

Classroom-embedded assessment based on subject differences for  
high school teachers focused on “learning to learn” behind the PISA

Masahiro Arimoto, Yoshiko Goda

Tohoku University, Kumamoto University

masaari@m.tohoku.ac.jp, ygod@kumamoto-u.ac.jp,

## ABSTRACT

The purposes of this study are: (1) to provide an overview of the use formative assessments in Japanese high school classrooms for students’ learning to learn and (2) to investigate formative assessment use based on subject matter. The survey questionnaires from 402 Japanese high school teachers were based on the Teaching and Learning Research Programme (TLRP) in the UK. The findings show the teaching practices related to assessments, including (1) providing guidance that helps the pupils assess their own learning and (2) giving the pupils opportunities to determine their own learning objectives. Teaching methods with significant differences between a charged-subject vs. the overall differ from charged-subjects. The adoption of methods that help students with peer assessment will be a challenge in the future. The study results provide significant insight that can be used to help future educational reform.

**KEYWORDS:** Classroom-embedded assessment, Japanese high school, learning to learn

## 1. Introduction

Learning to learn (L2) is a core competency for lifelong learning. According to the Programme for International Student Assessment (PISA) analysis report, formative assessment is an effective teaching tool for students’ mastery of L2 (OECD, 2005). In this study, formative assessments use by Japanese high school teachers in their classes was investigated.

Teaching methods vary and tend to reflect the subject matter being taught. For example, English as a foreign language (EFL) education uses the communicative approach in which EFL teachers use peer or group activities to increase student interactions in the classrooms (Ellis, 1994). In this study, the teaching methods based on

each teacher's subject were also investigated to identify how the methods enhance students' L2.

## 2. Background and Related Literature

L2 is defined as the capability and willingness to adapt to new tasks (Hautamäki et al., 2002). L2 concepts have been adapted to the current changing educational environment, include lifelong learning factors, and reflect a central position amid cross-curricular competencies (Hautamäki et al., 2010). L2 competence, including the knowledge, skills, attitudes, and values necessary for lifelong learning, is one of the eight key competences. The other seven are: communication in the mother tongue, communication in foreign languages, mathematical competence and basic competence in science and technology, digital competence, social and civic competence, a sense of initiative and entrepreneurship, and cultural awareness and expression (Hoskins & Fredriksson, 2008).

L2 and similar program concepts have emerged in educational curriculum in Portugal, Austria, Finland, and England, and L2 has been introduced into the national curriculum in Spain, Italy, Cyprus, and France (Hoskins & Fredriksson, 2008). In Japan, L2 is relatively new, but similar concepts had been employed in the educational curriculum and in classroom teaching. Katsuno et al. (2000) report the importance of the integration of 21<sup>st</sup> Century competencies that are similar to L2 concepts into educational curriculum beyond subject matters. Hautamäki et al. (2002) mention that L2 is formed through good educational practices and accompanies all achievement. It is essential to encourage students to acquire L2 competency across subjects and learning domains.

Hautamäki et al. (2002) state that L2 assessment is easy to execute and is cost-effective. One of the frameworks for L2 promotion in schools is Assessment for Learning (AfL) (Hautamäki et al., 2010). There were several projects related to AfL. James et al. (2006) clarifies the primary role of AfL and provides a self-evaluation questionnaire with 30 statements about classroom assessment practices. Part of James et al.'s self-evaluation questionnaire was used in the current study to assess high school teachers' L2 teaching methods. James et al.'s original questionnaire consisted of three factors: making learning explicit, promoting learning autonomy, and performance orientation.

Since the 1990s, teachers have been required to perfect their teaching methods

through interacting with other teachers in primary and junior high schools. "Experienced [teachers] assume responsibility for advising and guiding their young colleagues. Head teachers [principals] organise meetings to discuss teaching techniques...Meetings at each school are supplemented by informal district-wide study groups" (Stevenson & Stigler, 1992). The current study is founded upon the prerequisite that there is much to be learned in the study of the unique and significant cultural attribute of "group quest of knowledge," so it focuses on the classroom-embedded assessment of Japanese high school teachers. This topic does not have a lot of previous research. In the current study, Japanese teachers' daily educational practices were investigated, focusing on AfL, and different teaching methods based on subject matter of their teaching were examined.

### 3. Research Methods

#### 3.1 Research Participants

A total of 401 (300 male, 100 female) Japanese teachers working at nine public high schools participated in this survey research. Their work experience varied from less than 5 years to more than 20 years.

#### 3.2 Instrument

The questionnaire used in this research had two parts: demographic information and the type of formative assessment used in the classroom. The former section included a multiple choice question about the subject the teachers taught with 16 options (national language (Japanese), geography & social study, mathematics, science, foreign language (English), P.E., art, home economics, information, commerce, technology, fishery, farming, nursing, welfare, and others). The latter section of the questionnaire consisted of 12 statements (Table 1). All the statements were evaluated on a four-point Likert scale. Of the statements, nine were selected from the Teaching and Learning Research Programme (TLRP; James et al., 2006), one was from Kinoshita et al. (2005), and the final two statements (2 & 3) were developed by the researchers and were based on the unique features of Japanese teaching methods.

Question 2 asked for class size, typically greater than 40 students, to determine the largest classes. To provide a formative assessment of each student, a teacher should walk around the classroom and observe the individual learning progresses of each

student. Therefore, question 3 examined the guidance teachers provided for poorly structured problem solving.

### 3.3 Data Collection and Data Analysis

The questionnaire data were collected from the schools ( $n = 393$ ) and through the mail ( $n = 8$ ) in July and August of 2012. The data were analyzed using descriptive statistics for the first research purpose, to provide an overview of Japanese high school teachers' use of formative assessments in their classrooms. To investigate the differences in teaching methods based on subject, t-tests were conducted for five subjects (first language, geography, mathematics, science, and foreign language). These five subjects were chosen because they were mandatory courses in the all schools, and more than 50 teachers taught these subjects, possibly providing enough data to conduct a comparative inferential statistical analysis.

## 4. Results

### (1) Overall Tendency to Use Formative Assessments in Classrooms

The overall results for the 12 formative assessment statements are shown in Table 1. Statements 1, 2, 5, and 6 averaged higher than other items. The highest average was from statement 5 ( $m = 3.48$ ), I provide guidance to help pupils assess their own learning. The second highest was statement 1 ( $m = 3.39$ ), pupils are given opportunities to decide their own learning objectives.

Statements 8, 9, 10, and 11 had lower averages. Although, the standard deviation for statements 8 and 9 was greater than 0.8, these statements' scores were higher than other statements' scores. The lowest averaged score was from statement 9 ( $m = 2.65$ ), I provide guidance to help pupils to assess one another's work. The second lowest score was statement 10 ( $m = 2.67$ ), pupils are given opportunities to assess one another's work.

### (2) Subject Differences of Teaching Methods

The demographic section of the questionnaire was used to categorize the teachers' subject groups. The five subject groups (first language, geography, mathematics, science, and foreign language) were compared with the averages of the other subjects.

The results of the t-tests are shown in Table 2. The mean for the number of teachers who teach First Language is significantly higher than the other subjects on statements 3, 7, and 9. Statements 2, 5, 6, and 10 concerned Foreign Language teachers. A significant positive t-score was found in mathematics for statement 2. Negative significances in geography and social studies were found for statements 2, 5, 9, and 10. Those in science were statements 2, 3, and 4. Statement 1 in the foreign language classes was inversely related to the other subjects, and it was also significant.

Table 1. Questionnaire Statements, Sources, and Descriptive Analysis Results

Source	item #	Question	<i>n</i>	<i>m</i>	<i>sd</i>
James, et al. (2006)	1	Pupils are given opportunities to decide their own learning objectives.	401	3.39	0.67
Researchers-made	2	I identify pupils' understanding through walking around the class and checking how they are doing.	401	3.29	0.73
Researchers-made	3	I provide guidance to help pupils solve ill-structured problems.	399	2.91	0.79
James, et al. (2006)	4	I identify pupils' strengths and advise them on how to develop them further.	401	2.91	0.72
James, et al. (2006)	5	I provide guidance to help pupils assess their own learning.	401	3.48	0.57
James, et al. (2006)	6	Pupils are encouraged to view mistakes as valuable learning opportunities.	401	3.31	0.72
James, et al. (2006)	7	I use questions mainly to elicit reasons and explanations from my pupils.	401	3.09	0.72
James, et al. (2006)	8	I provide guidance to help pupils assess their own work.	401	2.80	0.75
James, et al. (2006)	9	I provide guidance to help pupils to assess one another's work.	401	2.65	0.80
James, et al. (2006)	10	Pupils are given opportunities to assess one another's work.	401	2.67	0.84
Kinoshita, et al. (2005)	11	Pupils are told how well they have done in relation to their own precious performance.	401	2.76	0.75
James, et al. (2006)	12	I use questions mainly to elicit factual knowledge from my pupils.	401	3.19	0.67

Table 2. The Subject Differences of Formative Assessment Use in Daily Classes

Subject	n	Question #											
		1	2	3	4	5	6	7	8	9	10	11	12
First Language (Japanese)	63			++				+++		++			
Geography & Social Study	51		---			-				-		---	
Mathematics	70		++							--		---	
Science	55		--	-	-								
Foreign Language (English)	87	-	+				++	+++				+++	

*Note.* The mean of the subject that the teacher is responsible for teaching is significantly higher at +++  $p < 0.001$ , ++  $p < 0.01$ , +  $p < 0.05$ . The mean of the subject the teacher is responsible for teaching is significantly lower at ---  $p < 0.001$ , --  $p < 0.01$ , -  $p < 0.05$ .

The detailed results of the t-tests for statements 2, 9, and 10 are organized in Table 3. Statement 2 had a higher overall average (Table 1), but science was significantly negative ( $t = -3.24$ ,  $p < 0.05$ ). Mathematics and geography and social studies had a negative significance for statements 9 and 10. English had a positive significance for statement 10.

Table 3. Results of the t-test for Statements 2, 9, and 10

Subject	Question#2				Question#9				Question#10			
	Charged		Other		Charged		Other		Charged		Other	
	Subject's mean	Subjects' mean	t	p	Subject's mean	Subjects' mean	t	p	Subject's mean	Subjects' mean	t	p
First Language (Japanese)					2.95	2.59	3.50	**				
Geography & Social Study	2.96	3.34	3.52	**	2.41	2.68	-2.24	*	2.29	2.73	-3.51	**
Mathematics	3.51	3.25	2.79	*	2.39	2.70	-3.02	*	2.36	2.74	-3.53	**
Science	3.00	3.34	-3.24	*								
Foreign Language (English)	3.46	3.25	2.39	*					3.05	2.57	4.82	**

*Note.* \*\*:  $p < 0.01$ , \*:  $p < 0.05$ .

## 5. Discussion and Future Implications

Overall, the results showed that teachers most often guided students' self-evaluations (#5) and self-directed learning (#1), and peer-assessment (#9, #10) was used the least in the classroom. This indicates teachers may be challenged to provide and encourage collaborative work, including peer-assessments, in their classes. Both

first and foreign language teachers used AfL methods significantly more often than the teachers of other subjects. Geography and social studies and science teachers employed AfL less often than other subjects' teachers. Geography and social studies and science teachers do not appear to provide peer-assessment activities. To enhance peer reviews based on the results, these teachers should be encouraged to introduce collaborative work in their classes. However, some subjects' unique features might hinder the use of group activities. In this case, to encourage teachers' adoption of group activities and provide peer assessment opportunities, the learning environment should provide opportunities for students to collaborate.

In the future, learning environments that encourage teachers to use AfL effectively should be designed and established. The integration of new tools in the learning environment could lead to a rich AfL environment for L2. For example, a pilot study using a tablet-type iPad in the classroom and school designated by MEXT (2013-15) has been just started. The research results may provide useful information for future educational reforms.

#### **Note**

The main part of this study was conducted by Mr. Kazumasa Ikeda, Sendai Ni High School, as a part of a master's thesis submitted to the Graduate School of Education at Tohoku University. This article was organized with his permission.

#### **Acknowledgement**

This work was supported by JSPS KAKENHI Grant Numbers Scientific Research (B) 22330240, (A) 25245079.

#### **References:**

- Ellis, R. (Ed.) (1994). *The Study of Second Language Acquisition*. Oxford, U.K.: Oxford University Press.
- Hautamäki, J., Arinen, P., Eronen, S., Hautamäki, A., Kupianen, S., Lindblom, B., Niemivirta, M., Pakaslahti, L., Rantanen, P., & Scheinin, P. (2002). *Assessing Learning-to-Learn: A Framework*. Helsinki: Centre for Educational Assessment, Helsinki University/National Board of Education.
- Hoskins, B. & Fredriksson, U. (2008). *Learning to Learn: What is It and Can It Be*

*Measured?* (JRC Scientific and Technical Reports EUR 23432 EN 2008). Ispra: Joint Research Centre of European Commission.

Hautamäki A., Hautamäki J., & Kupiainen S. (2010). Assessment in Schools— Learning to Learn. In P. Peterson, E. & Baker, B. McGaw, (Eds.), *International Encyclopedia of Education* (Vol. 3, pp. 268–272). Oxford: Elsevier.

James, M., Black, P., Carmichael, P., Conner, C., Dulley, P., Fox, A., Frost, D., Honour, L., MacBeaath, J., McCormick, R., Marshall, B., Pedder, D., Procter, R., Swaffield, S., & Wilaim, D. (2006). *Learning How to Learn: Tools for Schools*. London, U.K.: Routledge.

Kinoshita, H., Matsuura, T., & Kadoya, S. (2005). A fundamental study on Teacher's awareness and situation for meta cognition—Through questionnaire survey on the teaching of science. *Japan Curriculum Research and Development Association Journal*, 28(3), 83–89.

OECD. (2005). *Formative Assessment Improving Learning in Secondary Classrooms*. OECD.

Stevenson, H., & Stigler, J. W. (1994). *The Learning Gap: Why Our Schools are Failing and What We Can Learn from Japanese and Chinese Education*. New York: Simon & Schuster.