

## **An Analysis of the TOEIC® as a Placement at Asia University**

**Daniel Bates**, Asia University

### Abstract

This article questions Asia University's use of the Test of English for International Communication (TOEIC®) as a placement test for Freshman English courses for students belonging to the International Relations department. First, it considers the current use of placement tests for English courses at Asia University and the significance of the TOEIC® for students and educators at Asia University. Next, distinctions between different language tests are defined, followed by a review of current research in the field regarding the use of the TOEIC® as a placement test in English as a Second Language (ESL) courses. Then, the paper considers if the TOEIC® is a relevant indicator of the skills students are expected to master during the Freshman English (FE-IR) classes at Asia University and whether there is a sufficient relationship between the TOEIC® and the content of the curriculum to make the TOEIC® a valid placement test in this context. Furthermore, the paper compares the TOEIC® test with Asia University's own Freshman English Placement test (FEPT) by using an inter-item correlation matrix to compare the relationship between the two tests and determine whether each test is using a common measuring scale to form sub-groups of students based on English proficiency. Finally the article speculates on the future of TOEIC® as a placement test at Asia University, as well as in a wider ESL context, and offers suggestions on how placement testing may take place at Asia University in the future. While this paper focuses specifically on the context of placement testing within Asia University's English Program, the findings and recommendations may be of interest to other universities throughout Japan since, as of 2006, 66 university

institutions in Japan are using the TOEIC® test as a placement test for their English programmes (ETS 2007).

## **Placement Tests at Asia University**

Freshman students entering Asia University are required to study an English course five times a week during their first year of study at the institution. In order to place students into sub-groups of similar English proficiency, a placement test is administered before the opening week of semester. Currently, there are two separate placement tests at Asia University. Students studying in the Business Administration, Economics or Law departments will sit the Freshman English Placement Test (FEPT), a placement test designed by the Assessments Committee at Asia University and comprising of three parts (namely, listening, grammar and vocabulary). Students from the International Relations Department sit the listening and reading sections of the TOEIC® test (rather than the FEPT), as means of placing students into classes of similar English abilities. The historical reason behind using separate placement tests remains unclear. However, the TOEIC® test is used in a number of additional ways by the International Relations faculty beyond just a placement test (which are discussed in greater detail in section three), and this may explain why it was originally adopted as a placement test. As with the FEPT, the results from the TOEIC® placement test are used to assign students to an appropriate English class for the semester. A second TOEIC® placement test is taken by IR students at the end of the first semester in order to place them in to new sub-groups, depending on their levels of improvement or otherwise (as determined by the two TOEIC® scores) throughout the semester. These placement tests are used for two specific purposes: to avoid multilevel proficiency among students in the same class groups, and to allow teachers to prepare materials and use textbooks suitable to that specific group of students' abilities. The placement tests are vital in giving students the best possible environment in which to learn, and also aid teachers in producing materials that appropriately challenge the students' abilities. Therefore, the tests

themselves must be fulfilling the specific requirements of a placement test effectively in order to do this.

### **Test Distinctions**

In order to assess a test's effectiveness as a placement test, one must consider the different purposes tests have in language education. Language testers have made a number of distinctions between tests in order to give a better comprehension of an individual test's strengths and aims. First, tests are classified as either criterion-referenced tests (CRT) or norm-referenced tests (NRT). CRTs are based on a particular curriculum and the scores of such tests quantify the amount of knowledge a student has of those specific criteria. Hence, there needs to be a close correlation between the content of a program and what is being tested. Typically, a CRT is commonly used as an end of unit test in order to determine how much of the material taught in class has been mastered by the students. In contrast, NRTs are used to measure more general language proficiency. The TOEIC® is characterized as an NRT. The aim of NRTs is to disperse the scores of test takers out so that they are normally distributed along a bell curve (Besette, 2007). Unlike the scores of CRTs which show the extent of knowledge a student has of a subject, NRT scores are relative to each other and show us only how well a student has done in comparison to the other test takers.

Brown (1996) further divided tests into the following four categories: achievement, diagnostic (CRTs), proficiency, and placement (NRTs). Achievement tests are typically used as end of course tests to see how well a student has mastered the material from the course, while a diagnostic test is often used for research purposes by educators gathering information about the objectives of a course. Proficiency tests are very general in content, seldom have a connection to a course or syllabus, and are used to compare students, or classes of students, with each other. Placement tests are similarly

general, but unlike proficiency tests, the content of a placement test should cover all levels and skills that are to be mastered in a course (Wilson, 1989). Therefore, the relationship between the TOEIC® placement test and the curriculum content at Asia University is particularly significant to this study and will be discussed in greater detail in part three. Brown (1996) stated that the TOEIC® test does not usually make for a good placement test as there is rarely a strong connection with the TOEIC® test and the content of a program or curriculum. Cohn (2007) concurs with Brown's assessment and states that "Of the four uses, proficiency, placement, achievement and diagnostic, the TOEIC® can only legitimately be used for the first" (p.41). This research indicates that the TOEIC® is not suitable as a placement test for the majority of English courses. To justify its use as such, a particular emphasis on TOEIC® preparation and a focus upon listening and reading skills would be required in the curriculum to form a suitably strong association between the curriculum and the placement test. As such, the next section of this paper will focus on the current content of the IR Freshman English curriculum and consider to what extent this relates to the TOEIC® placement test itself and speculate whether this relationship is adequate to justify its use as a placement test at Asia University.

### **Freshman English for International Relations (FE-IR) Curriculum and TOEIC®**

The first semester of the FE-IR curriculum is divided into two distinct areas: communicative English and TOEIC® preparation. Teachers of these courses are required to cover a communicative-based text book for 50% of the course which focuses primarily on students' speaking and listening skills. The other 50% of class time is spent working on a separate TOEIC® preparation book. Textbooks are prescribed for both areas of the course according to the class level with a suggested pacing of

materials. However, each teacher has the flexibility to choose materials and activities that will encourage communicative abilities among the students in the class. In the second semester FE-IR students study from a textbook which focuses on critical thinking skills alongside the four main language skills. Crucial to note, however, is that beyond the first semester there is no additional element of TOEIC® study within the FE-IR course curriculum. However, while there is no further TOEIC® instruction after the first semester, IR students are required to take the TOEIC® on a number of occasions thereafter. Students studying on the Asia University American Program (AUAP) are also required to take the TOEIC® test before and after attending this course. Here, the score is used to place students in a suitable level at the beginning, but also as a tangible measure of improvement in students' language skills at the end of the course. Finally, a TOEIC® score of 600 or above is a requirement of all students in the IR department in order to graduate from Asia University. Evidently, the TOEIC® score is of importance to IR students and this is reflected in the FE-IR curriculum, but this alone does not justify its use as a placement test. What can be surmised from the FE-IR curriculum is that with only 25% of the IR student's English study devoted to TOEIC® preparation, and the remaining 75% spent on communicative language and critical thinking skills, the TOEIC® test as a placement test neither has a strong relationship with the content of the FE-IR program, nor covers all of the skills to be mastered in the curriculum. As this is a prerequisite for placement testing, stated by a number of researchers in part two of this paper, the TOEIC® test could be considered an insufficient measurement for placing students in sub-groups according to the skills required for the courses at hand at Asia University.

The Freshman English program is designed primarily as a way to improve and use communicative abilities in English in an authentic and realistic manner, focusing heavily on speaking and listening skills. While some research indicates that inferences

of oral proficiency can be garnered from one's TOEIC® score (Wilson, 1989), there is significant counter evidence suggesting the results of the TOEIC® do not reflect the communicative abilities of its takers. TOEIC® is a test of receptive skills and therefore has the inherent problem that it does not tell us the functional ability to use the target language. Therefore, some scholars argue that it does not directly test productive skills (Daller & Phelan, 2006). In a related study, Cunningham (2002, p.58) concluded that there was no correlation between the gains in TOEIC® scores and any improvements in one's communicative ability. According to Cunningham, high TOEIC® scores do not necessarily reflect high communicative ability and low TOEIC® scores do not necessarily reflect low communicative ability. With this weight of conflicting research in mind, the TOEIC® test may not be the best representative of students' communicative abilities, and does not test the skills that students will master on the FE-IR course.

At the end of the first semester, the TOEIC® placement test is taken again by the IR students to adjust the sub-groups for the second semester classes. This test is in place to account for any changes in the language abilities of the students that may have occurred from the instruction in semester one. This particular placement test can be problematic for a number of additional reasons. Once again, one should consider if the improvements in one's TOEIC® score are indicative of improvements in the skills required for, and to be mastered in the second semester. The curriculum for the second semester does not feature any TOEIC®-related material so the placement test does not have any connection at all with the content of the second semester. Furthermore, a TOEIC® score, as part of an NRT, is relative to the other scores of other test takers, so the results and sub-grouping of students based upon them could have a psychological impact on some students. For example, a student who improves their score by 100 points over the semester, an achievement in itself, could potentially move to a lower-

level class if a significant number of his/her peers' scores increase more dramatically. This could negatively affect a student's motivation or give them an inferiority complex among their peers in the English classroom (Shimizu, 1999, p.240). Of course, while this is something of a risk among all placement tests, it seems to be particularly unfair on a student to have their motivation damaged by a test with a limited connection to the course they have just completed, and with no direct connection to the skills that will be mastered in the upcoming course in which they are being placed.

Another troublesome issue in using the TOEIC® as a placement test is the broad nature of the test itself. As mentioned in section two, the goal of NRTs is to spread the scores of test takers out so that they are normally distributed along a bell curve. As the TOEIC® is designed to be taken by many thousands of candidates, the distribution is normed from a large population of test takers. Comparing this to an institution with only a few hundred test takers (for instance, 343 test takers at Asia University in 2016) could negate the value of the scores as good proficiency indicators (Brown, 1996).

### **Correlation Results Between the TOEIC® and FEPT Placement Tests**

Having established that there is a weak connection between the TOEIC® placement test and the IR-FE curriculum, and raising some additional concerns related to the use of the TOEIC® as a placement test at Asia University, this paper will now proceed by analyzing the results of a correlation test with the other placement test in place at Asia University, the FEPT. Testing the correlation between the two tests will determine if, and to what extent, the two placement tests are testing students for the same thing. Aside from the TOEIC® preparation element of FE-IR, the two curricula for which the differing placement tests are used are similar in nature so, one would expect some kind of correlation to be found between the two placement tests. One caveat in assessing the correlation between these two placement tests in particular is the



significantly different numbers of students taking the two tests. The FEPT covers three large departments at Asia University and had 1438 students sit the test in April 2016. In comparison, the TOEIC® placement test is only administered to the IR department, and as such, only saw 343 students sit the test during the same test period. While the sample sizes are unequal, it is hoped that a correlation matrix will determine, to some degree, whether or not a relationship between the two tests exists.

In order to understand the correlation data gathered in Table 2 and Table 3, it is pertinent to review the consensus of what is considered a lower limit for acceptable correlation. While there are no concrete rules, Nunnally (1978) suggests 0.7 as the lower limit. Others have suggested that 0.6 will be sufficient for research purposes. (Daller & Phelan, 2006, p.107). Accordingly, we can assess the correlation matrix (tables 2 and 3) by referring to the table 1 below.

<b>Correlation coefficients magnitude</b>	<b>Correlation relationship</b>
0.9-1.0	Very high correlation
0.7-0.9	High correlation
0.5-0.7	Moderate correlation
0.3-0.5	Low correlation
<0.3	Little (if any) correlation

*Table 1.* Understanding Correlation Coefficients (Calkins, 2005).

Referring to Table 2, it is clear that absolutely no correlation exists between the TOEIC® and FEPT placement tests in either the listening, reading or total scores. In Table 3, the total correlation, between the TOEIC® placement and FEPT placement tests is .008. This indicates that there is no correlation between the total scores of the two placement tests. From this we can surmise that the two placement tests are not

testing the same skills of the examinees. Furthermore, Table 2 shows that the listening section of the TOEIC® correlates to 0.27 with the FEPT listening section, while the reading section of the TOEIC® placement test correlates to -.030 with the FEPT counterpart. In essence, neither part of the TOEIC® placement test has any significant correlation with the FEPT placement test, and by definition, there is no correlation between the two placement tests as a whole. Having already determined that the TOEIC® does not have a strong relationship with the FE-IR curriculum, the lack of correlation between the two placement tests is concerning and this data indicates that the questions on the two placement tests are not testing for the same skills.

	TOEIC® Listening	TOEIC® Reading	TOEIC® Total	FEPT Listening	FEPT Reading
TOEIC® Listening	-				
TOEIC® Reading	.540	-			
TOEIC® Total	.913	.837	-		
FEPT Listening	.027	-.010	.013	-	
FEPT Reading	.028	-.030	.003	.759	-
FEPT Total	.029	-.022	.008	.932	.944

Table 2. Inter-Item Correlation Matrix (a).

	FEPT Total
TOEIC® Listening	.029
TOEIC® Reading	-.022
TOEIC® Total	.008
FEPT Total	1.000

Table 3. Inter-Item Correlation Matrix (b).

### Conclusion

This paper set out to determine the validity of the TOEIC® test as a general placement test for L2 English classes, and more specifically, as a placement test at Asia University for FE-IR students. Furthermore, the paper looked at the relationship between the TOEIC® placement test and the FEPT placement test given to Freshman English students from departments other than IR.

In conclusion, there is significant research in the field that suggests the TOEIC® is not suitable as a placement test for ESL courses in general due to the prerequisite that a placement test must have a strong connection to the curriculum of the course. Specifically at Asia University, the students taking the TOEIC® as a placement test will study TOEIC®-related materials for approximately 25% of the following year of English study (the rest of the course being made up of communicative English and critical thinking skills). Thus, we cannot surmise that a TOEIC® placement test is sufficiently covering the skills required to complete and master the skills taught during the year. Therefore, I argue that the TOEIC® placement test is not placing students into appropriate groups based on the skills they will need in order to master the majority of the course ahead.

In addition, there is no significant correlation between either of the major sections (listening and reading) of the TOEIC® and FEPT placement tests, or indeed the tests in general. This indicates that the skills being assessed are different on each placement test, which leads to potential irregularities in the placing of students into “appropriate” levels.

I suggest that the TOEIC® should be discontinued as a placement test at Asia University; and, on a wider level, it should not be used for English courses which are not based on 100% TOEIC®-based curricula. The results of the correlation matrix and the test’s lack of connection with the vast majority of the curriculum studied thereafter suggests two salient points; one, that the TOEIC® is not fulfilling the same role as the FEPT, and two, that it is not a suitable choice as a placement test for a curriculum that is largely based on communicative, productive skills. Replacing the TOEIC® placement test at Asia University should not be of great difficulty. While the validity and reliability of the FEPT is beyond the scope of this paper, the FE courses at Asia University are of a similar nature to the FE-IR course and thus a single placement test for all FE and FE-IR classes would be more appropriate and serve as a more valid indicator of assessment of the skills to be mastered upon both of these courses.

## References

- Bessette, A. (2007). TOEIC: Uses and Misuses: "Testing improves education the same way that bombing promotes democracy" Steve Cohn, Education professor at Tufts University. *プール学院大学研究紀要*, 47, 35-45.
- Brown, J.D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.
- Calkins, K.G. (2005). Correlation coefficient. Retrieved from <https://www.andrews.edu/~calkins/math/edrm611/edrm05.htm>.
- Cunningham, C. R. (2002). *The TOEIC Test and Communicative Competence: Do Test Score Gains Correlate With Increased Competence?* University of Birmingham (Internal Publication).
- Educational Testing Service (ETS). (2007). TOEIC test data and analysis for 2006. Retrieved from [http://www.toEIC.or.jp/library/toEIC\\_data/toEIC\\_en/pdf/data/TOEIC\\_Utilization\\_2007.pdf](http://www.toEIC.or.jp/library/toEIC_data/toEIC_en/pdf/data/TOEIC_Utilization_2007.pdf).
- Daller, H. and Phelan, D. (2006) The C-test and TOEIC® as measures of students' progress in intensive short courses in EFL. In Grotjahn, R. (ed.) *The C-Test: Theoretical basis and practical applications*. Vol.4. (pp. 101-119). Frankfurt: Lang Verlag.
- Moritoshi, P. (2001). *The Test of English for International Communication (TOEIC®): necessity, proficiency levels, test score utilization and accuracy*. The University of Birmingham (Internal Publication).
- Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Shimizu, Y. (2002). Survey Research on the Use of Placement Tests at Four-Year Universities in Japan. *立命館言語文化研究*, 14(1), 231-243.
- Wilson, K.M. (1989). *Enhancing the interpretation of a norm-referenced second language test through criterion-referencing: A research assessment of experience in the TOEIC testing context (TOEIC Research Report Number 1)*. Princeton, NJ: Educational Testing Service.