

The Use of the SW TOEIC Test in University English Education

大学英語教育における TOEIC SW テスト導入の有意性

Gerardine McCROHAN (大学教育開発センター講師)

Introduction

More than 30 years ago the traditional TOEIC LR test was conceived in Japan and created by the Educational Testing service (ETS), an institution for developing and administering tests to measure the proficiency of English as a foreign language. ETS was asked to create a “test system which accurately measures communication skills in English, give standards as to where the skills are in terms of expertise and occupation, and where they are in relation to objectives and goals” (Mitsuhashi, 2010). Additionally, the test content was to be able to measure a wide range of abilities from beginner to advanced level. The results were to “allow the test taker to measure his or her own skills, while at the same time allowing companies or schools to utilize the scores for making many important decisions” (Mitsuhashi, 2010). ETS was also asked to design a test that would be “capable of permitting test taking by large numbers of people at once” and to be “capable of being given relatively affordable and conveniently” (Mitsuhashi, 2010).

The test created under these conditions consists of two separate sections, listening and reading. The test, which takes 2 hours, has two sections each with 100 questions for a total of 200 questions. The score for each section ranges from 4 to 495, with totals ranging from 10 to 990.

The TOEIC LR test was created at a time when corporations still had not much need for employees with high levels of speaking and writing skills. Because of this, ETS decided not to directly test speaking and writing abilities. ETS did carry out validity tests to examine the correlation between the TOEIC Reading/Listening score and a separate speaking/writing test. They found a correlation rate of 0.83, indicating that the LR TOEIC test scores could be used to predict a test taker’s speaking and writing abilities (Mitsuhashi, 2010).

This then raises the question that if the LR TOEIC test is an accurate test of a test taker’s speaking and writing abilities why did the ETS introduce the TOEIC Speaking and Writing tests (TOEIC SW tests), in January 2007?

One reason is that although the correlation between a test taker’s TOEIC listening and speaking score, and reading and writing score, is quite high (0.83) it is not a perfect match (1.0), so it could not be concluded that for some test takers the difference between their listening and reading abilities, and their speaking and writing abilities may be statistically significant.

In addition, there was an increasing demand by companies and universities for a more direct and accurate measurement of speaking and writing skills due to the increased use of English in the workplace and because of the advancement of technology such as e-mail and video conferencing.

One of the main differences between the TOEIC LR and TOEIC SW tests is that the LR test is norm-referenced which indicates the position of a test taker relative to other test takers. In contrast, the SW test is a criterion-based test, in which the test taker is evaluated on how well a set of certain standards have been met. Therefore, the two tests are qualitatively different. The IIBC (The Institute for International Business Communication) released information regarding the statistical relationships between the TOEIC LR and the TOEIC SW. This information is shown in Table 1.

Table 1 Statistical relationships between the TOEIC LR and the TOEIC SW tests

Level (LW)	L	R	S	W
5 (860-990)	430-495	430-495	150-200	160-200
4 (730-860)	365-430	365-430	130-150	140-160
3 (470-730)	235-365	235-365	100-130	110-140
2 (220-470)	110-235	110-235	70-100	70-100
1 (5-220)	5-110	5-110	0-70	0-70

(http://www.toEIC.or.jp/toEIC_en/about/)

Because the TOEIC LR has been around so long, it is well known and respected by employers and teachers as a measure of a test taker's proficiency level. The type of questions, the scoring system and proficiency levels are widely understood and there is a multitude of books giving advice on how to achieve better results.

The same cannot be said for the TOEIC SW test. For this reason, what follows is a detailed description of the test itself, how it is administered and how the scores relate to a test taker's proficiency level.

Description of the TOEIC SW test.

The TOEIC SW test provides an accurate way for score users to assess the test taker's ability to read and write in real life situations. As opportunities for using English are growing in both business and daily life situations, the need for English communication skills, particularly for speaking and writing has risen. The TOEIC SW tests were developed by ETS as a way to directly assess the speaking and writing skills to answer this need.

The TOEIC SW test measures language skills that are used in daily life and in a business context. Test takers speak and write in English their responses to real-world questions and scenarios. The tests use common everyday business phrases, key expressions and vocabulary.

Both sections (speaking and writing) of the test are delivered through the Internet. Each test taker has a computer, monitor and head set. In the speaking test, test takers hear and respond to questions using the head set. In the writing sections, test takers respond to questions on their monitors. Their written answers are typed and saved.

The rating process for the TOEIC SW takes place over the ETS's "Online Scoring Network", which is a secure Internet-based system for evaluating test taker's responses. Responses are sent to the OSN, where they are

scored by certified ETS test raters.

The students are then given a proficiency level for both the speaking and writing sections. (See Appendix 1 for a description of the proficiency levels)

a. TOEIC SW Speaking Test

The TOEIC Speaking test includes 11 questions that measure different aspects of test-takers' speaking ability. This section of the test lasts about 20 minutes.

Table 2 Description of the TOEIC SW Speaking Test

Task	# of items	Response time	Description of task	Evaluation criteria
Read a text aloud	2	45 sec. per item (+ 45 sec. prep. time per item)	Read aloud a short text such as an announcement or advertisement.	Pronunciation Intonation and stress
Describe a picture	1	45 secs. (+30sec. prep time)	Give a verbal description of a photograph	All of the above, & Grammar Vocabulary Cohesion
Respond to questions	3	Q1 & 2 15 secs. Q3. 30secs. (No prep. time)	Respond to everyday questions.	All of the above, & Relevance of content and Completeness of content
Respond to questions using information provided	3	Q1 & 2 15 secs. Q3. 30secs. (No prep. time)	Respond to questions based on written information such as a schedule of events.	All of the above
Propose a solution	1	60 secs. (+ 30 sec. prep time)	Listen to a voice mail message describing a problem give a response recognition of the problem and propose a solution.	All of the above
Express an opinion	1	60 secs. (+ 15 secs. prep time)	Express an opinion with reasons about a specific topic.	All of the above

Speaking test scores are rated as follows.

- Questions 1-9 are rated on a scale of 0-3
- Questions 10-11 are rated on a scale of 0-4
- Sum of all ratings is converted to a scale score of 0-200
- 8 proficiency levels are provided.

(Stafford, 2008)

b. TOEIC Writing Test

The TOEIC Writing test is composed of 8 questions to measure test-takers' writing ability. There are 3 separate tasks and it takes 60 minutes. Table 3 describes this section.

Table 3 TOEIC SW Writing Test

Task	# of items	Response time	General outline	Evaluation criteria
Write a sentence based on a picture	5	8 mins. for 5 items	Write one sentence per picture, using 2 words or phrases provided	Grammar Relevance sentences to pictures
Respond to a written request	2	10 mins for each item	Read an e-mail message and write an e-mail response.	Quality and variety of sentences Vocabulary Organization
Write an opinion essay	1	30 mins.	Write an essay expressing an opinion on the topic provided, giving reasons &/or examples	Whether the opinion is supported with reasons and/or examples Grammar Vocabulary Organization

Speaking test scores are rated as follows:

- Questions 1-3 are rated on a scale of 0-3
- Questions 6-7 are rated on a scale of 0-4
- Question 8 is rated on a scale of 0-5
- Sum of all ratings is converted to a scale score of 0-200
- 9 proficiency levels are provided.

(Stafford, 2008)

TOEIC SW at Kagawa University

The TOEIC LR has long been a feature of the first year general education English classes for all students at Kagawa University but the TOEIC SW was first taught and taken by a small group of students in the spring semester of the 2011-2012 academic year.

In total sixteen 3rd and 4th year students from the Faculties of Education, Technology, Economics and Law sat the exam at the Faculty of Education campus on July 30th, 2011.

All 16 students were enrolled in the General Education Advanced English classes held at the Education and Technology campuses in the 2011 spring semester. Although the class focused on preparation for the TOEIC SW using a textbook specially written for TOEIC SW preparation (*Successful keys to the TOEIC Speaking and Writing Tests*, by M.D. Stafford and published by Pearson Longman) students were not required to take the exam in order to get credit for the class. As a result, 11 of 14 students at the Advanced English class held on the Faculty of Education campus and 5 of 12 students from the class held at the Faculty of Technology campus chose to take the test.

At the start of the semester students were asked for their TOEIC LR score (if they remembered it) so that their scores from both tests could be compared. Some students could not remember their exact TOEIC LR score so their scores are listed in the LR 1-5 score bands (See Table 1 for a description of these bands).

Table 4 Results of TOEIC SW Test

Student/Faculty	LR 1-5 score band	Speaking/ proficiency level	Writing/ proficiency level	Total
1. Technology	5	140 6	160 7	300
2. Education	4	140 6	150 7	290
3. Education	4	140 6	140 7	280
4. Education	4	130 6	140 7	270
5. Technology	4	130 6	140 7	270
6. Education	4	120 5	150 7	270
7. Law	3	130 6	130 6	260
8. Technology	4	120 5	140 7	260
9. Technology	4	120 5	130 6	250
10. Education	4	110 5	140 7	250
11. Technology	3	120 5	120 6	240
12. Economics	3	100 4	120 6	220
13. Education	3	100 4	90 5	190
14. Economics	3	80 4	110 6	190
15. Economics	2	70 3	120 6	190
16. Economics	2	80 4	100 5	180
		Mean=114.4 Sd. Dev=22.5	Mean=130.0 Sd. Dev.18.97	Mean=244.37 Sd. dev. 38.3

The average in Japan for the speaking section is 124 and for the writing section is 142. Students 1 to 5 had a higher score than the Japanese average in the listening section and students 1, 2 and 6 also had a higher score in the writing section. However, students 3, 4, 5, 8 and 10 had a writing score very close to the Japanese average. As can be seen from the table, nearly all students did better in the writing section than in the speaking, some with a significant difference in scores. Two students had equal scores in both sections and one student did better in the speaking section than in the writing section. All students reported after the exam that they found the speaking very difficult, mainly because of lack of time. This suggests that they need more fluency building practice. In practice tests taken before the official TOEIC SW test, students showed a similar pattern. Most of them struggled with questions where there was no vocabulary given in the task such as the picture description or in the final question where they had to express an opinion. Where there was information given which included vocabulary, students were very able to use it efficiently when doing the task. In question 10 of the speaking section, where student had to propose a solution to a problem, students with lower proficiency levels found this extremely difficult. In this question, test takers listen to a voice message describing a problem. In this description a lot of necessary but also unnecessary information is included. Students with weaker listening skills struggled to differentiate between those details that were crucial to answering the question and those which were not. Most students also found it difficult to catch the names of people and places that were usually needed to give a high-scoring reply.

Students found the writing section much easier and more familiar than the speaking section. Most students have had extensive practice forming complex sentences (sentences with more than one clause) at high school and

in other university English classes and were able to use this when answering questions 1-5 of the writing test. Questions 6 and 7, writing responses to business e-mails, were less familiar to students but most found it not too difficult to learn the fixed phrases and more formal vocabulary used in business English. The final question, an opinion essay, was again a familiar task for all students. Students were already familiar with the basics of essay writing in English but depending on the topic, some found it more or less difficult. In practice essays, students with a weaker vocabulary sometimes found it difficult to express their ideas and opinions clearly. Many students also exhibited weakness in using discourse markers with over use of “In addition” a common problem. Students were usually aware of other discourse markers but were not used to using them. Giving specific examples and details caused some problems too, but again this depended on the essay topic.

The types of problems students had with the essay question suggests that an academic writing course offered in the general education classes would be of considerable benefit to students who, in the future, will need to write research papers and articles.

One final point to note was that all responses in the writing section had to be typed. Many, if not most, students were unaccustomed to typing in English and were unaware of basic typing rules to do with punctuation, spacing and indentation. Basic computer skills such as the ability to use word processing software is a skill that all people need today and something our students should learn before leaving university.

Conclusion

For students who wish to work in an English-speaking environment, preparation for the TOEIC SW test will give them the skills to improve their ability to use English in a real-life situations. For many university students, their exposure to English has often been more focused on receptive skills such as reading and listening. By preparing for and taking the TOEIC SW test, students are given extensive exposure to the more productive skills of speaking and writing.

In the General Education Advanced English classes, all students sat a mini TOEIC SW test prior to the start of the course. They repeated a similar mini TOEIC SW test as their final examination. When the results of both mini tests were compared, they all showed considerable improvement in their abilities to speak more fluently, had a better understanding and ability to use “fillers” to gain thinking time, and were able to write basic business letters in English. These are all skills needed to work in an English-speaking environment. Since the TOEIC SW was introduced in 2007, a number of universities around Japan have been using it including Kobe University and Tokiwa University.

At Tokiwa University, research has focused on classifying students according to the results of the four scores of the TOEIC LR and TOEIC SW tests. They have also tried to determine what causes gaps in a student’s proficiency in any of the 4 skills tested by the two types of TOEIC tests. To date, they have not been able to identify any factors that cause gaps in specific skills but they have successfully classified students according to gaps between the four skills. They believe it is important to take both the TOEIC LR and TOEIC SW tests, in order to examine the balance of the four skills (Koizumi, 2010). Studies done at Kobe University have found that a test taker’s speaking skill is the most highly correlated with

over all proficiency (Ishikawa, 2010). As with the students at Kagawa University, they found that the speaking test was the most difficult section of the TOEIC SW tests no matter what the test taker's proficiency level. When all 4 skills were analyzed they found that the section with the highest score was in the listening test for advanced learners, but in the writing test for intermediate and lower level test takers. This suggests that high level listening abilities is characteristic of advanced learners. Moreover, when they examined the gap between average scores for elementary, intermediate and advanced learners, they found that the advanced group had smaller score differences between the tests, but as the level of achievement decreased, the greater the gap became. They concluded that since the correlation between a test takers speaking skill and their over all proficiency is highest and that the speaking test only takes 10 minutes to administer, the speaking test may be a way of extrapolating the overall proficiency of the four skills.

At present we do not have the facilities to administer the TOEIC SW speaking section to all first year students at Kagawa University, but because of the high correlation of over 0.9 between the TOEIC LR and TOEIC SW tests, our present system of placing students into graded classes according to their TOEIC LR test results should be sufficient.

However, since the estimated precision of the LR test in measuring overall proficiency differs according to the test taker's level of achievement, it is more desirable to directly measure proficiency in all four skills if we wish to develop a more accurate picture of our students' abilities in all four skills, and to determine any gaps they may have. This is particularly true of students who have lived and studied abroad as they often have very good listening skills but weaker knowledge of more formal vocabulary and sentence patterns used in academic writing and in the business world. Using both scores of the TOEIC LR and the TOEIC SW test would be an effective way of doing this and would allow us to further develop the overall proficiency of our more advanced students.

Appendix 1 TOEIC® Speaking Writing Test - Proficiency Level Descriptors (The Institute for International Business Communication)

Speaking

Proficiency Level 8 Score 190-200

Typically, test takers at level 8 can create connected, sustained discourse appropriate to the typical workplace. When they express opinions or respond to complicated requests, their speech is highly intelligible. Their use of basic and complex grammar is good and their use of vocabulary is accurate and precise. Test takers at level 8 can also use spoken language to answer questions and give basic information.

Their pronunciation and intonation and stress are at all times highly intelligible.

Proficiency Level 7 Score 160-180

Typically, test takers at level 7 can create connected, sustained discourse appropriate to the typical workplace. They can express opinions or respond to complicated requests effectively. In extended responses, some of the following weaknesses may sometimes occur, but they do not interfere with the message: -minor difficulties with pronunciation, intonation, or hesitation when creating language -some errors when using complex grammatical structures -some imprecise vocabulary Test takers at level 7 can also use spoken language to answer questions and give basic information. When reading aloud, test takers at level 7 are highly intelligible.

Proficiency Level 6 Score 130-150

Typically, test takers at level 6 are able to create a relevant response when asked to express an opinion or respond to a complicated request. However, at least part of the time, the reasons for, or explanations of, the opinion are unclear to a listener. This may be because of the following: -unclear pronunciation or inappropriate intonation or stress when the speaker must create language -mistakes in grammar

-a limited range of vocabulary Most of the time, test takers at level 6 can answer questions and give basic information. However, sometimes their responses are difficult to understand or interpret. When reading aloud, test takers at Level 6 are intelligible.

Proficiency Level 5 Score 110-120

Typically, test takers at level 5 have limited success at expressing an opinion or responding to a complicated request. Responses include problems such as: -language that is inaccurate, vague, or repetitive -minimal or no awareness of audience -long pauses and frequent hesitations -limited expression of ideas and connections between ideas -limited vocabulary Most of the time, test takers at level 5 can answer questions and give basic information. However, sometimes their responses are difficult to understand or interpret. When reading aloud, test takers at Level 5 are generally intelligible. However, when creating language, their pronunciation, intonation and stress may be inconsistent.

Proficiency Level 4 Score 80-100

Typically, test takers at level 4 are unsuccessful when attempting to explain an opinion or respond to a complicated request. The response may be limited to a single sentence or part of a sentence. Other problems may include: -severely limited language use -minimal or no audience awareness

-consistent pronunciation, stress, and intonation difficulties -long pauses and frequent hesitations -severely limited vocabulary Most of the time, test takers at level 4 can not answer questions or give basic information. When reading aloud, test takers at Level 4 vary in intelligibility. However, when they are creating language, speakers at level 4 usually have problems with pronunciation, and intonation and stress. For more information, check the “Read Aloud Pronunciation and Intonation and Stress ratings.”

Proficiency Level 3 Score 60-70

Typically, test takers at level 3 can, with some difficulty, state an opinion, but they cannot support the opinion. Any response to a complicated request is severely limited. Most of the time, test takers at level 3 cannot answer questions and give basic information. Typically, test takers at level 3 have insufficient vocabulary or grammar to create simple descriptions. When reading aloud, speakers at level 3 may be difficult to understand. For more information, check the “Read Aloud Pronunciation and Intonation and Stress ratings.”

Proficiency Level 2 Score 40-50

Typically, test takers at level 2 cannot state an opinion or support it. They either do not respond to complicated requests or the response is not at all relevant. In routine social and occupational interactions such as answering questions and giving basic information, test takers at level 2 are difficult to understand. When reading aloud, speakers at level 2 may be difficult to understand. For more information, check the “Read Aloud Pronunciation and Intonation and Stress ratings.”

Proficiency Level 1 Score 0-30

Test takers at level 1 left a significant part of the TOEIC Speaking Test unanswered. Test takers at level 1 may not have the listening or reading skills in English necessary to understand the test directions or the content of the test questions.

TOEIC® Speaking Test Pronunciation Levels

- | | |
|----------|---|
| 3 HIGH | When reading aloud your pronunciation is highly intelligible. |
| 2 MEDIUM | When reading aloud your pronunciation is generally intelligible with some lapses. |
| 1 LOW | When reading aloud your pronunciation is not generally intelligible. |

TOEIC® Speaking Test Intonation and Stress Levels

3 HIGH	When reading aloud your use of intonation and stress is highly effective.
2 MEDIUM	When reading aloud your use of intonation and stress is generally effective.
1 LOW	When reading aloud your use of intonation and stress is generally not effective.

Writing

Proficiency Level 9 Score 200

Typically, test takers at level 9 can communicate straightforward information effectively and use reasons, examples, or explanations to support an opinion. When using reasons, examples, or explanations to support an opinion, their writing is well-organized and well-developed. The use of English is natural, with a variety of sentence structures and appropriate word choice, and is grammatically accurate.

When giving straightforward information, asking questions, giving instructions, or making requests, their writing is clear, coherent, and effective.

Proficiency Level 8 Score 170-190

Typically, test takers at level 8 can communicate straightforward information effectively and use reasons, examples, or explanations to support an opinion. When giving straightforward information, asking questions, giving instructions, or making requests, their writing is clear, coherent, and effective. When using reasons, examples, or explanations to support an opinion, their writing is generally good. It is generally well-organized and uses a variety of sentence structures and appropriate vocabulary. It may also include one of the following weaknesses:

-occasional unnecessary repetition of ideas or unclear connections between ideas -noticeable minor grammatical mistakes or incorrect word choices

Proficiency Level 7 Score 140-160

Typically, test takers at level 7 can effectively give straightforward information, ask questions, give instructions, or make requests but are only partially successful when using reasons, examples, or explanations to support an opinion. When attempting to explain an opinion, their writing presents relevant ideas and some support. Typical weaknesses at this level include: -not enough specific support and development for the main points

-unclear connections between the points that are made -grammatical mistakes or incorrect word choices When giving straightforward information, asking questions, giving instructions, or making requests, their writing is clear, coherent, and effective.

Proficiency Level 6 Score 110-130

Typically, test takers at level 6 are partially successful when giving straightforward information or supporting an opinion with reasons, examples, or explanations. When giving straightforward information, asking questions, giving instructions, or making requests, the message omits important information or is partly unintelligible.

When attempting to explain an opinion, their writing presents relevant ideas and some support. Typical weaknesses at this level include: -not providing enough specific support and development for the main points -unclear connections between the points that are made -grammatical mistakes or incorrect word choices

Proficiency Level 5 Score 90-100

Typically, test takers at level 5 are at least partially successful when giving straightforward information. However, when supporting an opinion with reasons, examples, or explanations, they are mostly unsuccessful. When giving straightforward information, asking questions, giving instructions, or making requests, the message omits important information or is partly unintelligible.

When attempting to explain an opinion, significant weaknesses that interfere with communication occur, such as: -not providing enough examples, explanations, or details to support the opinion or they are inappropriate -inadequate organization

or connection of ideas -limited development of ideas
-serious grammatical mistakes or incorrect word choices

Proficiency Level 4 Score 70-80

Typically, test takers at level 4 have some developing ability to express an opinion and to give straightforward information. However, communication is limited. When attempting to explain an opinion, significant weaknesses that interfere with communication occur, such as: -not providing enough examples, explanations, or details to support the opinion or they are inappropriate

-inadequate organization or connection of ideas -limited development of ideas -serious grammatical mistakes or incorrect word choices When giving straightforward information, asking questions, giving instructions, or making requests, the responses do not successfully complete the task because of

-missing information -missing or obscure connections between sentences and/or -many grammatical mistakes or incorrect word choices At level 4, test takers have some ability to produce grammatically correct sentences, but they are inconsistent.

Proficiency Level 3 Score 50-60

Typically, test takers at level 3 have limited ability to express an opinion and to give straightforward information. When attempting to explain an opinion, the responses show one of the following serious flaws: -serious disorganization or underdevelopment of ideas -little or no detail or irrelevant specifics

-serious and frequent grammatical mistakes or incorrect word choices When giving straightforward information, asking questions, giving instructions, or making requests, the responses do not successfully complete the task because of -missing information -missing or obscure connections between sentences and/or -many grammatical mistakes or incorrect word choices At level 3, test takers have some ability to produce grammatically correct sentences, but they are inconsistent.

Proficiency Level 2 Score 40

Typically, test takers at level 2 have only very limited ability to express an opinion and give straightforward information. When attempting to explain an opinion, the responses show one of the following serious flaws: -serious disorganization or underdevelopment of ideas -little or no detail or irrelevant specifics

-serious and frequent grammatical mistakes or incorrect word choices At level 2, test takers cannot give straightforward information. Typical weaknesses at this level include: -not including any of the important information -missing or obscure connections between ideas -frequent grammatical mistakes or incorrect word choices At level 2, test takers are unable to produce grammatically correct sentences.

Proficiency Level 1 Score 0-30

Test takers at level 1 left part or parts of the TOEIC Writing Test unanswered. Test takers at level 1 may need to improve their reading ability in order to understand the test directions and the content of test questions.

References

Ishikawa, S. (2010). Aspects of the L2 Proficiency of Japanese Learners of English: What Direct Measurement of Four Skills Shows. Presentation given at *University-level English language education and the TOEIC test* at Campus Plaza Kyoto, March 21, 2010

Koizumi, R. (2010). Gaps Between Four skills Assessed by the TOEIC Test: Examining Learner Groups and Predictor Variables. Presentation given at *University-level English language education and the TOEIC test* at Campus Plaza Kyoto, March 21, 2010

Mitsubishi, M. (2010). Using the TOEIC Test for College English Education: Review of the Characteristics of the TOEIC Test. Presentation given at *University-level English language education and the TOEIC test*

at Campus Plaza Kyoto, March 21, 2010

Stafford, M.D. (2008). Successful keys to the TOEIC Speaking and Writing Tests. Pearson Longman Tokyo
TOEIC® Speaking Writing Test - Proficiency Level Descriptors (The Institute for International Business
Communication)

http://www.toeic.or.jp/toeic_en/pdf/sw/SW_Descriptors_EN.pdf Accessed Dec. 22nd 2010

“Statistical relationships between the TOEIC LR and the TOEIC SW tests” taken from

About the TOEIC® Speaking and Writing Test http://www.ets.org/toeic/speaking_writing/about Accessed
December 22, 2010