

Combining course- and program-level outcomes assessments through embedded performance assessments at key courses: A proposal based on the experience from a Japanese dental education program

Kayo Matsushita, Kazuhiro Ono, and Yugo Saito*

doi: [http://dx.doi.org/10.18543/tjhe-6\(1\)-2018pp111-142](http://dx.doi.org/10.18543/tjhe-6(1)-2018pp111-142)

Received: 27.09.2018

Accepted: 13.11.2018

Abstract: This paper addresses how to combine the course- and program-level assessments and presents a new method illustrated by a case of dental education program in Japan. Performance assessments are considered effective for evaluating knowledge integration and higher-order skills, while placing a burden on faculty, hence their feasibility as the program-level assessment is regarded lower than standardized tests or questionnaire surveys. We have developed several performance assessments at the course level, such as Modified Triple Jump for the PBL course. Based on this experience, we propose Pivotal Embedded Performance Assessment (PEPA) as a method for combining assessment at the course and program levels. The method limits the range of performance assessment to key courses directly linked to program goals and placed at the critical juncture points of curriculum, while entrusting the assessment of other courses to expert judgment of individual teachers.

* Kayo Matsushita (corresponding author, matsushita.kayo.7r@kyoto-u.ac.jp), PhD, is a professor at the Center for the Promotion of Excellence in Higher Education and the Graduate School of Education, Kyoto University, Japan.

Kazuhiro Ono (k-ono@dent.niigata-u.ac.jp), PhD, is a professor at the Division of Oral Science for Health Promotion and the Division of Dental Educational Research Development (concurrent post), Graduate School of Medical and Dental Sciences and a chair of the Student Affairs of the Faculty of Dentistry, Niigata University, Japan.

Yugo Saito (ugo.saito@gmail.com, y-saito@pt-u.aino.ac.jp), PhD, is an assistant professor at the Department of Physical Therapy, Faculty of Health Science, Aino University, Japan. He is a member of the Japanese follow-up program “The Assurance of Higher Education through the Development of a Tuning Test Item Bank Global Quality.”

More information about the authors is available at the end of this article.

Acknowledgements: This work was supported by JSPS KAKENHI Grant Numbers JP15H03473 and JP18H00975.

PEPA consists of the following procedures: systematization of curriculum and selection of key courses; design and implementation of performance assessments by a faculty team; setting passing criteria with incorporating the function of formative assessment; certifying the completion of the degree program. PEPA thus enables maintaining assessment feasibility and compatibility with a credit system, while ensuring assessment validity and reliability.

Keywords: Performance Assessment; Embedded Assessment; Program-Level Assessment; Curriculum; PBL (Problem-Based Learning); Dental Education.

I. Problem and Purpose

I.1. Diversity of learning outcomes assessments

The question of how to assess learning outcomes is currently an important issue in higher education across many countries.

The variety of learning outcomes assessments has been increasing in recent years and can be classified into (1) direct and indirect assessment, (2) qualitative and quantitative assessment, and (3) assessment at the course/program/institution level.¹ Of these, the difference between direct and indirect assessment is whether the assessment method is based on direct or indirect evidence. For example, students can demonstrate their knowledge and skills either directly (through what they really know and can do), which constitutes a direct assessment, or indirectly via a self-report (through what they think they know and can do), which constitutes an indirect assessment.²

Classifying along these axes makes it easy to grasp the characteristics of assessment methods. For example, K. Matsushita³ used the intersection of (1) and (2) to elucidate the characteristic features of four types of learning outcomes assessments (Figure 1).

Also, Middle States Commission on Higher Education,⁴ one of American accreditation associations, employed the intersection of (1) and (3) to

¹ Kayo Matsushita, "Making Learning Outcomes Visible," *Japanese Journal of Higher Education Research* 20 (2017): 94-96 [in Japanese].

² Trudy W. Banta and Catherine A. Palomba. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*, 2nd ed. (San Francisco: Jossey-Bass, 2015), 93-144.

³ Matsushita, "Making Learning Outcomes Visible," 102.

⁴ Middle States Commission on Higher Education, *Student Learning Assessment: Options and Resources*. 2nd ed. (2007), 29, https://www.msche.org/publications/SLA_Book_08080728085320.pdf.

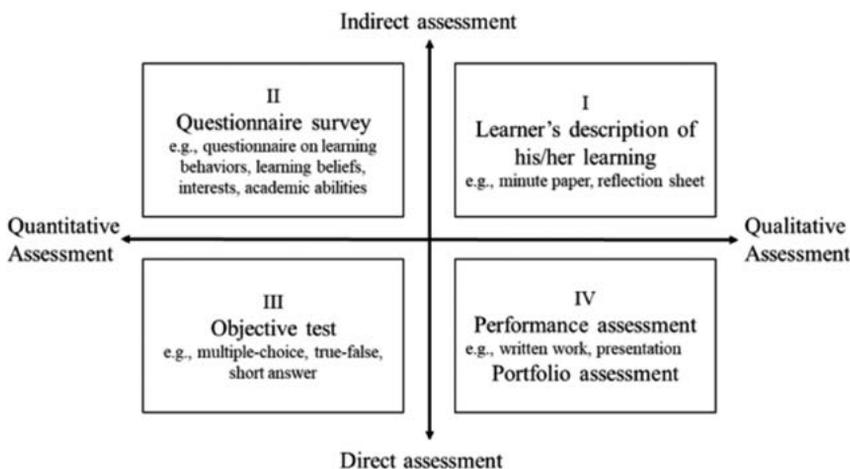


Figure 1

Four Types of Learning Outcomes Assessment

categorize assessments according to direct vs. indirect measures and institution-, program- and course-level assessments.

Use and integration of such a variety of assessment methods for different purposes allows for a multifaceted understanding of student learning, which connects with educational improvements. This approach is increasingly being adopted by universities across different countries.

I.2. The assessment gap between course, program and institution levels

However, looking closely at the assessment methods that are actually being used, large differences between countries become apparent. For example, the assessment of student learning outcomes at the institution level in the U.S. mostly utilizes three tools, namely national student surveys (85%), rubrics (69%) and classroom-based performance assessments (66%), which are regarded as “most valuable or important” approaches for assessing undergraduate student learning outcomes.⁵ On the other hand, in Japan it is

⁵ George D. Kuh et al., *Knowing What Students Know and Can Do: The Current State of Student Learning Outcomes Assessment in U.S. Colleges and Universities* (Champaign: National Institute for Learning Outcomes Assessment, 2014), 12-13, <http://www.learningoutcomeassessment.org/documents/2013%20Abridged%20Survey%20Report%20Final.pdf>.

standardized tests (32.2%), questionnaire surveys (20.2%), learning portfolios (12.7%), and rubrics (6.8%) that are used as “assessment methods of student learning outcomes at a program level.”⁶ With the growing demands on assessment at program and institution levels in recent years, standardized tests (type III in Figure 1) and questionnaire surveys (type II) have been increasingly employed. However, the utilization of learning portfolios and rubrics (type IV) at program and institution levels remains limited to some universities.

Why are the qualitative and direct assessment methods (type IV) not used more often for program- and institution-level assessments in Japan? Certainly, at the course level there are many kinds of performance measures in use in Japan, including products (e.g., essays, art works) and demonstrations (e.g., oral presentations, simulations), but they are not connected with program- and institution-level assessments due to the lack of knowledge and skills as well as human and time resources.

In contrast, the criticism of standardized tests in the U.S. led the Association of American Colleges and Universities (AAC&U) to develop Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics,⁷ which increased the share of program- and institution-level via classroom-based performance assessments and rubrics.⁸ According to T. W. Banta and C. A. Palomba, “The [VALUE] rubrics were developed to help link the assessment work done by faculty in individual classrooms to the assessment work that is often done separately by faculty and evaluators at the program or institution level”⁹ and they are considered to fulfil these functions in practice.

Nonetheless, even in the U.S. the question of how to implement assessments at the program and institution level remains problematic. Based on the survey results of assessment practices at the program level, P. Ewell, K. Paulson, and J. Kinzie find that locally developed assessments at universities, such as capstone courses and rubrics, are used more often than standardized tests, although a great deal of variation exists across disciplines.¹⁰

⁶ Ministry of Education, Culture, Sports, Science and Technology, “Japanese University Reforms Including Those in Educational Contents,” (November, 2017), 22, http://www.mext.go.jp/a_menu/koutou/daigaku/04052801/_icsFiles/afieldfile/2017/12/13/1398426_1.pdf [in Japanese].

⁷ Terrel Rhodes, *Assessing Outcomes and Improving Achievement: Tips and Tools for Using the Rubrics* (Washington D.C.: Association of American Colleges and Universities, 2009).

⁸ Kuh et al., *Knowing What Students Know*, 33.

⁹ Banta and Palomba, *Assessment Essentials*, 101.

¹⁰ Peter Ewell, Karen Paulson, and Jillian Kinzie. *Down and In: Assessment Practices at the Program Level*. (Champaign: National Institute for Learning Outcomes Assessment, 2014),

1.3. Purpose and outline

Our purpose in this paper is to introduce the state of learning outcomes assessment in Japan, to discuss the current assessment research regarding how to connect the course- and program-level assessments, and to present a concrete proposal based on the experience from the Faculty of Dentistry at Niigata University, Japan.

Niigata University is one of Japan's 86 national universities and its Faculty of Dentistry is known for being at the forefront of dental education. In our paper, we present some of the results of practical research related to assessment and curriculum that stems from the collaboration between a faculty member from the Faculty of Dentistry at Niigata University and specialists in higher education assessment over the past few years.

We describe performance assessment in a course Problem-Based Learning (PBL), one of the key courses in Niigata University's dental education program, as an example of assessment at the course level. There have been various approaches related to PBL assessment,¹¹ but we will here propose a new method of performance assessment of PBL called "Modified Triple Jump (MTJ)." We will also focus on several key courses besides the PBL course and illustrate how their performance assessments in fact connect with program-level assessment.

Course assessments that do double duty, providing information not only on what students have learned in the course but also on their progress in achieving program or institutional goals are called "embedded assessment."¹² Using this concept in this paper, we will investigate how to realize program-level assessment by arranging and integrating several course-embedded performance assessments. These practical findings form the basis of the

<http://www.learningoutcomeassessment.org/documents/NILOASurveyreport2011%20-%20Down%20and%20In%202010-20.pdf>.

¹¹ Ntombifkile G. Mtshali and Lyn Middleton, "The Triple Jump Assessment: Aligning Learning and Assessment," in *New Approaches to Problem-Based Learning: Revitalising Your Practice in Higher Education*, eds. Terry Barrett and Sarah Moore (New York: Routledge, 2011), 187-200; Tracy Winning, Elaine Lim, and Grant Townsend, "Student Experiences of Assessment in Two Problem-Based Dental Curricula: Adelaide and Dublin," *Assessment & Evaluation in Higher Education* 30, no. 5 (2005): 489-505.

¹² Linda Suskie, *Assessing Student Learning: A Common Sense Guide*, 2nd ed. (San Francisco: Jossey-Bass, 2009), chap. 2, Kindle; Rhoda Cummings, Cleborne D. Maddux, and Aaron Richmond, "Curriculum-Embedded Performance Assessment in Higher Education: Maximum Efficiency and Minimum Disruption," *Assessment & Evaluation in Higher Education* 33, no. 6 (2008): 599-605.

discussion, through which we intend to contribute to the development of international assessment research.

According to T. W. Banta and C. A. Palomba, the practical use of performance measures in program-level assessment can be largely categorized into two methods.¹³ The first one selects samples of courses from the program, then selects samples of student work in the chosen courses, and finally the faculty team conducts a second scoring for the purpose of program assessment. The second method directly aggregates assessment results of each course by each teacher for the purpose of program assessment. They introduce, as one typical example, the so-called All-in-One approach developed at Prince George's Community College, which uses digital technology in order to integrate assessment of course, program, and general education outcomes and connect outcomes assessment with classroom grading.¹⁴

The first method is used in many American universities, but it imposes a great assessment burden on faculty members, hence the assessment feasibility is considered to be low. In contrast, the second method seems to be more efficient and promising, but since there are only few cases, it is considered to be in its trial stage. The Faculty of Dentistry at Niigata University utilizes the second method; however, it differs from the All-in-One approach and is quite innovative.

Below we will first introduce the MTJ, which represents a new performance assessment method for PBL courses. Then we will discuss the alignment between curriculum and assessment, which underpins assessment at the course and program levels. Based on that, we will propose a method of integrating course-level and program-level assessments, and after comparing various program-level assessments we will finally demonstrate the effectiveness of our postulated method – Pivotal Embedded Performance Assessment.

II. PBL as a key course and its performance assessment

II.1. PBL at the Faculty of Dentistry at Niigata University

The Faculty of Dentistry at Niigata University consists of two departments: the Department of Dentistry and the Department of Oral Health and Welfare.

¹³ Banta and Palomba, *Assessment Essentials*, 104-5.

¹⁴ W. Allen Richman and Laura Ariovich, *All-in-One: Combining Grading, Course, Program, and General Education Outcomes Assessment* (Champaign: National Institute for Learning Outcomes Assessment, 2013), [http:// learningoutcomesassessment.org/documents/Occasional%20Paper%2019%20FINAL.pdf](http://learningoutcomesassessment.org/documents/Occasional%20Paper%2019%20FINAL.pdf).

Study at the Dentistry department requires six years, whereas the Oral Health and Welfare takes four years, and the class size is 40 and 20 students, respectively. The Dentistry department formulates its educational goals as “nurturing skills for solving various current problems in the rapidly changing modern society while properly collaborating with persons concerned, as well as providing high dental clinical competences for practicing holistic medicine.” In order to realize these goals both departments actively adopt PBL from early stages. Furthermore, several key courses, directly linked with these educational goals, implement performance assessments with rubrics.

PBL at the Faculty of Dentistry at Niigata University is based on the model developed by the Faculty of Odontology at Malmö University in Sweden.¹⁵ In this model, students form groups of seven to eight with teachers as facilitators and PBL is conducted in the following steps.¹⁶

1st step: Group learning in the classroom

- (1) First, students identify the facts from the case written in a form of scenario.
- (2) They discuss questions and ideas related to these facts and make solution strategies (hypotheses) linked to the problems included in the scenario.
- (3) Next, students confirm what kind of knowledge they lack in order to examine their own hypotheses and they set their learning tasks.

2nd step: Individual learning outside the classroom

- (4) Outside the classroom, students then individually conduct investigations related to the learning tasks.

3rd step: Group learning in the classroom (one week later)

- (5) One week later, the student groups combine the individual findings and integrate their pre-existing knowledge with their new knowledge, which was acquired through investigations.
- (6) Students examine whether their originally proposed solution strategies (hypotheses) are valid or not, and create their final version of the solution strategies.

¹⁵ Madeleine Rohlin, Kerstin Petersson, and Gunnel Svensäter, “The Malmö Model: A Problem-Based Learning Curriculum in Undergraduate Dental Education,” *European Journal of Dental Education* 2 (1998): 103-114.

¹⁶ Kazuhiro Ono and Kayo Matsushita, “PBL Tutorial Linking Classroom to Practice: Focusing on Assessment as Learning,” in *Deep Active Learning: Toward Greater Depth in University Education*, ed. Kayo Matsushita (Singapore: Springer, 2017), 185-86.

PBL's effects are listed as follows: acquisition of a body of integrated deep knowledge and understanding, fostering the ability to analyze and solve problems, cultivation of interpersonal skills and nurturing a desire to continually learn.¹⁷

II.2. *Development of performance assessment method for PBL: Modified Triple Jump*

In order to assess the problem-solving skills that students acquired through PBL, we developed the Modified Triple Jump (MTJ). The Triple Jump as a method for the assessment of problem-solving and self-directed learning skills in PBL was designed in 1975 by the Department of Medicine at McMaster University in Canada.¹⁸ A student and a teacher conduct PBL one on one, similarly to the learning process of regular PBL comprising of three steps, replacing the usual student group learning in the 1st and 3rd steps with a teacher-student interaction, thereby assessing the student. In order to examine their solution strategy, the student visits libraries, for example, in the 2nd step, collects reliable information, and engages in individual learning. Afterwards, in the 3rd step, the student returns to the classroom, integrates the knowledge obtained with the pre-existing knowledge, and explains his/her final version of solution strategy to the teacher.

Assessment progresses through the same process as the regular PBL, so that assessment validity, in particular, face validity is considered high. Moreover, scenarios are created and examined through cooperation of various experts, so that the content validity is deemed to be maintained. On the other hand, the assessment reliability is generally considered to be low for the following reasons: it is subjective; there is no assessor who can observe the interaction between a teacher and a student; the teacher could miss the student's verbal explanations; the assessment results are easily influenced by the quality of the material (case) used for assessment, student personality, as well as the level of assessor's proficiency.¹⁹ As for assessment feasibility, it has been pointed out that triple jump not only requires time for

¹⁷ Howard S. Barrows, "The Essentials of Problem-Based Learning," *Journal of Dental Education* 62 (1998): 630.

¹⁸ Jennifer M. Blake, Geoffrey R. Norman, and E. Kinsey M. Smith, "Report Card from McMaster: Student Evaluation at a Problem-Based Medical School," *The Lancet* 345 (1995): 899-902.

¹⁹ Mtshali and Middleton, "The Triple Jump Assessment," 199.

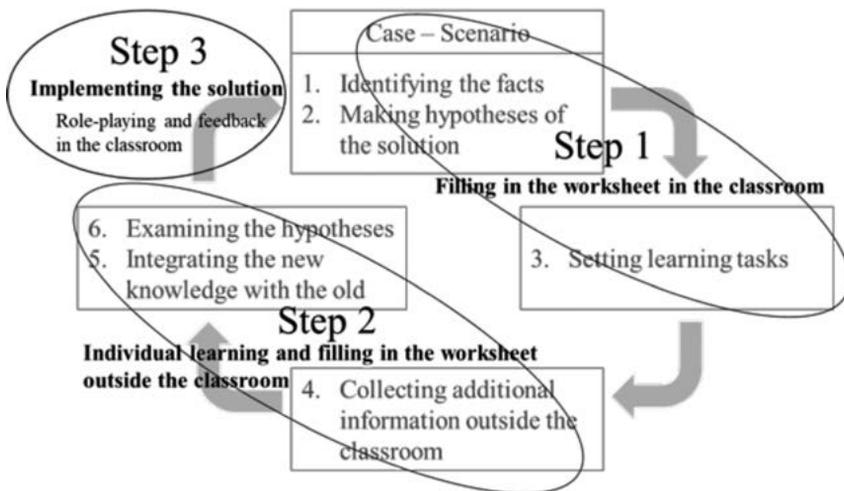


Figure 2

Process and Steps of the Modified Triple Jump (Ono and Matsushita 2017, 194)

learning, but also that assessment is time-intensive and thus places a burden on teachers.²⁰ As a result, there are few universities that currently conduct triple jump as an assessment method of PBL. However, since there is no other assessment method that could replace triple jump and at the same time fulfil the criteria of validity, reliability and feasibility, our study aims to improve upon triple jump.

The whole process of our MTJ is depicted in Figure 2. In step 1, like the original triple jump, a student identifies problems, proposes hypothetical solution strategies and sets learning tasks, but, unlike the original triple jump, he/she also describes the process on a worksheet, all within a 60-minute time limit. In step 2, the student not only investigates learning tasks, but also based on the findings he/she also examines the solution strategies and within a week formulates a final version of solution strategy, while recording the process on a worksheet. That is to say, MTJ can be largely characterized by reducing the original triple jump to steps 1 and 2, replacing verbal assessment with written assessment, and also by employing rubrics for the purpose of assessment. In addition, step 3 of MTJ is newly

²⁰ Mark J. Newman, "Problem Based Learning: An Introduction and Overview of the Key Features of the Approach," *Journal of Veterinary Medical Education* 32 (2005): 17.

designed and added, where the student is assessed through role play with the teacher as a simulated patient acting out scenario situations, and the assessment is continuously conducted all the way until the implementation of solution strategy when another rubric is adopted. Immediately after the role play in step 3, feedback on assessment results is provided for the total duration of 15 minutes.²¹

Let's be more specific. For example, in MTJ within the course titled "Link between oral cavity and the body" a student, assigned the scenario shown in Figure 3, conducts written tasks based on the worksheet (steps 1 & 2) and a demonstration task based on the role play (step 3), while teachers conduct assessment according to two different rubrics. The rubric for the worksheet in steps 1 & 2 (Table 1) comprises of six dimensions from "identifying a problem" to "proposing a solution" that are assessed on a four-level scale. The seventh dimension called "implementing a solution" is assessed in step 3, where the rubric (Table 2) contains four sub-dimensions such as "sympathetic attitudes." We showed the rubric for steps 1 and 2 to the students before they tackled the task; however, we didn't show them the one for step 3 because it is a task-specific rubric.

"I don't know what to do"

You are a starting dentist at the Niigata University Hospital.
Today your patient is Yuzo Takahashi (67 yrs, male) comes for his second visit. He has a medium case of periodontal disease. During his first visit he was given a diagnosis and an explanation of his condition.

You : Mr. Takahashi, how is your condition? Last time we talked about quitting smoking, but have you stopped?

Takahashi : No, I haven't. In my case, I think it's better to die than to quit smoking. The other day, I even filled in the government questionnaire that "I cannot live without smoking." I still smoke 40 cigarettes a day. I don't get why even at the dentist I am told to stop smoking.

You : But you also have diabetes, so I really think it would be best to stop...

Takahashi : Diabetes has nothing to do with it. I am at a dentist office, or are you my doctor?

You : Well I am a dentist. But I am going to tell you the same thing as the last time, you should first stop smoking, alright?

Takahashi : I see, I see. You hate smoking.

Figure 3
Scenario Example of a Modified Triple Jump

²¹ Ono and Matsushita, "PBL Tutorial," 193-94.

Table 1
Rubric for Steps 1 and 2 of the Modified Triple Jump (partial excerpt)

Dimensions	1. Identifying a problem	2. Conceiving solution strategies	3. Setting learning tasks	4. Learning results and resources	5. Examining solution strategies	6. Proposing a solution
Explanation of dimensions	Identifies the problem based on the facts of the scenario.	Determines the objective of the solution and proposes a number of solution strategies.	Sets out the necessary learning tasks to solve the problem.	Learning tasks undertaken using credible resources.	Considers the effectiveness and feasibility of the solution strategies.	Proposes a solution to the problem.
Level 3
Level 2	Identifies and explains the problem based on the facts of the scenario.	Proposes a number of solution strategies and explains the process by which they were developed.	Identifies learning tasks and explains their necessity from their relation to the proposed solution strategies, but misses some key learning tasks.	Selects resources based on their credibility and generally obtains correct information.	Compares a number of solution strategies with regard to the effectiveness and feasibility of each.	Proposes a reasonable solution appropriate for the scenario situation.
Level 1
Level 0	Students not satisfying the Level 1 criterion shall be given a zero.					

Note: The descriptors of levels 3 & 1 are omitted here.

Table 2
Rubric for Step 3 of the Modified Triple Jump (partial excerpt)

7. Implementing a solution				
Dimensions	7-1. Gathering additional information (gathering additional information and reformulating the problem)	7-2. Integration of additional information and correction of the preexisting solution)	7-3. Sympathetic attitudes (sympathy for a partner)	7-4. Communication (expressing the solution in the way that partner can grasp)
Explanation of dimensions	To persuade the patient to stop smoking, the student gathers additional information through a conversation with the patient and, if necessary, reformulates the problem.	To persuade the patient to stop smoking, the student integrates useful information with additional information and, if necessary, modifies the proposed solution.	The student urges the patient to stop smoking while respecting patient's thinking and values.	The student explains his or her thinking to the patient in simple terms.
Level 3
Level 2	The student gathers some of the required additional information, such as patient's needs of periodontal treatment, investigates patient's claim "I cannot live without smoking," and collects information about diabetes and other ailments.	Upon integrating the additional information from the patient, the student achieves an adequate understanding of the importance of quitting smoking for periodontal treatment, not only due to the link between periodontal disease and smoking, but also due to the link between periodontal disease and diabetes.	After acknowledging patient's claim "I cannot live without smoking," the student urges the patient to stop smoking while paying attention to patient's feelings.	The student largely considers the topics and their organization, achieves patient's understanding, but there are some problems regarding the structure of the communication toward the patient.
Level 1
Level 0	Students not satisfying the Level 1 criterion shall be given a zero.			

Note: The rubric for step 3 of the Modified Triple Jump is task-specific, hence its descriptors depend on the scenario content. Here we display the rubric of step 3 for the scenario shown in Figure 3. The descriptors of levels 3 & 1 are omitted here.

The results of implementing PBL courses in the Oral Health and Welfare department over the past five years have revealed the following advantages of MTJ when compared with the original triple jump. First, in regard to assessment reliability, the assessment conducted by three teachers utilizing rubrics in steps 1 & 2 and step 3, respectively, showed an overall high level of absolute agreement between the assessors in each dimension, indicating a sufficient level inter-rater reliability (average ICC(2,3) = .76).²² As for the assessment feasibility, due to the introduction of worksheets in steps 1 & 2, many students could take the examination simultaneously, thereby considerably reducing the time demands on teachers for face-to-face assessment tied to the assessment setting. In addition, the results of student free-answer questions in the aftermath of MTJ implementation showed numerous positive responses, such as “In order to gain understanding of the disease from the simulated patient, I closely examined the disease and deepened my own understanding,” “Through conducting step 3, I could understand how PBL can become helpful in my future workplace,” so that the assessment was not merely an “assessment of learning” but it itself became a learning experience for students, or “assessment as learning.”²³ Based on the above, we can assert that MTJ is a well-designed assessment method for PBL.

III. Alignment of assessment with curriculum

III.1. Outline of the curriculum and assessment

The educational goals of the Faculty of Dentistry at Niigata University stated earlier can be further divided into 24 items: seven items within “knowledge & understanding,” six items within “specialized skills (subject-specific skills),” eight items within “generic skills,” and three items within “attitudes and values.” These are the *intended learning outcomes*²⁴ of the dental education program.

²² Kazuhiro Ono, Kayo Matsushita, and Yugo Saito, “Prospects for Direct Assessment of Problem Solving Competence: Development of Modified Triple Jump in Problem-Based Learning,” *Journal of the Liberal and General Education Society of Japan* 36, no. 1 (2014): 128-29 [in Japanese].

²³ Alverno College Faculty, *Student Assessment-as-Learning at Alverno College* (Milwaukee: Alverno College Institute, 1994); Lorna M. Earl, *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning* (Thousand Oaks: Corwin Press, 2003); Ono and Matsushita, “PBL Tutorial,” 183-84.

²⁴ John Biggs and Catherine Tang, *Teaching for Quality Learning at University*, 4th ed. (Berkshire: The Society for Research into Higher Education & Open University Press, 2011), chap. 7, Kindle.

At institution level, Niigata University aims to equip “skills for identifying and solving problem,” “ability to learn independently knowledge and skills necessary for solving problems,” and “communication skills for collaborative tackling of problems” to all of its graduates. These goals are included in all individual educational programs, and the intended learning outcomes are therefore subdivided into the categories of “knowledge & understanding,” “specialized skills (subject-specific skills),” “generic skills,” and “attitudes and values.” The intended learning outcomes of the dental education program are established in accordance with the institutional diploma policy, but they are also consistent with the “core” that must be followed by all dental universities in Japan, namely “the basic qualities and abilities required of a dentist” listed in Model Core Curriculum for Dental Education.²⁵ In other words, the intended learning outcomes of the Faculty of Dentistry at Niigata University are on the one hand institutional goals, and on the other hand they transcend the institution, as they have been designed in accordance with the field goals on a national level.

In order for the students to acquire such qualities and abilities, we link the intended learning outcomes with each course on the curriculum map, and compose the dental education program containing general education (Figure 4). As shown in this curriculum tree, the program takes six years, and based on its learning content it can be largely divided into four stages: 1st school year to early 2nd school year, late 2nd school year to 3rd school year, 4th school year to early 5th school year, and late 5th school year to 6th school year. Also, courses are classified into the following eight groups: liberal arts, English, study/research skills, basic oral science, clinical dentistry, integration of knowledge and skills, professionalism, and international activities.

The first stage emphasizes “transformation to autonomous learning and studying liberal arts.” In the course “University study skills” we attempt to transform learning attitudes of freshmen, cultivate problem-solving skills and ability to think logically and communicate effectively necessary for completing the coming dental education program, and assess the learning outcomes through performance assessment.

The second stage focuses on “study of basic oral science and gaining self-awareness as a dentist,” during which basic oral science courses are basically delivered in a lecture-style format for knowledge acquisition, while the awareness and attitudes as a medical professional are cultivated through patient

²⁵ Model Core Curriculum Revision Coordination Committee and Model Core Curriculum Revision Specialist Research Committee. *Model Core Curriculum for Dental Education: AY 2016 Revision* (2017), http://www.mext.go.jp/component/b_menu/shingi/toushin/_icsFiles/afieldfile/2017/07/07/1383961_02_3.pdf [in Japanese].

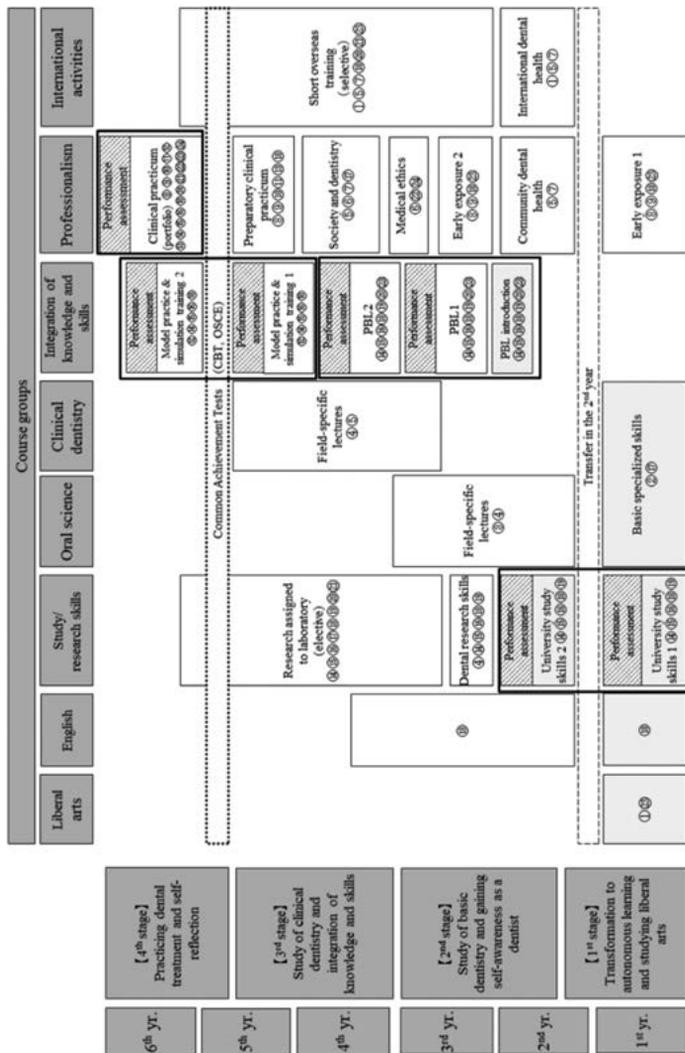


Figure 4
Outline of the Curriculum and Assessment

contact in the “Early exposure” courses. Moreover, in order to integrate the knowledge gained from lectures and foster problem-solving skills, PBL is implemented in parallel with related lectures, and its learning outcomes are assessed through performance assessment, that is, the above described MTJ.

The third stage concentrates on “study of clinical dentistry and integration of knowledge and skills,” during which clinical dentistry courses are largely delivered in a lecture-style format. At the same time, PBL from the second stage is continued, and “Model practice & simulation training” is newly implemented, thereby integrating knowledge with skills, including those of basic oral science, as well as fostering problem-solving skills that are highly specialized. Hence, these learning outcomes are assessed by means of a different type of performance assessment.

The fourth stage is dedicated to “practicing dental treatment and self-reflection.” First, the Common Achievement Tests Organization (CATO) administers Common Achievement Tests for Dental Students Prior to Clinical Clerkship, assessing whether the qualities and abilities for conducting clinical training are acquired through Computer-Based Testing (CBT) for knowledge and Objective Structured Clinical Examination (OSCE) for skills and attitudes. Upon passing these tests, students are admitted to join “Clinical practicum,” where they improve dental clinical competences through their experience with patient care. The assessment of dental clinical competences is conducted via e-portfolio continuously and formatively, whereas performance assessments are carried out as the direct assessment of patient care at the end of “Clinical practicum.”

III.2. *Performance assessments at key courses*

Within this curriculum the following courses are placed as key courses in each stage: “University study skills” as a general education course; “PBL,” “Model practice & simulation training,” and “Clinical practicum” as courses in the major at the Faculty of Dentistry (Figure 4). Based on the integration of previously acquired knowledge and skills, these key courses require generic problem solving skills as well as subject-specific problem solving skills, both of which are departmental educational goals as mentioned above. The items under “Problem solving skills” in Table 3 represent each dimension of AAC&U’s problem solving VALUE rubric,²⁶ while “Dental clinical competences” represent problem solving skills in the field of dentistry.

²⁶ Rhodes, *Assessing Outcomes*, 40-41.

Table 3
Fostering and Assessing Subject-specific Problem Solving Skills in a Dental Education Program

Problem solving skills	Dental clinical skills	University study skills (no patient)	PBL (paper patient/ simulated patient)	Model practice & simulation training (model)	Clinical practicum (real patient)
Define problem	Information gathering & analysis	√	√	√	√
	Diagnosis	√	√	√	√
Identify strategies	Determining treatment policy	√	√	√	√
	Propose solutions/ hypotheses	√	√	√	√
Evaluate potential solutions	Designing treatment plan	√	√	√	√
	Implement solution		Step 3 of the Modified Triple Jump	√	√
Evaluate outcomes	Evaluating treatment results and revising treatment plan		Step 3 of the Modified Triple Jump	Comments/ feedback from a teacher	√

↑ Degree of specificity, comprehensiveness and authenticity

Progressing from “University study skills” (1st stage) to “PBL” (2nd stage), “Model practice & simulation training” (3rd stage), and then to “Clinical practicum” (4th stage), problem solving skills become increasingly specialized, including an increasing number of problem solving dimensions, becoming ever more comprehensive. Moving from problem solving on paper to real patients in clinical situations thus raises the authenticity.

Discussed below are the different kinds of performance assessments that are implemented within the key courses. The dimensions of all the rubrics are shown in Table 4.

Table 4

Quality Assurance of Graduates Through Embedded Performance Assessment

Dimensions		Level 3	Level 2	Level 1	Level 0
(Freshmen)	University study skills 1 & 2				
	Background and problems		2	1	
	Claims and conclusions		2	1	
	Warrant and facts/data	2	1		
	Examination of rebuttals		2	1	
	Overall structure		2	1	
	Rules of expression	2		1	
Problem-solving/dental clinical skills	PBL 1 & 2				
	Identifying a problem		2	1	
	Conceiving solution strategies		2	1	
	Setting learning tasks		2	1	
	Learning results and resources	2	1		
	Examining solution strategies		2	1	
	Proposing a solution		2	1	
	Implementing a solution		2	1	
	Model practice & simulation training 1 & 2				
	Pathosis and diagnosis	2	1		
	Setting of treatment policy		2	1	
	Treatment plan		2	1	
	Reflection after the treatment		2	1	
	Technical terms and expressions	2	1		

Dimensions		Well done	Acceptable	Unacceptable
Seniors	Clinical practicum			
	Interviewing and gathering information			
	Diagnosis and selection of procedures			
	Preparation and use of equipment			
	Reflection after procedures			
	Consideration for patients			
	Safety of treatment			

Note: This table provides an example of learning progression (represented by ■) of one graduate as a successful case of quality assurance. Numbers 1 & 2 in the rubrics correspond to the order of course series.

III.2.1. Performance assessment at “University study skills” (1st stage)

In “University study skills” students are given essay assignments, which form the basis for the assessment of problem solving, logical thinking and written communication skills.²⁷ Teachers offer broad themes for the assignment, and out of many possible problems students take up a specific problem for their essays. Then, they not only investigate their problems, but they themselves formulate their own claims and conclusions. Teachers then assess essays using a rubric containing six dimensions and four levels.

III.2.2. Performance assessment of “PBL” (2nd stage)

As mentioned above, the MTJ employs two different rubrics for the assessment of two types of performances, namely written tasks based on scenarios about paper patients and role plays with simulated patients.

²⁷ Kayo Matsushita, Kazuhiro Ono, and Yusuke Takahashi, “Development of a Rubric for Writing Assessment and Examination of Its Reliability,” *Journal of the Liberal and General Education Society of Japan* 35, no. 1 (2013): 109-12 [in Japanese]; Kazuhiro Ono and Kayo Matsushita, “Assessment of Writing in First-Year Education,” *In Assessment of Active Learning*, eds. Kayo Matsushita and Terumasa Ishii (Tokyo: Toshindo, 2016), 28-39 [in Japanese].

III.2.3. Performance assessment of “Model practice & simulation training” (3rd stage)

In “Model practice & simulation training,” patients’ problems are identified based on model representation of patient’s oral cavity, patient scenarios, roentgen photographs and examination findings. Next, students propose appropriate treatment policies and treatment plans, which is carried out in a model, and based on the judgment of results they modify their treatment plan, and teachers then assess the content of student worksheets that recorded the process by using a four-level rubric.²⁸

III.2.4. Performance assessment of “Clinical practicum” (4th stage)

In order to assess dental clinical competences, “Clinical practicum,” where patient treatment is conducted, implements portfolio assessment as formative assessment and clinical performance assessment as summative assessment.

In portfolio assessment, students set their own target of the day before they begin dental treatment for their patients, and after the practicum they record in e-portfolios details about procedures and content of dental treatment, acquired knowledge and special skills, and cultivated attitudes and values as a medical professional. Students also conduct a self-assessment on a five-level scale, in order to assess the degree to which they could accomplish the treatment by themselves. Likewise, the teacher, who is in charge of his/her students, assesses the degree to which they could accomplish the treatment on a five-level scale and assigns them teacher comments/instructions for further learning.²⁹

As for the clinical performance assessment, at the end of the clinical practicum, students’ dental treatment, to which students apply as their final exam is assessed in a clinical setting by a professional dentist, who is also the teacher in charge of the practicum. The performance assessment is implemented under six dimensions such as “interviewing and gathering information.” These dimensions are assessed on a three-level scale, namely “well done,” “somewhat

²⁸ Nami Akiba et al., “An Introduction to the Undergraduate Comprehensive Model Practice Course at the Faculty of Dentistry, Niigata University,” *The Journal of Japanese Dental Education Association* 33 (2017): 110 [in Japanese].

²⁹ Yohei Oda et al., “Development and Use of a Web-Based E-Portfolio for Dental Clinical Training,” *The Journal of Japanese Dental Education Association* 33 (2017): 67-68 [in Japanese].

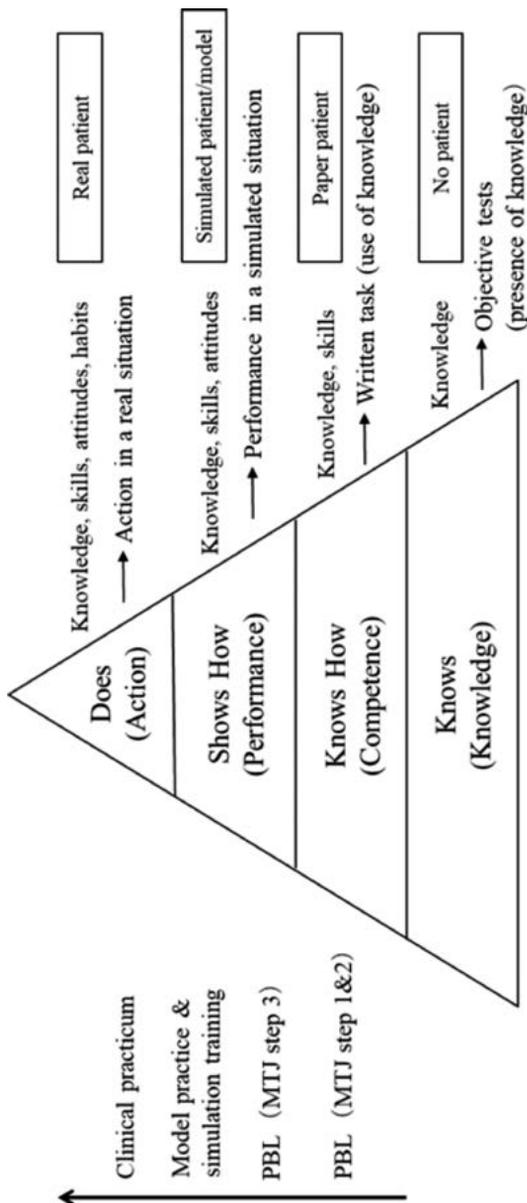


Figure 5
Relationship between Performance Assessments at Key Courses

lacking, but within the acceptable range (acceptable),” “somewhat lacking, outside the acceptable range (unacceptable).”³⁰

As described above, there is no patient in “University study skills,” whereas “PBL,” “Model practice & simulation training,” and “Clinical practicum” include a paper patient and its simulated patient based on a scenario, model of oral cavity, and real patient, respectively. Each performance is assessed through faculty-developed rubrics. Using Miller’s Pyramid,³¹ we can also represent the hierarchy of assessments of “PBL,” “Model practice & simulation training,” and “Clinical practicum,” as illustrated in Figure 5. While PBL’s written tasks regarding the paper patient, conducted within steps 1 & 2 of MTJ, relate to *Knows How* (knows how to apply acquired knowledge), role plays conducted in step 3 of MTJ and Model practice & simulation training relate to *Shows How* (shows how to apply that knowledge), and Clinical practicum relates to *Does* (actually applies that knowledge in practice). Hence, through different types of multi-layered performance assessments we can assess the increasingly higher-order dental clinical competences.

How can we then combine performance assessments at the course-level with program-level outcomes assessment? Do they really function as embedded assessments and provide information on student progress at both the course- and program-levels?

IV. Relationship between course- and program-level assessments

IV.1. *Quality assurance of graduates through embedded performance assessment*

Table 4 depicts the image of quality assurance of graduates through embedded performance assessment. Within the four key courses, “University study skills,” “PBL,” and “Model practice & simulation training” require Level 2 and above on the rubrics in order to pass, whereas “Clinical practicum” requires a level of at least “acceptable” (out of “well done,” “acceptable” and “unacceptable”) in order to pass. Each of the firstly mentioned three key courses extends over two school semesters/years and

³⁰ Noritaka Fujii et al., “Competency Assessments for Undergraduate Students in Clinical Clerkships at the Faculty of Dentistry, Niigata University,” *The Journal of Japanese Dental Education Association* 33 (2017): 6-7 [in Japanese].

³¹ George E. Miller, “The Assessment of Clinical Skills/ Competence/ Performance,” *Academic Medicine* 65, no. 9 (September Supplement 1990): S63.

includes two rounds for each, such as “PBL1” and “PBL2.” Even if passing criteria in the number 1 course (e.g., PBL1) are not achieved, it is sufficient to achieve level 2 and above in the subsequent number 2 course (e.g., PBL2). As for the last key course “Clinical practicum,” students can surpass the passing criteria of final clinical performance assessment by following the daily portfolio assessments and teachers comments (this process is recorded in the e-portfolio). In this manner, each key course’s performance assessment incorporates the function of formative assessment in the way that it is designed for students to achieve the passing criteria by the time they complete each kind of key courses.

Of course, as shown in Figure 4, many courses besides the key courses are included in the dental education program. Many of those courses aim for students’ acquisition of “knowledge” or “knowledge and skills” (see Figure 5), where the assessment is entrusted to the expert judgment by individual teachers who award credits.

This is how students are conferred the undergraduate degree through fulfilling the passing criteria in key courses and earning the credits necessary for completing the whole curriculum. In addition, by completing the undergraduate program students become eligible for the National Board Dental Examination.

IV.2. Comparison of program-level assessments

Out of the four types of learning outcomes assessment described in the outset of this paper (Figure 1), the typical assessments that have been adopted at the program level are: questionnaire surveys (type II), objective tests at the end of degree program (type III), portfolio assessments and performance assessments, such as in capstone courses.³² Table 5 organizes the respective assessment methods according to their characteristics: assessment validity, reliability, feasibility, and compatibility with a credit system, where graduation is approved only upon the acquisition of credits in all courses.

First, as substituting questionnaire survey of students with direct assessment is considered difficult,³³ we excluded it from our analysis in this

³² Suskie, *Assessing Student Learning*, chap. 1; Ewell, Paulson, and Kinzie, *Down and In*, 9.

³³ Gary R. Pike, “Limitations of Using Students’ Self-Reports of Academic Development as Proxies for Traditional Achievement Measures,” *Research in Higher Education* 37, no. 1 (1996): 89-114; Yugo Saito, Kazuhiro Ono, and Kayo Matsushita, “Correlations of Direct Measures Based on Performance Assessment and Indirect Measures Based on Student Self-report,” *Japan Journal of Educational Technology* 40 (Suppl.) (2016): 157-60 [in Japanese].

Table 5
Comparison of Program Level Assessments

Type	Program-level assessment	Validity	Inter-rater reliability	Feasibility	Compatibility with a credit system
II	Questionnaire survey (indirect, quantitative assessment)	Substitute of direct assessment is problematic.	–	High	Low
III	Objective test as add-on assessment (direct, quantitative assessment)	Suitable for assessing factual knowledge, but not necessarily for integration of knowledge and higher-order skills.	–	High	Not high (Sometimes the test results don't match with the expected results based on the acquired credits.)
IV	Portfolio assessment (direct, qualitative assessment)	Suitable for assessing learning and growth within a selected time period.	Assessor's training is required.	Medium or low (It requires second scoring. Assessment burden depends on the volume of assessment objects and the methods of second scoring.)	Not high (Sometimes the results of second scoring don't match with the expected results based on the acquired credits.)
IV	Performance assessment at all courses (direct, qualitative assessment)	Suitable for assessing knowledge integration and higher-order skills, and it can cover the whole program.	Assessor's training is required.	Low (Although second scoring is not required, assessment burden at all courses is high.)	High (The assessment result of each course can be directly used in a credit system)
IV	Performance assessment at key courses (direct, qualitative assessment)	Suitable for assessing knowledge integration and higher-order skills, but it cannot cover the whole program.	Assessor's training is required.	Relatively high (Second scoring is not required. Although assessment burden at each key course is high, the number of courses is limited.)	High (The assessment result of each course can be directly used in a credit system)

paper. Even though their feasibility is high, objective tests conducted at the end of the program as an add-on assessment are not suitable for the assessment of higher-order skills nor knowledge integration. At the same time, their compatibility with a credit system is not high. As for portfolio assessment, the requirement of second scoring by the faculty team is thought to decrease the assessment feasibility.³⁴ In case performance assessment is conducted in all courses, second scoring is not required and moreover the compatibility with a credit system is high, but the assessment burden becomes significant.

In contrast to the four methods discussed above, the method of performance assessment conducted only at key courses imposes a considerable assessment burden, but as it is limited only to those courses, its assessment feasibility is relatively high. Likewise, it is adept at directly assessing knowledge integration as well as higher-order skills, which are included in the program's goals. A prime example is the performance assessment at capstone courses. However, capstone courses cannot cover the whole program.

Our proposed method is performance assessment at key courses only by a faculty team with other courses left to the expert judgment of individual teachers, thereby we connect assessments at the course and program levels while covering the whole curriculum. We call the method of embedded performance assessment at key courses "Pivotal Embedded Performance Assessment (PEPA)."

Considering the assessment burden, our method focuses the object of performance assessment on selected key courses requiring the integration of knowledge and higher-order skills, which are placed at critical juncture points in curriculum (divided into four stages in the case of the Faculty of Dentistry at Niigata University), thereby insuring assessment validity as well as assessment feasibility and compatibility with a credit system. Furthermore, reliability, especially inter-rater reliability, is guaranteed through the collaboration of a faculty team, consisting of numerous teachers in charge of the courses, that develops rubrics and implements assessment including calibration and moderation.³⁵ In this way, by arranging the key courses directly linked to program-level goals sequentially within the curriculum, we combine performance assessments conducted by a faculty team at key

³⁴ Banta and Palomba, *Assessment Essentials*, 103-05.

³⁵ Suzanne Lane, "Performance Assessment: The State of the Art," in *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*, eds. Linda Darling-Hammond and Frank Adamson (San Francisco: Jossey-Bass, 2014), chap. 5, Kindle; Terrel Rhodes and Ashley Finley, *Using the VALUE Rubrics for Improvement of Learning and Authentic Assessment* (Washington D.C.: Association of American Colleges and Universities, 2013), 17-25.

courses with assessment of knowledge and skills conducted by individual teachers at other courses. This is the method of program-level assessment presented in our research.

As described above, Japanese dental education programs require students to pass standardized common achievement test and clinical examination (CBT and OSCE) before they can start their clinical practicum (Figure 4) and the National Board Dental Examination upon graduation, so that benchmarking can be conducted along these external standards. However, as CBT and the national board examination represent an objective test of individual knowledge and thinking skills, it is difficult to assess the knowledge integration and higher-order skills. OSCE assesses skills and attitudes in the simulation settings, but it cannot assess the competences that are expected to be cultivated by clinical practicum.

Moving from “University study skills,” “PBL,” “Model practice & simulation training” to “Clinical practicum,” the performance assessments in this research are designed to evaluate generic problem-solving skills as well as subject-specific problem-solving skills represented in this case by dental clinical competences. In accordance with this process, the scope of cultivation and assessment of problem-solving skills becomes increasingly specialized as well as comprehensive, while the assessment setting is designed to embody higher state of authenticity (see Table 3 and Figure 5). The results of student questionnaire after the last performance assessment in “Clinical practicum” showed numerous positive responses, such as “I could understand my clinical competences more objectively,” “This assessment told me what should be improved,” “CBT and OSCE cannot replace this assessment.”³⁶

As shown in Table 4 under the “Clinical practicum,” out of the three levels (well done, acceptable, unacceptable), students can pass only if they achieve the level of “acceptable,” and the level of “well done” indicates the standards to be met in postgraduate clinical training they experience as interns. This is how the rubric attempts to capture a learning progression up to clinical training after graduation. This also sends a message to students that their own learning should become a life-long continuous journey.

V. Conclusion

It is difficult to implement program-level learning outcomes assessments even after you are equipped with good course-level assessments. In this

³⁶ Fujii et al., “Competency Assessments,” 7-8.

paper, building on the performance assessments we developed in some courses, we proposed Pivotal Embedded Performance Assessment (PEPA) as a method for combining assessment at the course and program levels. PEPA denotes that the method adopts the idea of embedded assessment and embodies it in the form of performance assessment in key courses of the curriculum. It consists of the following set of procedures:

- (1) Faculty systemize curriculum, clarify the relation between program goals and each course, and segment curriculum into several parts. Faculty select key courses directly linked with program goals, in which students are required to integrate the various knowledge included in each segment and to cultivate higher-order skills.
- (2) The faculty team in charge of each key course designs tasks and rubrics of the performance assessment, and implements it. The newly developed different types of performance assessment are sequentially arranged so that their specificity, comprehensiveness, and authenticity increase as student learning progresses. Conversely, the assessment of other individual courses is entrusted to the expert judgment of individual teachers.
- (3) In performance assessment at key courses, students have to pass all the courses by demonstrating performance that surpasses a certain level in all rubric dimensions. However, each course's performance assessment incorporates the function of formative assessment, by giving students a second opportunity to achieve the passing criteria in the second round of key course with the same name. Also, each rubric not only applies to the course itself, but can be used longitudinally after the course's completion, indicating the direction of learning progressions.
- (4) By fulfilling passing criteria in performance assessment at key courses and obtaining a specified number of credits from a variety of stated regular courses, students are awarded the degree of the program.

Our method called Pivotal Embedded Performance Assessment enables to maintain assessment feasibility and compatibility with a credit system while insuring assessment validity and reliability, through limiting the range of performance assessment to key courses that require knowledge integration and higher-order skills and consequently are placed at the critical juncture points of curriculum, and through designing and implementing the assessment by a faculty team. Although the idea of curriculum-embedded performance

assessment is not a new way of thinking,³⁷ our PEPA is original in the way that it clarifies what key courses are selected for embedded assessments through alignment with curriculum design, while upholding a formative assessment function and sequentially arranging those embedded assessments.

In this paper, we have elucidated our method taking up the case of dental education program in Japan. We believe that PEPA is a powerful method for combining course- and program-level outcomes assessments. However, it is not clear whether the concept and procedures of PEPA will function effectively in other academic fields. We suppose that it can be utilized not only in dental education but also in medical and pharmaceutical education, which are similar to each other in their curriculum structure. Furthermore, it could be applicable in fields such as education for teachers and legal professions, which includes the progressions in cognition and behavior from Knows, Knows How, Shows How to Does, as depicted in Miller's Pyramid. Our challenge is to explore the ranges of applicability of PEPA and its potential constraints in further research.

Bibliography

- Akiba, Nami, Masako Nagasawa, Kazuhiro Ono, Takeyasu Maeda, and Katsumi Uoshima. "An Introduction to the Undergraduate Comprehensive Model Practice Course at the Faculty of Dentistry, Niigata University." *The Journal of Japanese Dental Education Association* 33 (2017): 106-14. [In Japanese.] Alverno College Faculty. *Student Assessment-as-Learning at Alverno College*. Milwaukee: Alverno College Institute, 1994.
- Banta, Trudy W., and Catherine A. Palomba. *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education*. 2nd ed. San Francisco: Jossey-Bass, 2015.
- Barrows, Howard S. "The Essentials of Problem-Based Learning." *Journal of Dental Education* 62 (1998): 630-33.
- Biggs, John, and Catherine Tang. *Teaching for Quality Learning at University*. 4th ed. Berkshire: The Society for Research into Higher Education & Open University Press, 2011. Kindle.
- Blake, Jennifer M., Geoffrey R. Norman, and E. Kinsey M. Smith. "Report Card from McMaster: Student Evaluation at a Problem-Based Medical School." *The Lancet* 345 (1995): 899-902.
- Cummings, Rhoda, Cleborne D. Maddux, and Aaron Richmond. "Curriculum-Embedded Performance Assessment in Higher Education: Maximum Efficiency

³⁷ Cummings, Maddux, and Richmond, "Curriculum-Embedded Performance Assessment," 599-605.

- and Minimum Disruption.” *Assessment & Evaluation in Higher Education* 33, no. 6 (2008): 599-605.
- Earl, Lorna M. *Assessment as Learning: Using Classroom Assessment to Maximize Student Learning*. Thousand Oaks: Corwin Press, 2003.
- Ewell, Peter, Karen Paulson, and Jillian Kinzie. *Down and In: Assessment Practices at the Program Level*. Champaign: National Institute for Learning Outcomes Assessment, 2014. <http://www.learningoutcomeassessment.org/documents/NILOAsurveyreport2011%20-%20Down%20and%20In%2010-20.pdf>.
- Fujii, Noritaka, Syoji Takenaka, Koichi Tabeta, Naoko Sato, Nami Akiba, Yohei Oda, Yuji Katsumi, Kazuhiro Ono, and Takeyasu Maeda. “Competency Assessments for Undergraduate Students in Clinical Clerkships at the Faculty of Dentistry, Niigata University.” *The Journal of Japanese Dental Education Association* 33 (2017): 4-11. [In Japanese.]
- Kuh, George D., Natasha Jankowski, Stanley O. Ikenberry, and Jillian Kinzie. *Knowing What Students Know and Can Do: The Current State of Student Learning Outcomes Assessment in U.S. Colleges and Universities*. Champaign: National Institute for Learning Outcomes Assessment, 2014. <http://www.learningoutcomeassessment.org/documents/2013%20Abridged%20Survey%20Report%20Final.pdf>.
- Lane, Suzanne. “Performance Assessment: The State of the Art.” In *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*, edited by Linda Darling-Hammond and Frank Adamson, chap. 5. San Francisco: Jossey-Bass, 2014. Kindle.
- Matsushita, Kayo. “Making Learning Outcomes Visible.” *Japanese Journal of Higher Education Research* 20 (2017): 93-112. [In Japanese.]
- Matsushita, Kayo, Kazuhiro Ono, and Yusuke Takahashi. “Development of a Rubric for Writing Assessment and Examination of Its Reliability.” *Journal of the Liberal and General Education Society of Japan* 35, no. 1 (2013): 107-15. [In Japanese.]
- Middle States Commission on Higher Education. *Student Learning Assessment: Options and Resources*. 2nd ed. 2007. https://www.msche.org/publications/SLA_Book_0808080728085320.pdf.
- Miller, George E. “The Assessment of Clinical Skills/Competence/Performance.” *Academic Medicine* 65, no. 9 (September Supplement 1990): S63-S67.
- Ministry of Education, Culture, Sports, Science and Technology. “Japanese University Reforms Including Those in Educational Contents.” November, 2017. http://www.mext.go.jp/a_menu/koutou/daigaku/04052801/___icsFiles/afieldfile/2017/12/13/1398426_1.pdf. [In Japanese.]
- Model Core Curriculum Revision Coordination Committee and Model Core Curriculum Revision Specialist Research Committee. *Model Core Curriculum for Dental Education: AY 2016 Revision*. 2017. [In Japanese.] http://www.mext.go.jp/component/b_menu/shingi/toushin/___icsFiles/afieldfile/2017/12/26/1383961_02_3.pdf.
- Mtshali, Ntombifikile G., and Lyn Middleton. “The Triple Jump Assessment: Aligning Learning and Assessment.” In *New Approaches to Problem-Based*

- Learning: Revitalising Your Practice in Higher Education*, edited by Terry Barrett and Sarah Moore, 187-200. New York: Routledge, 2011.
- Newman, Mark J. "Problem Based Learning: An Introduction and Overview of the Key Features of the Approach." *Journal of Veterinary Medical Education* 32 (2005): 12-20.
- Oda, Yohei, Kazuhiro Ono, Noritaka Fujii, Tadaharu Kobayashi, and Takeyasu Maeda. "Development and Use of a Web-Based E-Portfolio for Dental Clinical Training." *The Journal of Japanese Dental Education Association* 33 (2017): 65-73. [In Japanese.]
- Ono, Kazuhiro, Kayo Matsushita, and Yugo Saito. "Prospects for Direct Assessment of Problem Solving Competence: Development of Modified Triple Jump in Problem-Based Learning." *Journal of the Liberal and General Education Society of Japan* 36, no. 1 (2014): 123-32. [In Japanese.]
- Ono, Kazuhiro, and Kayo Matsushita. "Assessment of Writing in First-Year Education." In *Assessment of Active Learning*, edited by Kayo Matsushita and Terumasa Ishii, 26-43. Tokyo: Toshindo, 2016. [In Japanese.]
- . "PBL Tutorial Linking Classroom to Practice: Focusing on Assessment as Learning." In *Deep Active Learning: Toward Greater Depth in University Education*, edited by Kayo Matsushita, 183-206. Singapore: Springer, 2017.
- Pike, Gary R. "Limitations of Using Students' Self-Reports of Academic Development as Proxies for Traditional Achievement Measures." *Research in Higher Education* 37, no. 1 (1996): 89-114.
- Rhodes, Terrel. *Assessing Outcomes and Improving Achievement: Tips and Tools for Using the Rubrics*. Washington D.C.: Association of American Colleges and Universities, 2009.
- Rhodes, Terrel, and Ashley Finley. *Using the VALUE Rubrics for Improvement of Learning and Authentic Assessment*. Washington D.C.: Association of American Colleges and Universities, 2013.
- Richman, W. Allen, and Laura Ariovich. *All-in-One: Combining Grading, Course, Program, and General Education Outcomes Assessment*. Champaign: National Institute for Learning Outcomes Assessment, 2013. <http://learningoutcomesassessment.org/documents/Occasional%20Paper%2019%20FINAL.pdf>.
- Rohlin, Madeleine, Kerstin Petersson, and Gunnel Svensäter. "The Malmö Model: A Problem-Based Learning Curriculum in Undergraduate Dental Education." *European Journal of Dental Education* 2 (1998): 103-14.
- Saito, Yugo, Kazuhiro Ono, and Kayo Matsushita. "Correlations of Direct Measures Based on Performance Assessment and Indirect Measures Based on Student Self-report." *Japan Journal of Educational Technology* 40 (Suppl.) (2016): 157-60. [In Japanese.]
- Suskie, Linda. *Assessing Student Learning: A Common Sense Guide*. 2nd ed. San Francisco: Jossey-Bass, 2009. Kindle.
- Winning, Tracy, Elaine Lim, and Grant Townsend. "Student Experiences of Assessment in Two Problem-Based Dental Curricula: Adelaide and Dublin." *Assessment & Evaluation in Higher Education* 30, no. 5 (2005): 489-505.

About the authors

KAYO MATSUSHITA (matsushita.kayo.7r@kyoto-u.ac.jp), corresponding author, has been a professor at the Center for the Promotion of Excellence in Higher Education and the Graduate School of Education, Kyoto University, Japan, since 2004. She received her B.A., M.A., and Ph.D. in education from Kyoto University, Japan. Her research themes include teaching & learning, curriculum, and assessment in higher education. She advocates “deep active learning,” which combines active learning and deep learning, in her edited book *Deep Active Learning: Toward Depth in University Education* (Springer, 2017). Classifying learning outcomes assessment into four types along two axes, direct vs. indirect and qualitative vs. quantitative, she mainly focuses on performance assessment. She has developed several performance assessments collaborating with faculty members in fields such as dentistry, physical therapy, and philosophy. She is currently leading a project “Building Disciplinary Reference Points for Curriculum Design and Quality Assurance of University Education” in the field of education studies, as a member of the Science Council of Japan. She is also a council member of the Tuning Japan National Center.

KAZUHIRO ONO (k-ono@dent.niigata-u.ac.jp) received D.D.S., Ph.D. from Niigata University, Japan in 1990. He is a professor at the Division of Oral Science for Health Promotion and the Division of Dental Educational Research Development (concurrent post), Graduate School of Medical and Dental Sciences and a chair of the Student Affairs of the Faculty of Dentistry, Niigata University. He specializes in oral surgery and dental education. He has been leading curriculum and assessment reforms at the Niigata University Faculty of Dentistry, and is one of the influential members in charge of educational programs of both the Department of Dentistry and the Department of Oral Health and Welfare which comprise the Faculty of Dentistry. His recent research topics include active learning, especially problem-based learning, performance assessment of higher-order integrated abilities, and program design. Having been a vice president of Niigata University since 2018, he is also in charge of establishing a system of quality assurance for higher education at Niigata University.

YUGO SAITO (ugo.saito@gmail.com, y-saito@pt-u.aino.ac.jp) received B.A., M.A., and Ph.D. in education from Kyoto University, Japan. He is an assistant professor at the Department of Physical Therapy, Faculty of Health Science, Aino University, Japan. He is a member of the Japanese follow-up program “The Assurance of Higher Education through the Development of a Tuning Test Item Bank Global Quality.” He works with a team of engineers seeking to develop a shared understanding of expected learning outcomes in the field of mechanical engineering. His research topics include assessment of higher education learning outcomes, performance assessment, institutional research, learning analytics, and deep active learning. The focus of his research is on how to bridge and combine direct and indirect as well as quantitative and

qualitative assessments to support students' learning and development. In addition, he teaches statistics and information science to paramedical students using "deep active learning" to increase students' research ability and higher-order thinking skills.

Combining course- and program-level outcomes assessments through embedded performance assessments at key courses: A proposal based on the experience from a Japanese dental education program

Kayo Matsushita, Kazuhiro Ono, and Yugo Saito

doi: [http://dx.doi.org/10.18543/tjhe-6\(1\)-2018pp111-142](http://dx.doi.org/10.18543/tjhe-6(1)-2018pp111-142)

Received: 27.09.2018

Accepted: 13.11.2018

Copyright

Copyright for this article is retained by the Publisher. It is an Open Access material that is free for full online access, download, storage, distribution, and or reuse in any medium only for non-commercial purposes and in compliance with any applicable copyright legislation, without prior permission from the Publisher or the author(s). In any case, proper acknowledgement of the original publication source must be made and any changes to the original work must be indicated clearly and in a manner that does not suggest the author's and or Publisher's endorsement whatsoever. Any other use of its content in any medium or format, now known or developed in the future, requires prior written permission of the copyright holder.