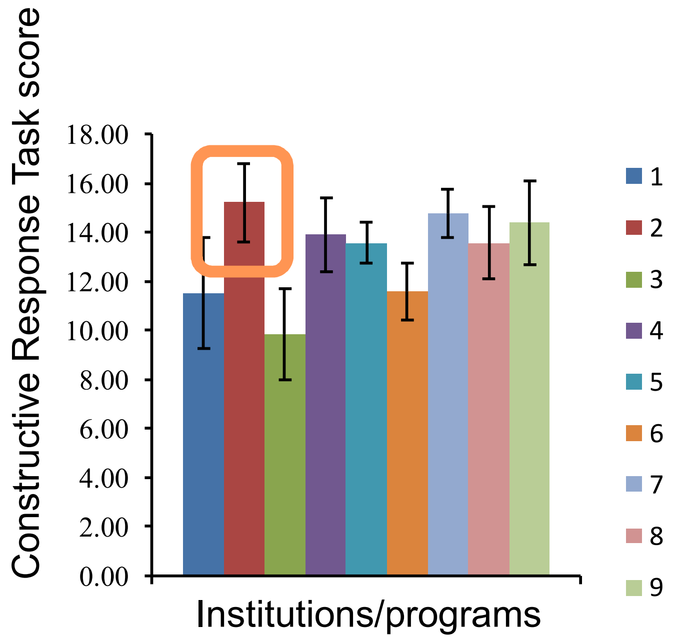
THE CORRELATION BETWEEN MULTIPLE CHOICE QUESTION (MCQ) SCORES AND CONSTRUCTIVE RESPONSE TASK (CRT) SCORES WAS *R* = .17 (*P*<.01)

HIGH MCI SCORES DO NOT NECESSARILY LEAD TO HIGH CRT SCORES, VICE VERSA.

FOCUS ON INSTITUTION 2 WITH LOW MCI BUT HIGH CRT SCORES.



95% confidence interval



95% confidence interval

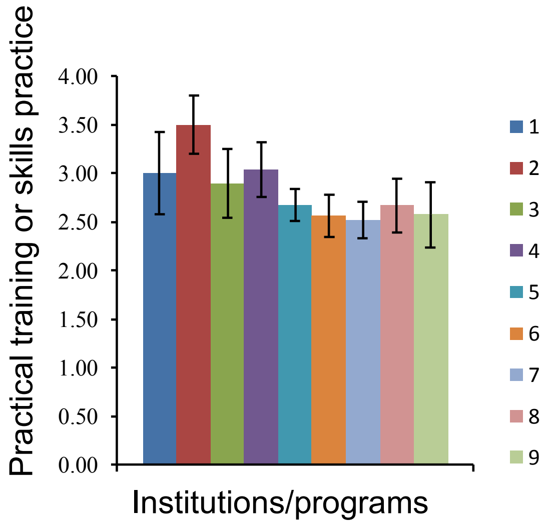
Analysis by Yugo Saito

RESULTS OF LARGE SCALE IMPLEMENTATION 2016

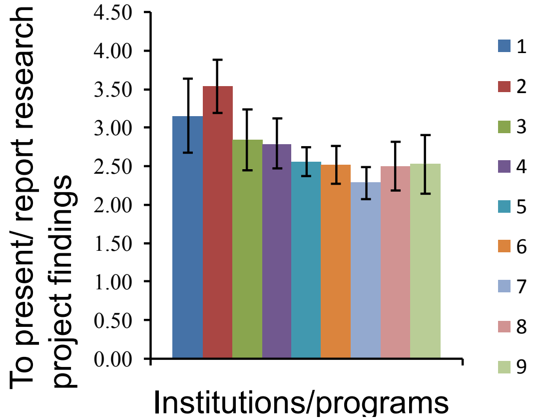
WHAT ARE THE EDUCATIONAL CHARACTERISTICS OF INSTITUTION 2?

STUDENTS IN INSTITUTION 2 RESPONDED MORE AFFIRMATIVELY THAT THEY:

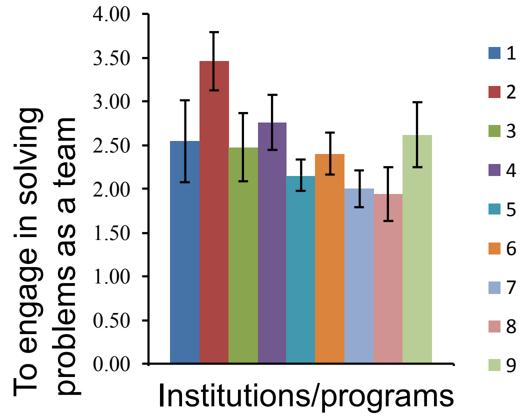
COMMITTED THEMSELVES TO “PRACTICAL TRAINING OR SKILLS PRACTICE”
BEFORE JOINING THE LABORATORY; AND
HAD OPPORTUNITIES “TO PRESENT/ REPORT RESEARCH PROJECT FINDINGS” AND
“TO ENGAGE IN SOLVING PROBLEMS AS A TEAM.”



95% confidence interval



95% confidence interval



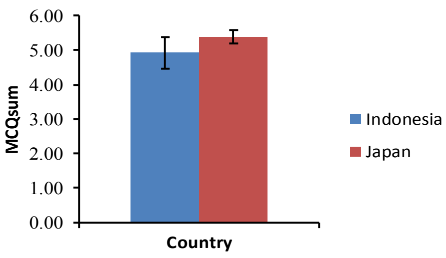
95% confidence interval

Analysis by Yugo Saito

RESULTS OF LARGE SCALE IMPLEMENTATION 2016

COMPARISON OF INDONESIAN AND JAPANESE UNIVERSITIES.

NO SIGNIFICANT DIFFERENCE IN THE TOTAL SCORE FOR MCQ.
SOME DIFFERENCES IN CRT, POSSIBLY REFLECTING CURRICULUM CONTENT COVERAGE, SEQUENCE,
AND EDUCATIONAL EXPERIENCE.



MCQ Total
95% confidence interval

・Before joining the laboratory, more Indonesian students responded that they committed themselves to studies in “foreign language,” “general education subjects,” and co-curricula engineering activities,” whereas more Japanese students responded that they committed themselves to “paid part-time jobs.” After joining the laboratory, more Japanese students responded that they committed themselves “writing graduation thesis.”

・ More Indonesian 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.)”

Constructive Response Task 計量

	Indonesia				Japan			
	Mean	SD	Minimum	Maximum	Mean	SD	Minimum	Maximum
CRT1	9.39	2.07	3.33	11.33	7.34	2.20	0.00	12.00
CRT2	2.11	1.41	0.00	6.00	1.62	1.07	0.00	5.33
CRT3	1.24	0.72	0.00	2.67	1.41	0.85	0.00	4.00
CRT4	0.74	0.66	0.00	2.33	0.86	0.97	0.00	4.00
CRT5	0.22	0.42	0.00	1.00	0.18	0.39	0.00	1.00
CRT6	0.49	0.71	0.00	3.33	2.22	1.79	0.00	6.00
CRT sum score	14.18	3.59	4.00	20.00	13.62	4.33	0.00	23.67

Members from both countries agreed on the importance of educational benchmarking. Indonesia will be contributing test items starting 2017

Analysis by Yugo Saito

PUBLICATION

Jeffrey S. Cross, Estiyanti Ekawati, Satoko Fukahori, Shinnosuke Obi, Yugo Saito, Nathanael P. Tandian, Farid Triawan. (2017). "Development of a Mechanical Engineering Test Item Bank to promote learning outcomes-based education in Japanese and Indonesian higher education institutions."

Tuning Journal for Higher Education, Vol. 5, No 1 (2017)

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

FINDINGS FROM THE TECHNICAL ANALYSIS

1. Measuring a wide range of basic/engineering science learning outcome with just 10 MCQs (in 30 minutes) was unrealistic.

- The results of the 10 items revealed to be largely uncorrelated, meaning that the items were measuring the achievement of different areas of knowledge and skills.
  - Solution: Implement more items (which require much more time) or focus on just one area of knowledge and skills.

2. Variability in inter-rater reliability scores of CRTs.

- Scorers at each university were trained by the project team members.
- For student x at University A: Two scorers from University A and one scorers from University B.
  - Solution: Intensive calibration at all universities. (jointly scoring several student responses).

3. The need for more intentional and closer alignment of tasks to learning outcomes.

- Engineering experts first developed the items, and then defined the competences that each question was trying to address.
  - Solution: Develop items tailored to the learning outcomes to be measured, applying optimal forms of item design.
- Multiple competences were measured with one question (one scoring gui one score)
  - Solution: There should be one scoring guide and one score for each learning outcome to be measu
- The problem of defining levels of achievement is still unsolved.
  - Solution: Develop meta-rubrics with benchmarks of achievement (while narrowing down assessable competences).

Keep it simple, but informative!

REVISING THE CONSTRUCTIVE RESPONSE TASK
MORE INTENTIONAL AND CLOSER ALIGNMENT OF TASKS TO LEARNING OUTCOMES.

- Decomposition of questions that assess multiple learning outcomes: an example.
  - 2016
    - Describe the merits and demerits of a system depicted in Figure X. (BES1.EA1.EA6).
  - 2017
    - 1) Pick one functional element and explain how it is functionally related to X (BES3).
    - 2) Select one element of system performance and explain what it depends on based on the figure (EA1).
    - 3) Explain possible causes behind an issue and how you would resolve the issue.
    - 4) Propose one way to increase the function of a system by X times and explain the design features you would add to assure safety (ED1, EP6; prepare two scoring guides).

DISCUSSION ON THE RECONSTRUCTION OF A MORE PARSIMONIOUS ASSESSMENT FRAMEWORK. A MODIFIED ANDERSON AND KRATHWOHL'S TAXONOMY OF EDUCATIONAL OBJECTIVES

Knowledge Dimension	Cognitive Process Dimension							
	Remember (1)	**Understand** (2)		**Apply** (3)		**Analyze** (4)	**Evaluate** (5)	**Create** (6)
Facutual (A)		BES1						
Conceptual (B)		BES2		EA2(1)		EA2(2)		EA1(2)
		BES3		ED1(1)		EA5(2)		ED1(2)
		EP2				EA6		ED3(1)
Procedural (C)		ED2(1)		EA1(1)	EA4			
		EP3		EA3	EA5(1)			
		EP5		EP1	ED2(1)			
				EP8	ED3(2)			
Social/inter-disciplinary		EP4	EGS1					
		EP6	EGS2					
		EP7	EGS3					
			EGS4					

Fukahori, S. (2018). "Towards a Theory of Disciplinary Relationships – A proposal of an Analytical Framework of Disciplinary Learning Outcome Reference Points." In *Educational Studies in Japan International Yearbook.* No.12 pp.61-75.

LARGE SCALE IMPLEMENTATION 2017

- 9 universities (Japanese and Indonesian) & 409 students
- October 1–November 30 2017
- Test items
  - Multiple choice questions (15 items, 45 minutes)
  - Constructive response task (Revised "Machine Tools," 50 minutes)
  - Contextual Survey (10 minutes)
- Feedback reports delivered to the project team, individual participating universities and individual participating students.

RESULTS OF LARGE SCALE IMPLEMENTATION 2017 WEAK CORRELATION FOUND AMONG SOME ITEMS MEASURING THE SAME AREA OF KNOWLEDGE AND SKILLS

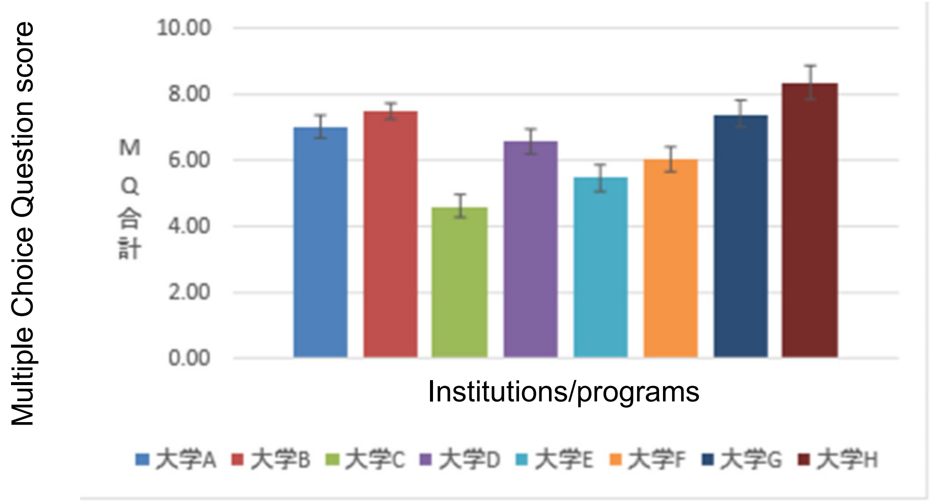
Categorical Correlation Coefficients between MCQs

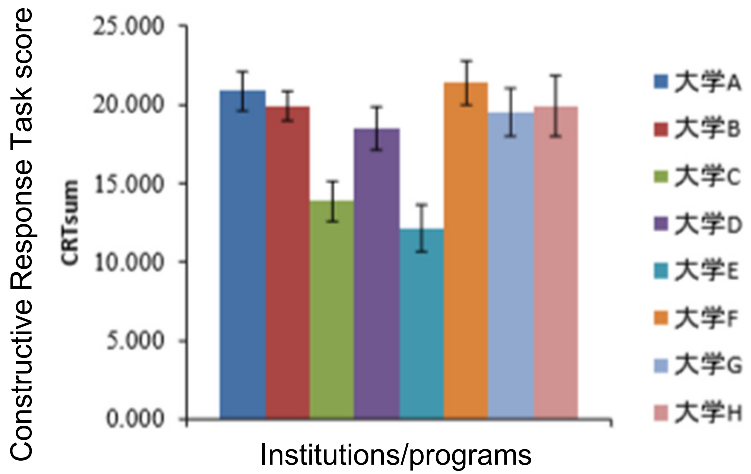
		%Correct	MQ1	MQ2	MQ3	MQ4	MQ5	MQ6	MQ7	MQ8	MQ9	MQ10	MQ11	MQ12	MQ13	MQ14	MQ15
Mathematics	MQ1	62.4%	-														
	MQ2	41.0%	.27 **	-													
Physics	MQ3	43.1%	.14	-.01	-												
	MQ4	43.8%	.29 **	.31 **	.29 **	-											
Strength of Materials	MQ5	4.1%	-.05	.30 +	.08	.37 *	-										
	MQ6	43.4%	.10	.07	.13	.26 **	.18	-									
Motion and Vibration	MQ7	55.2%	.00	.12	.16 +	.26 **	.14	.10	-								
	MQ8	33.8%	.07	.14	.14	.00	.29 +	.08	.05	-							
	MQ9	50.3%	.09	.07	.02	.20 *	.00	-.05	.20 *	-.01	-						
Energy and Flow	MQ10	49.3%	.24 **	.16 +	.16 +	.25 **	-.09	.08	.15 +	.06	.09	-					
	MQ11	48.3%	.06	.08	.15	.19 *	.22	.20 *	.25 **	.06	-.14	.13	-				
Information and Measurement Control	MQ12	49.0%	.01	-.23 *	-.05	-.18 *	-.09	-.17 +	-.05	-.07	-.05	-.15 +	-.06	-			
	MQ13	31.0%	.07	.03	.20 *	.09	.14	.10	.16 +	-.09	.07	-.13	-.04	-.05	-		
Processing and Production	MQ14	72.8%	.20 *	-.02	.29 **	.36 **	.17	.03	.22 *	.28 **	-.01	.10	.06	-.01	-.07	-	
	MQ15	34.5%	-.16 +	.12	.07	.05	.19	.22 *	-.07	-.07	-.03	-.12	.18 +	-.02	.05	-.19 +	–

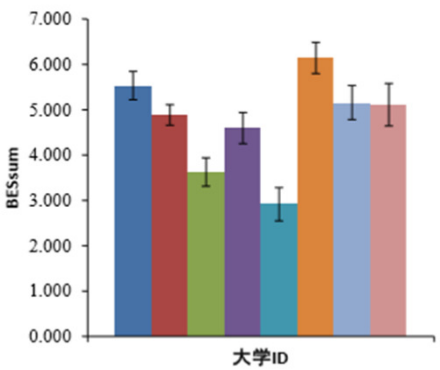
** $p < .01$, * $p < .05$, + $p < .10$

RESULTS OF LARGE SCALE IMPLEMENTATION 2017

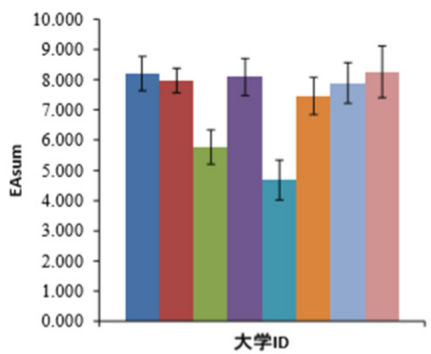
Analysis by Yugo Saito



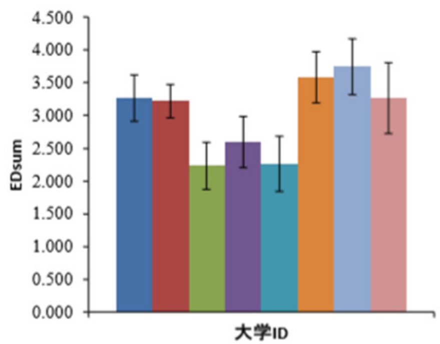




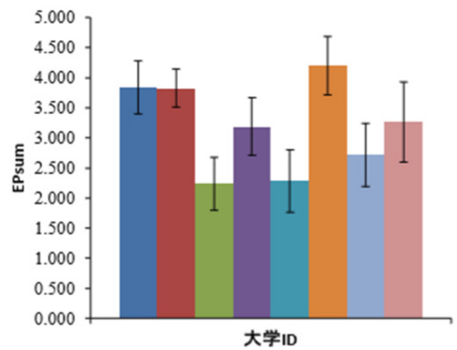
Basic & Engineering Science



Engineering Analysis



Engineering Design

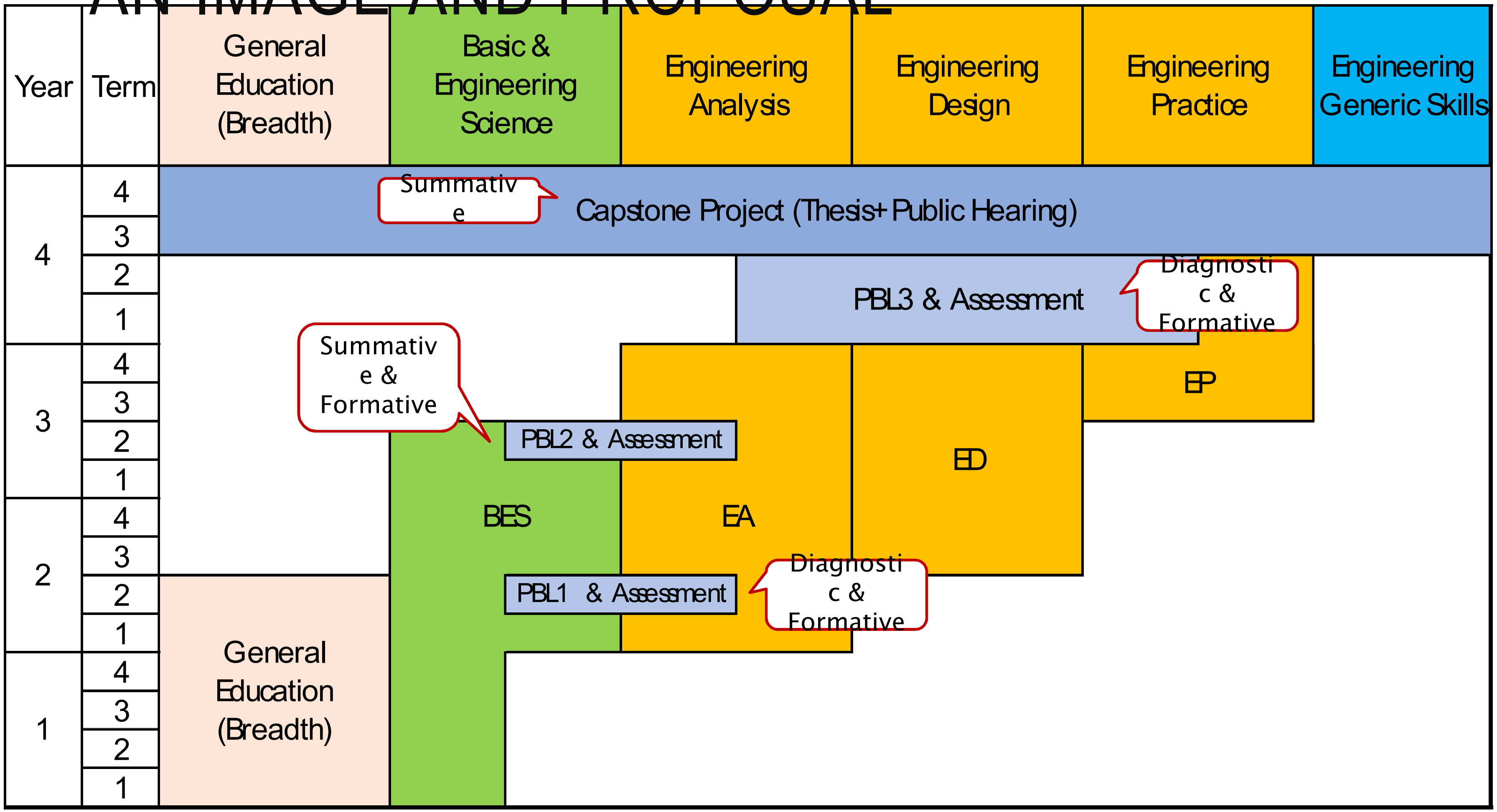


Engineering Practic

FUTURE PLANS

- Further sophistication of items.
  - Larger stock of high quality MCQs in each area of study.
  - Parsimonious assessment framework.
  - Closer alignment of CRTs to learning outcomes.
- Embedding the test items in coursework.
  - Backward design of curriculum.
  - Pivotal Embedded Performance Assessment, PEPA.
- Project Transfer from NIER to Japan Society of Mechanical Engineering, JSME.
  - Collaboration with industry to ensure societal relevance.

PIVOTAL EMBEDDED PERFORMANCE ASSESSMENT (PEPA) AN IMAGE AND PROPOSAL



Thank you for your attention!

Contact Information:
The Tuning Test Item Bank in Mechanical Engineering,
National Institute for Educational Policy Research (NIER), JAPAN
http://www.nier.go.jp/tuning/centre.html
fukahori@ueii.kyushu-u.ac.jp