高等学校英語授業に代替的生徒評価の導入を

Introducing Alternative Student Assessment
in Senior High School English Classes

市島清貴

Kiyotaka ICHISHIMA

要旨

　高等学校の英語授業は、学習指導要領が改訂される度にますますコミュニケーションが重視されるようになってきたのだが、生徒の評価方法に大きな変更は見られない。授業では発表力や表現力などのコミュニケーション能力を高める目的の活動が取り入れられることが多くなったが、活動ごとの評価を授業の内外で行うことはまれである。現場では旧態依然とした読み・書き中心のペーパー試験形式である定期考査の評価が絶対である。つまり授業内容がコミュニカティブになっても、生徒の個々の活動の評価が十分に行われているとは言えない。しかし、生徒の学習意欲を増大させ、積極的に授業に参加させ、日々の授業がより活発になるためには、ポートフォリオやルーブリックを用いた新たな評価法を活用すべきである。こういった代替的評価方法を採択することは、発音やスピーキング、プレゼンテーション能力を含む総合的な英語運用能力向上のために重要な意味を持つこととなる。

1. Introduction

　　　The Japanese Ministry of Education has published the New Course of Study for Senior High Schools that will be implemented in the year 2013. All senior high schools are now changing or modifying their curricula in hopes of coping with the new contents. The required course in English aims to nurture an ability to communicate in English by "comprehensively" utilizing listening, reading, speaking and writing skills. High school students will learn 1,800 English words instead of the current 1,300 words. Thus, in total, students are expected to learn 3,000 English words during their middle and high school years, 800 more words than the current level. In principle, spoken English is supposed to be used to teach English at high school (The Japan Times Online, 2009).

　　　Whereas the New Course of Study increases the content and hours to be taught, once again, the Ministry of Education (MEXT) has provided no useful

guidelines for the assessment of student learning. The Course of Study merely encourages teachers to foster a much broader scope of student competency, including problem-solving, critical thinking and group work abilities. However, the schools are left on their own in this regard.

> The Courses of Study seek to foster the qualities and abilities necessary to acquire steadily the rudimentary basics of education…and to learn, think and act for oneself as well as develop problem-solving skills. MEXT is working to ensure that children can actively engage in educational activities that offer individual instruction, review instruction, and hands-on, problem-solving learning, and making other improvements including the creation of the Periods for Integrated Study and the expansion of elective learning (MEXT, 2010).

However, from his experience as a former high school teacher of English, and as a parent of a student who receives report cards from a senior high school in Japan, the author acknowledges that student assessment highly depends on the results of regular mid-term and term-end examinations and focuses very little on student performance during daily classroom activities. In the United States of America, alternative assessments are commonly employed in public schools, however in Japan teachers are reluctant to move away from the traditional test-based method. Therefore, this paper will suggest that Japan consider using the alternative methods of assessment employed in American schools and explain how they might be adapted for use by Japanese teachers.

 (It should be noted that in this paper the terms "assessment" and "evaluation" are used interchangeably.)

2. Purpose of This Study

This paper focuses on student performance assessment in English class settings at Japanese senior high schools. Further, it will note that current assessment methods are inadequate and are not practiced properly. Accordingly, the purpose of this paper is to share with high school colleagues some practical teaching strategies that endeavor to introduce evaluation principles for enhancing students' motivation in English classes. One major variable that baffles Japanese English teachers is that there are no standard criteria to evaluate students' everyday performance. The author has retrospectively found that teachers tend to

use only the easiest methods of evaluation, and find it difficult to keep their class assessments in step with other teachers'. Teachers of English rarely have opportunities to discuss the effectiveness or significance of their schools' student assessment methods.

3. Significance of Student Assessment

One of a teacher's primary practices in any school setting is assessment. In this endeavor, a wide range of methods should be used; however, they have to meet the instructional needs of teachers and the learning needs of students. All teachers of English are required to teach four main skills, namely writing, reading, listening, and speaking, in a well-balanced manner. However, the methods used to assess these skills influence students' approaches to learning them.

Assessment is best described as a process by which information is used by teachers to adjust their teaching strategies and by students to adjust their learning strategies. Assessment, teaching and learning are inextricably linked, as each informs the other (New Zealand Government, 2009). Assessing student performance both in and out of class is not only required under the Ministry's guidelines, if handled properly, it can be a great factor toward developing student motivation in the classes in terms of "backwash effects" (Heaton, 1975, p. 162).

Assessment is important in that it "can help create positive attitudes toward instruction by giving students a sense of accomplishment and a feeling that the teacher's evaluation of them matches what he has taught them. Good English tests also help students learn the language by requiring them to study hard, emphasizing course objectives, and showing them where they need to improve" (Madsen, 1984, p. 5). Further, assessment can also be used to determine learner competence in several domains, including: cognitive - knowledge, analysis, communication; critical thought; meta-cognitive—planning, monitoring; self-regulation; social—leading discussions, persuading; working in groups, and affective—perseverance, internal motivation, responsibility, and independence (University of Warwick, 2009).

4. Formative and Summative Assessment

Formative assessment is an ongoing classroom process by which teachers can measure students' progress for their instructional goals. Hlebowitsh (2005) states that "the use of structured teacher observations can be part of a formative

evaluation, an ongoing and periodically employed method that reminds the teacher of the importance of some of the purposes in the school curriculum" (p. 192). Upon collecting students' data, teachers can judge whether their instructions were appropriate or in need of some alterations.

Data can be gained through such activities as tests, class performances, portfolios, and checklists. When used as an ongoing practice, it can help teachers better understand to what extent students acquire lesson content, skills, and knowledge. Formative assessment also gives students direction for further guided and self-learning. According to the Organization for Economic and Cooperation and Development (2005), "Formative assessment builds students' 'learning to learn' skills by emphasizing the process of teaching and learning, and involving students as partners in that process. It also builds students' skills at peer-assessment and self-assessment, and helps them develop a range of effective learning strategies" (OECD, 2005).

The summative evaluation literally means to summarize student progress. Summative assessments come in the form of achievement tests and curriculum-based assessments. This data can be gained through tests, class performance, portfolios, checklists, and other materials of this type. A report card is commonly used to record student progress during a pre-determined period. The data for a summative assessment can be used to help determine instructional content for the upcoming school year. Test results also can offer more specific information regarding the content of subjects. "Summative assessments are an efficient way to identify students' skills at key transition points, such as entry into the world of work or for further education" (OECD, 2005). Another benefit is that summative evaluation can be used to show accountability to the parents, local community, and administrators.

Teachers need not strictly differentiate between formative or summative evaluation in terms of evaluating students; both should also incorporate a remedial or diagnostic element. This can help determine remedial courses of action beneficial to students and provide essential information that will help teachers when designing instruction or curriculum.

Any type of assessment should help learners immediately recognize where they have a weakness and give them an opportunity to correct any mistakes. Learning occurs when students make a mistake and correct it. Thus, teachers

should not miss an important opportunity to promptly provide students with feedback in order to enhance their learning.

5. Over-reliance on Summative Assessment

Summative assessments are often implemented at the very end of a course, term, semester, or school year. In Japan, most schools have three terms in a year, during which, five high-stakes examinations are administered. During the first term, which starts in April, a school conducts mid-term examinations in May and term-end examinations in early July, after which, the first term ends. After a one-month summer vacation, the second term starts in late August, and schools, following the same pattern, conduct two more high-stakes examinations before the term ends in late December. After a two-week winter break, the third term starts in January and, as it is shorter than the other two, schools conduct only one high-stakes evaluation before the school year ends.

The final grade that senior high school students receive at the end of the school year in March is calculated using a five to one scale and is usually based on the average score of the five high-stakes examinations. At most, everyday class performance can be allotted 10% of the final grade. Traditionally in Japan's senior high school settings, however, mid-term and term-end examinations are the two high-stakes examinations that most decide a student's term-end grade.

Japan's senior high schools have had a long tradition of overreliance on such examinations and of ignoring everyday class performance. Pronunciation, public speaking and other performance skills are the very last to be tested perhaps because Japan is a country where students traditionally remain silent during their classes unless told to speak. Thus, students study only what will be tested, namely grammar, vocabulary, writing, and reading comprehension.

However, teachers should constantly collect informal and formal information about students' learning. This information can be gleaned from paper tests and assignments, observation of group activities, and other academic performances and activities. Dornyei (2005) states that teachers should take the students' learning very seriously and show students that their teachers care about their progress. Without properly considering this information, students will be unable to enhance their abilities or find enough motivation to study hard in class.

6. Suggestions for Student Assessment

Assessment, especially in-class evaluation, is an area that teachers of English at Japanese senior high schools have mostly ignored. Traditionally mid-term and end-of-term examinations have been of vital importance to teachers and students. Therefore, the students' everyday performance or attitude is believed to be, at most, negligible and, at worst, simply too hard to assess during class.

However, the author suggests that teachers should make classes more communicative and learner-centered and that they could enhance the attitude and the level of participation of the students with a simple change to the assessment system. To elevate its importance, in-class evaluation should be added to the regular high-stakes paper exams. Teachers should test pronunciation, and evaluate delivery and other performance skills both during and outside class.

Though often ignored by Japanese teachers of English, the Ministry of Education's guidelines do require high school teachers to "foster a positive attitude toward communication" (MEXT, 2009). Thus, the teachers of English are required to reconsider their assessment system even though classroom settings remain unchanged (e.g. an average class size of 40 students). Several examples of general assessment principles are shown below. It is imperative that these be taken into consideration when teaching English communicatively; to that end, this study introduces the following characteristics of assessment that, from now, should be used as general principles:

(a) Direct versus Indirect Testing

Conventionally, teachers tend to use non-communicative tests, such as cloze, structure, and dictation, on the grounds that they can be easily tested; however, more emphasis should be placed on direct testing in communicative classes so that the testing procedures aim to assess directly what has been taught. This is supported by Nunan (1988), who says "as assessment procedures inevitably seem to have a major influence on what happens in the classroom, it is desirable for these procedures to mirror what curriculum designers feel should be going on" (p. 118).

(b) Objectivity versus Encouragement

"The conventional way of evaluation is largely based on the concept of the relative evaluation, which means the teacher is supposed to grade the students from A to D

in the scale even if they obtain fairly good scores in examinations" (Yoshida, 1993, p. 39). This has been helpful in screening the students or letting them know their positions in a group in terms of academic achievement." However, when an emphasis is placed on "objectivity" and "discrimination," this does not seem to foster a positive participation attitude. Evaluations should function as a means of encouraging the students to study harder.

(c) Norm-referenced versus Criterion-referenced Testing

A student's performance, for example, is evaluated based on the relative extent to which he or she is able to attain the goals of a task. He or she is not evaluated in comparison with the achievements of other students; therefore, on one test, certain students might receive an 'A', while on another test, a 'D'. "Criterion-referenced tests are designed to enable the test user to interpret a test score with reference to a criterion level of ability or domain of content" (Bachman, 1990, p. 74). A primary aim of this testing, then, is to encourage students to attain higher goals, rather than discriminate against them based on their specific skills.

In criterion-referenced assessments, it is possible that none, or all, of the examinees will reach a particular goal or performance standard (Berman, 2007, p. 101). Therefore, reporting forms should show students and parents how well students meet specific goals rather than just telling them how their performance compares to a normal group of national or local students.

(d) Diagnosis and Feedback

A teacher has to give a general "grade" which shows the ranking of particular students, for example, "the third" in a class of 40 students, or "5" on the traditional grading scale of 5 to 1. However, this does not tell either the teacher or the students very much unless they know exactly what the grade is based on. A good listener may receive bad points on a test, if it focuses on spelling or grammar. "Diagnostic testing is of value in that it provides critical information to the student, teacher, and administrator that should make the learning process more efficient" (Grant, 1987, p. 2). It will be important to identify what sub-skills or areas a test is measuring.

(e) Teacher Collaboration

Collaboration among teachers and school administrators will ensure a successful assessment system. Therefore, creating a collective expectation among teachers regarding student performance is important. Teachers who collaborate by discussing issues related to student learning are more likely to take advantage of internally and externally generated information. They may be more willing recipients of research information if they are embedded in a setting where meaningful and sustained interaction with researchers occurs in an egalitarian context (Blankstein, Houston, & Cole, 2009).

Teachers, whatever they may teach, are encouraged and obligated to offer creative, interesting, and fun lessons to their students. However, when considering how to assess the students, teachers must use agreed-upon grading criteria and other assessing methods. The grades students receive have an impact on their future career courses, including their choice of upper-level schools and occupations. The assessment systems, therefore, must be reliable, valid, credible, objective, and convincing.

(f) Group Evaluation

Languages should be learned through interaction with teachers and classmates because this promotes two-way communication. In any society, good relationships and cooperation are necessary and require competent language and social skills. One way to encourage this is to have students provide critical assessments of friends' performances. This is known as "peer evaluation," and it helps the students maintain or enhance the quality of their work or performance through feedback from friends.

7. Alternative Assessment

Traditional forms of assessment are relatively easy to conduct. Teachers can easily create and grade multiple-choice, matching, and true or false tests. However, considering the general principles of assessment mentioned above, conventional grading systems will not be conducive to enhancing class activities, or fostering communicative competence. Alternative assessments, on the other hand, have rapidly gained popularity, particularly in America, but also among up-and-coming Japanese teachers, because of a need for a systemic strategy that improves student outcomes.

Alternative types of assessments include written essays, hands-on performance tasks and cumulative portfolios of diverse work products. In a nutshell, the biggest difference is that machines can do traditional assessments; however, only educators can perform alternative assessments. One of the advantages of alternative assessments is that they allow teachers to understand their students better. For example, true or false testing will not clarify the strengths and weaknesses of the students' writing skills, whereas portfolio and rubrics assessments, illustrated below, are designed to do that.

(a) Portfolio-Based Assessment

Portfolio-type assessments allow teachers to monitor student progress in such activities as writing by systematically collecting students' works. They encourage teachers to regularly monitor each individual student's performance and progress. The approaches to portfolios vary; however, they all rest on records kept by the teacher and on collections of the student's work called the "student portfolio." Portfolios also encourage students to reflect on the work that has been selected. Such reflection helps students think not only about what they have learned, but about their own learning processes, all of which contributes to the overall goal of improving student learning (NewSchools, 1999).

In some approaches, the teacher examines the portfolio and evaluates the work based on a scoring guide at the end of a marking period. Sometimes students or their peers also score their work. This approach is useful for the teacher and parent in determining how well a student is progressing. This classroom-based approach has several advantages. For example, the evaluation is based on a wide range of student work done over a long period of time rather than on a single, paper-and-pencil test taken over a few hours (NewSchools, 1999). Further, since a student portfolio is an ongoing process used throughout the year, this type of assessment should show incremental individual improvements.

The author used a portfolio-based assignment and assessment in his English technical writing class when he was a research student at a graduate school:

> Students were asked to write about such topics as "environmental issues,"
> "international relations," and "English pedagogy." After they submitted
> the first draft to the instructor, he corrected mistakes and provided
> comments and evaluations on the paper. They were again asked to write

and submit a revised work reflecting his feedback. This interactional practice occurred four times during each writing assignment, and the instructor stapled and accumulated each task of the assignment so that students had an opportunity to examine their progress. In this way, students learned how to organize, revise and edit their drafts.

(b) Rubrics

When assessing public-speaking performances, traditional forms of assessment are inadequate. Rubrics are employed to measure the success in achieving specific objectives. The rubric may take the form of a checklist, outline, or more formal matrix of skills and standards. Teachers should use rubrics that lay out the various components of the work being completed as well as assign a score or level of competence based on clearly articulated criteria (McGhee & Kozma, 2001). Students are regularly involved in a range of self-assessment and peer assessment activities using rubrics.

However, as this type of assessment becomes more complicated and student responses become more complex, scoring becomes increasingly difficult. A variety of methods should be developed to score student performances.

When he was an education graduate student of an American school, the author received rubric assessments as a formative evaluation during the course and a summative evaluation at the end of each course. The following is an example of a rubric grading report, which was the evaluation sheet of a research paper submitted during a course.

GRADING SCALE		
Excellent	5	
Above Average	4	
Average	3	
Below Average	2	
Poor	1	
Not Present	0	
GENERAL FORMATTING AND PRESENTATION		**COMMENTS**
Title page uses appropriate format.	**5**	(These and

Table of Contents is included and correct page numbers listed.	**5**	other
Document uses all first and second order headings appropriately.	**5**	comments
Appendices are included at the end of the report and sequentially labeled.	**5**	have been removed by
Body of text is double spaced throughout the entire document.	**5**	the author.)
Text correctly and appropriately refers to appendices.	**5**	
		COMMENTS
Introduction is appropriate and appealing, giving necessary background information.	**5**	
Content flows, effectively using transitions.	**5**	
Topic area (problem) is very clear to reader.	**5**	
Body of work thoroughly supports topic/problem.	**5**	
Examples are used effectively and appropriately.	**5**	
Topic is addressed critically; prior learning and common knowledge are synthesized appropriately; examination and explanation of thought processes are logical.	**5**	
Content is appropriate for a wide range of audiences while writing style remains appropriate for graduate-level work.	**5**	
Paper is presented in a professional manner (neat, typed/word processed, etc.).	**5**	
Paper avoids spelling and grammatical errors.	**5**	
Writing style is appropriate for graduate-level work.	**5**	
LITERATURE REVIEW/CITATIONS AND REFERENCES		**COMMENTS**
Literature review is appropriate and supports writer's position.	**5**	
Paper appropriately and effectively reviews and incorporates best practices in literature.	**5**	
Citations are accurate and current.	**5**	
Reviewed literature is appropriate for graduate-level work.	**5**	
References and citations utilize APA format.	**5**	
References and citations correspond with one another.	**5**	
TECHNICAL PRESENTATION		**COMMENTS**
Topic is appropriate for an Action Research Project.	**5**	
Problem statement is clear, precise, and accurate.	**5**	

The participants are clearly identified.	**5**
Goal(s) clearly stated and sets vision for the study.	**5**
Expected outcomes/objectives align with goal(s).	**5**
Expected outcomes/objectives are written in measurable form.	**5**
Measurements and analyses listed are appropriate for the collected data.	**5**
Pre-implementation data is included and sets the foundation for needing the study.	**5**
Solution strategy is well described and the calendar of events for the solution strategy is clearly written and placed in appendix.	**5**
Self-developed instruments have a copy in the appendix. Complete description of any borrowed instrument included.	**5**
Tables are used correctly and effectively (correct titles, incorporated into body of paper, etc.).	**5**
Text correctly and appropriately refers to appendices.	**5**
Figures are used correctly, appropriately and effectively, promoting understanding of the study.	**5**
FINAL REPORT	
Findings are presented to answer each expected outcome/objective.	**5**
Pre-implementation data are contrasted with final data in the findings.	**5**
A discussion follows findings that summaries across expected outcomes/objectives.	**5**
Conclusions are included that are appropriate, accurate, and corroborated.	**5**
Recommendations for future research are included.	**5**
TOTAL POINTS	200

	POSSIBLE	YOU
General Formatting and Presentation	10%	10
General Writing Skills	10%	10
Literature Review/Citations and References	30%	30
Technical Presentation	20%	20
Final Report	30%	30
PERCENT SCORE	100%	100

OVERALL COMMENTS:

Outstanding work! You have made all of the necessary changes. Chapter V is also solid work and shows you have a good understanding of how to analyze the results and provide recommendations for additional research.
(All further comments have been removed by the author.)

All the items in each grid make a segmentalized and specific point to evaluate the work in a well-balanced manner. This rubric can be massive or small in size; however, it can be effective when used as a summative assessment at the end of a term. It can also be applied to a formative assessment if it is simplified. No matter how it is used, it is of utmost importance that students receive rubrics before they start a specific task. This will help students recognize what points they should master, increase learning, and expect better outcomes.

8. Curriculum Assessment

The relationship between curriculum assessment and student assessment is very close and likened to two sides of a coin. Effective assessment enables students to recognize "achievement and make progress throughout their time at school, and teachers to shape and adapt their teaching to individual needs and aspirations. Assessment is at the heart of an effective curriculum and is a fundamental part of good teaching and learning" (QCA, 2009, p. 2). Done properly, it should also have a positive impact on student progress.

To ensure its continued success, the curriculum needs a mechanism for periodical modification. According to Hlebowitsh (2005), "the evaluation of the curriculum not only needs to speak to whether key purposes have been met, but also needs to give the curriculum developer some handle on how to make adjustments or improvements to the curriculum" (p. 214). The curriculum should be subjected regularly to the opinions of teachers, students, parents, and stakeholders to attain the school's educational success.

9. Conclusion

Teachers need and require more advanced skills to conduct complex alternative assessments. They should be carefully designed and students need to understand their purpose. In this respect, alternative assessments, in all likelihood, will increase the teacher's workload.

To do the job properly, teachers constantly need to discuss ideas, ask questions, and share practices regarding assessment methods. Faculty training will be necessary to best use the data and make it applicable to the students. However, educators should not simply use the data in a cold, unforgiving manner. There is still a place for intuition when teachers are dealing directly with their students.

Any assessment method should also require communication with the stakeholders to meet parent and student expectations regarding student learning. When teachers inform students and their parents about their academic achievement, reporting forms should incorporate effective evaluations to help students attain a well-balanced understanding of their progress and assist them in attaining sound educational development. The relationship between curriculum assessment and student assessment needs to be discussed further, and the author plans to explore this in future research.

This study proposes that teachers benefit from regularly scheduled mutual discussions in which information about class activities and ways to evaluate student performance is shared. By putting forward good ideas and sharing them with colleagues, teachers can make their classes more communicative and their students' performances more active. Hence, this study will be of value to those teachers of English who wish to look intently at their own English classes and evaluation systems.

References

Bachman, L.F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Berman, J. (2007). *Maryland charter school founder's manual* (2nd ed.). Maryland Charter School Network.

Blankstein, A. M., Houston, P. D., & Cole, R. W. (2009). *Building sustainable leadership capacity*. CA: Corwin Press.

Dornyei, Z. (2005). *Motivational strategies in the language classroom* (5th ed.). Cambridge: Cambridge University Press.

Grant, H. (1987). *A guide to language testing, development, evaluation, research.* Cambridge, MA: Newbury House Publishers.

Heaton, J. B. (1975). Writing English Language Tests. *A Practical Guide for Teachers of English as a Second or foreign Language*. London: Longman

Hlebowitsh, P. S. (2005). *Designing the school curriculum.* Pearson Education, Inc.

Japan Times Online. (January 7, 2009). Editorial: New high school guidelines. *The Japan Times*. Retrieved September 28, 2010, from http://search.japantimes. co.jp/cgi-bin/ed20090107a1.html

Madsen, H.S. (1984). *Techniques in Testing.* Oxford: Oxford University Press.

McGhee, R., & Kozma, R. (2001). *New teacher and student roles in the technology-supported classroom.* Retrieved August 28, 2010, from http://www.cehd.umn.edu/carei/reports/docs/NewRolesTechnology.rtf

MEXT. (2010). *Improvement of academic ability.* Retrieved August 18, 2010, from http://www.mext.go.jp/english/org/struct/014.htm

MEXT. (2009). *The course of study for upper-secondary school.* Tokyo: Ministry of Education, Culture, Sports, Science and Technology.

New Zealand Government. (2009). *Assessment online.* Retrieved December 10, 2009, from http://assessment.tki.org.nz/Assessment-in-the-classroom/ Assessment-for-learning

Nunan, D. (1988) *The learner-centred curriculum.* Glasgow: Cambridge University Press.

NewSchools Venture Fund. (1999). *Alternatives to standardized tests.* Retrieved August 11, 2010, from http://www.newschools.org/about/news/articles/ alternatives-to-standardized-tests

OECD. (November, 2005). Formative assessment: Improving learning in secondary classrooms. *The OECD policy brief. Nov.*

QCA. (2009). Qualifications and curriculum authority: *Assessing pupils' progress.* London: SecEd.

University of Warwick. (2009). *What and who is assessment for?* Retrieved July 10, 2010, from http://www.ukcle.ac.uk/resources/assessment/guide/who.html

Yoshida, K. (1993). *A guide to oral communication (A), (B), (C).* Tokyo: Buneido Publishing Co.