

Review of Language Proficiency Tests: TOEFL iBT, IELTS, PTE Academic

Sookhee Jeong

Colorado State University

### Abstract

English language proficiency tests are often used to make high-stakes decision about Korean test takers who are applying for Teaching English as a Foreign/Second Language (TEFL/TESL) MA programs. This paper explores three certified English proficiency tests which are generally accepted in most TEFL/TESL MA programs: TOEFL iBT, IELTS and PTE Academic. The purpose of this review is to decide which test would be best for Korean test takers after describing the structure, the item types, the scoring and the statistics, and evidence of reliability and validity of each test. I have concluded that IELTS would be most appropriate for Korean test takers after examining all factors of each test.

*Key words:* TEFL/TESL, proficiency test, TOEFL iBT, IELTS, PTE Academic

The Korean public English education has been test-oriented to enter universities and focuses on listening and reading with little attention given to communicative teaching. Therefore, the English education at public schools has been criticized because the students who graduate high school do not communicate in English successfully though they have studied English at school for 10 years. The English teachers have been at the center of criticism because it is believed that the lack of English teachers' proficiency has stunted communicative teaching. Therefore, English teachers in Korea are required to develop their own English proficiency to teach English in a communicative way. As a result, English teachers in Korea keep trying to enhance their English ability and some of them decide to study abroad, like me. In addition, the Korean government drives communicative English teaching strongly; the number of current teachers or future teachers who want to study Teaching English as a Foreign Language (TEFL) or Teaching English as a Second Language (TESL) in English speaking countries seems to keep increasing. The people who want to study abroad, especially in English speaking countries for a MA in TEFL/TESL, must submit a certified English proficiency test score to get admitted, as well as a GPA, a statement of purpose or letters of recommendation.

The test of English as a Foreign Language (TOEFL) or International English Language Test System (IELTS) are the two most popular tests of English proficiency in Korea; the Pearson Test of English (PTE) Academic is relatively new and gaining popularity nowadays. The majority of graduate schools in the US accept paper-based type TOEFL (TOEFL PBT) and internet-based TOEFL (TOEFL iBT). Specifically, as for Masters in TEFL/TESL at CSU, international applicants must submit one of these English proficiency tests (the required minimum scores are in parentheses): TOEFL iBT (80), TOEFL PBT (550), or IELTS (6.5). CSU has not accepted PTE Academic for admission yet. In the case of the University of San

Francisco, international students applying for master are required to submit TOEFL PBT (580), TOEFL iBT (92 with the sub score lower than 20), IELTS (7.0), or PTE Academic (62).

This paper explores three certified English proficiency tests: TOEFL iBT, IELTS and PTE Academic. Similarly, all three tests are designed to test applicants' language ability in the four skills: reading, writing, speaking and listening. These three tests provide applicants' overall proficiency indicating applicants' overall language proficiency. These four languages skills are necessary for international students to live successfully in the community they will belong to, as well as succeed academically. To be specific, English proficiency is required to ensure success in the TEFL/TESL Masters program, and English proficiency plays a crucial role in their daily lives for socializing, shopping, making appointments, reserving a hotel, etc.

The review is expected to contribute to judging which of the three tests is the best for current Korean English and future teachers who prepare for an English proficiency test to apply for the Masters program in TEFL/TESL in a country where English is used as the main instructional language. It might be valuable to give them useful information to decide which test they should take by reviewing these three tests, in terms of the purpose, the structure, the item types, the scoring, the statistics of each test and reliability and validity.

## The TOEFL iBT®

<i>Publisher:</i>	Educational Testing Service (ETS), ETS Corporate Headquarters is 660 Rosedale Road, Princeton, NJ 08541 USA; telephone 1-609-921-9000; fax 1-609-734-5410; <a href="http://www.ets.org/toefl">http://www.ets.org/toefl</a>
<i>Publication Date:</i>	September, 24, 2005
<i>Target Population:</i>	Those whose native language is not English who want to study at a higher education institution where English is used.
<i>Cost:</i>	Fees vary by country from US \$160 to US \$250 (US \$ 170 in case of Korea)

**Overview**

The TOEFL PBT was administrated in 1961 first, then a computer-based TOEFL (TOEFL CBT) was introduced in 1998. The TOEFL iBT® is the newest version of the TOEFL administrated by the Internet and ETS offers two formats for the test currently: the TOEFL iBT® and PBT. The TOEFL iBT® is designed to measure applicants' English proficiency in academic settings with the university level. It is made up of listening, reading, speaking and writing sections. TOEFL scores are accepted by more than 8,500 colleges or institutions in over 130 countries. The test is administrated 30–40 times a year at more than 4,500 authorized test centers in 165 countries. Specifically, the test is run in 28 cities throughout the country in Korea and in case of Seoul, the biggest city in Korea, and the test is administrated three times a month. A specific description of the test is provided (see Table 1).

Table 1

Specific description for The TOEFL iBT® (Adapted from ETS: *The TOEFL iBT®*, 2012)

Purpose	<p>The TOEFL iBT® measures the English proficiency of the people whose mother language is not English in academic settings where English is used. Sawaki et al. (2009) state that the main purpose of TOEFL iBT® is to decide on the admission of applicants who are non-native speakers of English for higher education institutions in North America (as cited in Alderson, 2009). Mainly, the applicants for a higher education institution in English-speaking countries are required to submit the TOEFL iBT® score to get admitted.</p> <p>Since all applicants are expected to have proven English proficiency to ensure their success in academic settings, and the TOEFL iBT® represents applicants' reading, listening, speaking and writing ability needed for their academic work, each admission administrator of college or university requires the minimum score of TOEFL. Generally, in case of a master in TEFL/TESL the applicants required minimum score is 79 to 100.</p> <p>In addition, the TOEFL iBT® score is used for English-language learning program admissions, and scholarship and certification candidates, too. In Korea, it is used as supplementary material for university entrance, and selecting scholarship students.</p>
Structure	TOEFL iBT is made up of four sections: reading, listening, speaking and writing. It takes about four and a half hours to complete. The reading section

	<p>has 3-4 passages and each passage has about 700 words with 12-14 questions. Students are allotted 60-80 minutes for the reading part.</p> <p>The reading section is designed to measure students' comprehension based on finding main ideas, inferring, vocabulary; test item types are multiple-choice, sentence insertion, and summarizing and sorting information using a given chart. The reading texts show a variety of topics from humanities to hard science. Therefore, students might read a topic that they do not have any topical knowledge about.</p> <p>The listening section consists of 2-3 conversations with 5 questions per conversation and 4-6 lectures with 6 questions per lecture. Test takers will hear the conversations and lectures only once, but they are allowed to take notes. A time of 60-90 minutes is allowed for the listening. The listening part is designed to measure test takers' ability to understand spoken English in academic and university-related life settings. The test items are multiple-choice, sequencing, categorizing.</p> <p>The speaking section deals with a variety of topics and includes 6 tasks: 2 independent and 4 integrated tasks (reading + speaking, listening + speaking). Specifically, test takers are required to talk about familiar topics, to talk about campus situation and to summarize and synthesize information after reading and listening passages, and they have 20 minutes for this section. Test takers respond via a computer and their responses are recorded and sent to raters.</p> <p>The writing section consists of two parts: independent task and integrated task (reading + listening + writing). Test taker are required to type their answers on the computer monitor. Test takers are asked to write an essay expressing their own opinion based on their knowledge and experience on the given topic for 30 min. with 150-225 words in length in the first task. In the second task, after reading a text and listening another text related to the reading, they are required to summarize the listening passage and explain how the summary are related to the reading passage for 20 min. with 150 and 225 words in length.</p>
Scoring	<p>Test takers have a separate score (0 to 30) in each section of four language skills and a total score of 0 to 120. Each task in the speaking section is rated on a holistic 0-4 point scale on the four criteria: general, delivery, language use, topic development and then the average score in each task is changed to a scaled score of 0 to 30.</p> <p>Similarly, each task in the writing section is rated on a holistic 0-5 scale. The independent writing task is scored on development, organization, appropriate and accurate use of grammar and vocabulary and the integrated task adds the completeness and accuracy of the content. Then the average score in two tasks is changed to a scaled score of 0 to 30.</p> <p>Test takers get the information about their scaled score in each section, total score, feedback about their performance, level according to scaled score range: <i>high, intermediate, low</i> in case of reading and listening, <i>good, fair, limited, weak</i> in case of speaking, <i>good, fair, limited, score of zero</i> in case of writing on the score report. In addition, scores are posted online about 10 days after the test date. Scaled scores are assigned in each section as follows:</p>

	Section	Scaled Score Range	Level																																				
	Reading	0-30	High (22-30) Intermediate (15-21) Low (0-14)																																				
	Listening	0-30	High (22-30) Intermediate (15-21) Low (0-14)																																				
	Speaking	0-30	Good (26-30) Fair (18-25) Limited (10-17) Weak (0-9)																																				
	Writing	0-30	Good (24-30) Fair (17-23) Limited (1-16) Score of Zero (0)																																				
	Total Score	0-120																																					
Statistical distribution of scores	<p>The following indicates means and standard deviations of section and total scores by gender for examinees tested between January and December 2011.</p> <table border="1"> <thead> <tr> <th>Group</th> <th></th> <th>Reading</th> <th>Listening</th> <th>Speaking</th> <th>Writing</th> <th>Total</th> </tr> </thead> <tbody> <tr> <td rowspan="2">Male</td> <td>Mean</td> <td>20.4</td> <td>20.0</td> <td>20.0</td> <td>20.9</td> <td>81</td> </tr> <tr> <td>SD</td> <td>6.9</td> <td>6.8</td> <td>4.6</td> <td>5.1</td> <td>21</td> </tr> <tr> <td rowspan="2">Female</td> <td>Mean</td> <td>19.9</td> <td>20.1</td> <td>20.8</td> <td>21.1</td> <td>82</td> </tr> <tr> <td>SD</td> <td>6.6</td> <td>6.6</td> <td>4.6</td> <td>4.9</td> <td>20</td> </tr> </tbody> </table> <p style="text-align: right;">(ETS, 2012a)</p>						Group		Reading	Listening	Speaking	Writing	Total	Male	Mean	20.4	20.0	20.0	20.9	81	SD	6.9	6.8	4.6	5.1	21	Female	Mean	19.9	20.1	20.8	21.1	82	SD	6.6	6.6	4.6	4.9	20
Group		Reading	Listening	Speaking	Writing	Total																																	
Male	Mean	20.4	20.0	20.0	20.9	81																																	
	SD	6.9	6.8	4.6	5.1	21																																	
Female	Mean	19.9	20.1	20.8	21.1	82																																	
	SD	6.6	6.6	4.6	4.9	20																																	
Standard error of measurement	<p>The following presents standard errors of measurement based on operational data from 2007 (ETS, 2011).</p> <table border="1"> <thead> <tr> <th>Score</th> <th>SEM</th> </tr> </thead> <tbody> <tr> <td>Reading</td> <td>3.35</td> </tr> <tr> <td>Listening</td> <td>3.20</td> </tr> <tr> <td>Speaking</td> <td>1.62</td> </tr> <tr> <td>Writing</td> <td>2.76</td> </tr> <tr> <td>Total</td> <td>5.64</td> </tr> </tbody> </table> <p>The SEM indicates “the degree of uncertainty associated with a test score” (ETS, 2005, p.2). In other words, SEM indicates how close a test taker’s score is to the test taker’s true ability score (Lawrence, 2011). Therefore, we might say SEM is related to the reliability of the test. The SEM assumes that any test score contains error and shows why a test score should be interpreted as a band of scores. The larger the standard error, the wider the band score, which the score might be interpreted as being less reliable (Miller et al. 2008).</p>						Score	SEM	Reading	3.35	Listening	3.20	Speaking	1.62	Writing	2.76	Total	5.64																					
Score	SEM																																						
Reading	3.35																																						
Listening	3.20																																						
Speaking	1.62																																						
Writing	2.76																																						
Total	5.64																																						
Evidence of reliability	<p>The following table presents reliability based on operational data from 2007 (Lawrence, 2011). It is believed that the TOEFL iBT® has high reliability.</p> <table border="1"> <thead> <tr> <th>Score</th> <th>Reliability</th> </tr> </thead> <tbody> <tr> <td>Reading</td> <td>0.85</td> </tr> <tr> <td>Listening</td> <td>0.85</td> </tr> </tbody> </table>						Score	Reliability	Reading	0.85	Listening	0.85																											
Score	Reliability																																						
Reading	0.85																																						
Listening	0.85																																						

	Speaking	0.88	As to test-retest reliability, it showed fairly high reliability: Reading 0.78, Listening 0.77, Speaking 0.84, Writing 0.77, Total 0.91 (Zhang, 2008, as cited in Alderson, 2008).
	Writing	0.74	
	Total	0.94	
Evidence of validity	<p>Lawrence (2008) argues the TOEFL iBT® is valid enough by stating the relevance and representativeness of test content; the appropriateness of task design and scoring rubrics; the relationship to academic language proficiency of the linguistic knowledge, processes and strategies that test takers use to respond to test tasks; the relationship between the test structure and theoretical views of the relationships among English language skills; and the relationship between TOEFL iBT scores and other criteria of language proficiency; test use and consequences (pp.4-10).</p> <p>In sum, the TOEFL iBT shows that it has construct validity because it measures test taker’s academic English proficiency through the integrative skills which they need in real university education settings, the use of long written and spoken texts which are regarded as more authentic in academic settings.</p>		



## International English Language Testing System (IELTS)

<i>Publisher:</i>	IELTS is managed jointly by University of Cambridge ESOL Examinations, British Council (UK) and IDP: IELTS Australia. Subject Manager, University of Cambridge ESOL Examinations, 1 Hills Road, Cambridge CB1 2 EU UK; <a href="mailto:ielts@ucles.org.uk">ielts@ucles.org.uk</a> ; <a href="http://www.cambridgeesol.org/">http://www.cambridgeesol.org/</a> . British Council, Bridgewater House 58 Whitworth Street Manchester M16BB; telephone +44 (0)161 957 7755; <a href="mailto:ielts@britishcouncil.org">ielts@britishcouncil.org</a> ; <a href="http://www.britishcouncil.org">http://www.britishcouncil.org</a> . IDP:IELTS Australia <a href="mailto:ielts@idp.com">ielts@idp.com</a> ; <a href="http://www.idp.com/ielts.aspx">http://www.idp.com/ielts.aspx</a>
<i>Publication Date:</i>	1989
<i>Target Population:</i>	Those whose mother language is not English and who are planning to study at above undergraduate school, to seek professional registration, or those who are planning to migrate to an English speaking country (Australia, Canada, New Zealand, UK), and to train or study at below undergraduate level.
<i>Cost:</i>	Fees vary by country. In general US \$ 185 to Dec. 31, 2012 and US \$ 190 from Jan. 2013 (US \$ 192: 210,000 Won in Korean currency in case of Korea)

**Overview**

Cambridge ESOL is in charge of designing the test and evaluating the candidates, while British Council and the IDP administer the test all over the world. IELTS originates from the English Language Test Service (ELTS) in 1980's. IELTS is influenced by “communicative language learning” and “English for specific purposes” (IELTS, 2012a).

There are two distinctive characteristics of IELTS. One is that it is paper-based test though a computer based IELTS was piloted in 2005 and the other is that it has two different types. Specifically, it is made up of Reading and Listening with paper text based, a face-to-face Speaking and Writing with an essay and report. Applicants can choose one of the two tests: *Academic and General Training* according to their purpose.

IELTS has been mainly for people who want to study in the UK, Australia and New Zealand (O'Sullivan, 2005) but it is administrated all over the world. IELTS is taken by over 1.7 million test takers each year. IELTS tests are held in over 800 centres with tests up to four times a month in 135 countries. IELTS is currently accepted by organizations in 135 countries; over 3,000 universities and organizations in America accept IELTS. In Korea, IELTS is taken in four cities three or four times a month. A specific description of the test is provided (see Table 2).

Table 2

Specific description for IELTS (Adapted from IELTS: *Information for candidates*, 2012b)

Purpose	<p>The purpose of IELTS is to measure English proficiency of people whose first language is not English and who want to study and work in an English speaking country or to immigrate to countries like Australia, Canada, New Zealand and UK.</p> <p>IELTS has two formats of the test and test takers choose the format according to their own purpose; one is the Academic module and the other is</p>
---------	---

	<p>the General Training module. The first is for people who wish to study above undergraduate level and seek professional jobs and the latter is for people who wish to migrate to UK , Australia, etc., and who want to train or study at below undergraduate level.</p>
Structure	<p>Both modules include listening, reading, writing and speaking. Listening and speaking tests are the same for both modules and reading and writing tests are different for each module. However, the review covers only Academic module for preparing for MA in TEFL/TESL.</p> <p>IELTS takes 2 hours and 45 minutes. In Listening section, there are 40 questions and it takes 30 minutes. A variety of task types are used, such as multiple choice, matching, labeling, summary, sentence completion, and short answer. It consists of 4 sections.</p> <p>Section 1 is a conversation in daily life as in a car Insurance company, gym or travel agency, section 2 is a monologue in daily life, section 3 is a conversation in an educational or training context as in planning a research project, and section 4 is a monologue on an academic subject as in a university lecture. Applicants can hear each section only once and a variety of accents are used.</p> <p>In Reading section, there are 40 questions and it takes 60 minutes. Task types include multiple choice, true/false/not given, matching headings, completion and short answers. There are 3 sections and the total text length is 2,150-2,750 words. Each section has one long text which includes authentic academic topics of general interest taken from books and journals. To be specific, according to Clapham (2000) the main text sources are from UK and Australia, which are intended for a non-specialist audience (as cited in Chalhoub-Deville &amp; Turner, 2000). If texts have technical terms, then a simple definition of the term is provided.</p> <p>In Writing section, test takers are required to complete two tasks within 60 minutes. In Academic writing, test takers describe or explain the information with a graph or diagram with at least 150 words in length and write an essay in response to the given general issue with at least 250 words in length.</p> <p>Lastly, Speaking section consists of a 3-part face-to-face oral interview with an examiner; the interview is recorded. It takes 11-14 minutes. It is made up of introduction, general questions about a test taker, long turn about a particular topic, and discussion based on the previous topic.</p>
Scoring	<p>Each correct answer in Listening and Reading receives 1 point and scores out of 40 are converted to the IELTS 9-band scale in whole and half bands (1:Non user ~ 9:Expert User).</p> <p>The Writing and Speaking tasks are scored by those who are trained and certified. The writing tasks are scored based on 4 criteria: task achievement/response, coherence and cohesion, lexical resource, grammatical range and accuracy. Task 2 is worth twice as much as Task 1 and it is reported in whole and half bands (1:Non user ~ 9:Expert User).</p> <p>The speaking tasks are scored based on 4 criteria: fluency and coherence, lexical resource, grammatical range and accuracy, pronunciation and the score is reported in whole and half bands (1:Non user ~ 9:Expert User).</p>

<p>In addition, the overall band score is provided in whole and half bands (1:Non user ~ 9:Expert User), too. Test results are informed 13 days after the test date. The overall band score explanations are as follows:</p>		
BAND 9	EXPERT USER	Has fully operational command of the language; appropriate, accurate and fluent with complete understanding
BAND 8	VERY GOOD USER	Has fully operational command of the language with only occasional unsystematic inaccuracies and inappropriacies. Misunderstandings may occur in unfamiliar situations. Handles complex detailed argumentation well.
BAND 7	GOOD USER	Has operational command of the language, though with occasional inaccuracies, inappropriacies and misunderstandings in some situations. Generally handles complex language well and understand detailed reasoning.
BAND 6	COMPETENT USER	Has generally effective command of the language despite some inaccuracies, inappropriacies and misunderstandings. Can use and understand fairly complex language, particularly in familiar situations.
BAND 5	MODEST USER	Has partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field.
BAND 4	LIMITED USER	Basic competence is limited to familiar situations. Has frequent problems in understanding and expression. Is not able to use complex language.
BAND 3	EXTREMELY LIMITED USER	Conveys and understands only general meaning in very familiar situations. Frequent breakdowns in communication occur.
BAND 2	INTERMITTENT USER	No real communication is possible except for the most basic information using isolated words or short formulae in familiar situations and to meet immediate needs. Has great difficulty understand spoken and written English.
BAND 1	NON USER	Essentially has no ability to use the language beyond possibly a few isolated words.
BAND 0	DID NOT	No assessable information provided

	<table border="1" style="width: 100%; text-align: center;"> <tr> <td style="width: 30%;"></td> <td style="width: 30%;">ATTEMPT THE TEST</td> <td style="width: 40%;"></td> </tr> </table> <p style="text-align: right;">(IELTS, 2013)</p>		ATTEMPT THE TEST							
	ATTEMPT THE TEST									
<p>Statistical distribution of scores</p>	<p>The reliability of Listening and Reading tests is provided below. The reliability of the Writing and Speaking tests cannot be reported in the same manner as Reading and Listening because they are not item-based.</p> <table border="1" style="width: 100%; text-align: center;"> <thead> <tr> <th></th> <th>Mean</th> <th>SD</th> </tr> </thead> <tbody> <tr> <td>Listening</td> <td>6.1</td> <td>1.3</td> </tr> <tr> <td>Academic Reading</td> <td>5.9</td> <td>1.2</td> </tr> </tbody> </table> <p style="text-align: right;">(IELTS, 2011)</p>		Mean	SD	Listening	6.1	1.3	Academic Reading	5.9	1.2
	Mean	SD								
Listening	6.1	1.3								
Academic Reading	5.9	1.2								
<p>Standard error of measurement</p>	<table border="1" style="width: 100%; text-align: center;"> <thead> <tr> <th></th> <th>SEM</th> </tr> </thead> <tbody> <tr> <td>Listening</td> <td>0.390</td> </tr> <tr> <td>Academic Reading</td> <td>0.379</td> </tr> </tbody> </table> <p style="text-align: right;">(IELTS, 2011)</p>		SEM	Listening	0.390	Academic Reading	0.379			
	SEM									
Listening	0.390									
Academic Reading	0.379									
<p>Evidence of reliability</p>	<p>The reliability of Listening and Reading tests is provided using Cronbach's alpha. The average alpha was 0.91 for the Listening, 0.92 for General Reading was 0.91, and 0.90 for Academic Reading (IELTS, 2011). It is assumed that the reliability is meaningful and valuable because it measures the internal consistency of the 40 items based on sufficient test takers' responses.</p> <p>The writing and speaking performances are rated by professionally trained raters according to detailed descriptive criteria and rating scales in terms of rater reliability. Recent G-studies showed coefficients of of 0.83-0.86 for Speaking and 0.81-0.89 for Writing (Shaw, 2004; Taylor &amp; Jones, 2001 , as cited in IELTS 2011).</p> <p>As to overall band score, Feldt &amp; Brennan (1989) claims that the composite reliability estimate for the Academic module was 0.96 and produced a composite SEM of 0.22 based on data from 2009 (as cited in IELTS 2011). In addition, it is reported that the speech and writing samples are re-rated if there is an inconsistency in the profile of the scores and IELTS centers are monitored in terms of reliability (Chalhoub-Deville &amp; Turner, 2000).</p>									
<p>Evidence of validity</p>	<p>It is noted that IELTS tries to cover English as an international language, which does not belong to one certain community and has a variety of accents and forms. IELTS contains a range of native-speaker accents as in North America, Australia, New Zealand and UK in the listening test and all standard varieties of English through all parts of the test (IELTS, 2012b).</p> <p>In addition, IELTS deals with everyday English as well as academic English that future applicants will encounter. What is better, Academic English contains topic of general interest for a non-specialist audience. However, Chalhoub-Deville &amp; Turner (2000) argue that research is needed to examine the appropriateness of IELTS score in universities of North America because IELTS has been intended mainly for use in the UK and Australia. Woodgrow (2006) claim that IELTS has the predictive validity based on comparison of the IELTS score and academic performance in postgraduate academic settings</p>									

	as in GPA arguing that there was no significant relationship between TOEFL and GPA (p.58).
--	--

## Pearson Test of English (PTE) Academic

<i>Publisher:</i>	Pearson Language Tests, 80 Strand, London, WC2R 0RL, UK; telephone +44 (0) 845 543 0243; fax: +44 (0) 207 010 6611; E-mail: <a href="mailto:pltsupport@pearson.com">pltsupport@pearson.com</a> ; <a href="http://pearsonpte.com/PTEAcademic/Pages/home.aspx">http://pearsonpte.com/PTEAcademic/Pages/home.aspx</a>
<i>Publication Date:</i>	Oct. 2009
<i>Target Population:</i>	Those who are non-native English speakers over 16 years old and need to demonstrate their level of academic English proficiency
<i>Cost:</i>	Fees vary by country. From US \$ 170 (Sri Lanka) to US \$ 270 (Belgium) (US \$ 192: 210,000 Won in Korean currency in case of Korea)

**Overview**

PTE Academic is a computer-based academic English language proficiency test created and administrated by Pearson Language Tests. PTE Academic measures test taker's academic English proficiency for universities, higher education institution or organizations requiring academic-level in terms of four skills: reading, writing, listening and speaking. PTE Academic focuses on real-life academic English with a variety of accents which test takers will encounter. It is endorsed by the Graduate Management Admission Council and it is accepted for admissions by thousands of programs in UK, America and Australia. It is expected that PTE Academic will compete with the previous two major English proficiency tests: iBT TOEFL and IELTS. PTE Academic is available in 43 territories including China, India, the US, Japan, South Korea, Australia, the UK, Hong Kong, Taiwan and Canada. In case of Korea, the test is held only in Seoul A detailed description of PTE Academic is provided (see Table 3).

Table 3

A detailed description of PTE Academic (Adapted from PTE: *PTE Academic overview*, 2012b)

Purpose	PTE Academic test is an international English proficiency test for those whose mother language is not English. The test is designed to measure test takers English ability by testing the level of academic English which they will use in real-life settings: universities, academic institution, education department, or other organizations where English is used as a main language and requires academic-level English proficiency. PTE Academic is accepted for admissions by above undergraduate level popular as TOEFL iBT and IELTS now. In addition, a variety of MBA programs, English language and Pathway programs are starting to use PTE Academic score for admission.
Structure	PTE Academic is a computer-based test and it measures listening, reading, speaking and writing skills in academic English. The test is composed of 20 different types of task and some tasks are integrated as in speaking with reading. PTE Academic takes 3 hours and is divided into timed parts. At first, introduction time is given, which is not counted in a test. Speaking and Writing is allotted 77-93 minutes. As for the speaking test, the test items are read aloud (6-7 items), repeat sentence (10-12), describe

	<p>image(6-7), re-tell lecture(3-4) and answer short questions(10-12) and test takers speak into a microphone and their answer is recorded.</p> <p>We can see two test items in writing: summarize written text (2-3) and write essay(1-2). Reading is 32-41 minutes and the test items are multiple choice (4-6), re-order paragraphs(2-3), fill in the blanks(9-11). Listening is 45-57 minutes and it has a variety of test items, such as summarize spoken text (2-3), multiple choice (4-6), fill in the blanks(2-3), highlight correct summary(2-3), select missing word (2-3), highlight incorrect words (2-3), write from dictation(3-4). Test takers hear different varieties of English: British, American and Australia. (PTE, 2012a)</p>												
Scoring	<p>The written and spoken portions of PTE Academic are scored with automated scoring technology, so test scores are provided within 5 days. PTE Academic reports provide an overall score (0-90) based on the test taker's performance on all items in the test, communicative skills scores (listening, reading, speaking, writing) and enabling skills scores (grammar, oral fluency, pronunciation, spelling, vocabulary, written discourse).</p> <p>The score range for each skill is 10–90 points. For example, in the case of oral fluency, the range of score is from 0 (Disfluent) to 5 (Native-like). Generally, postgraduate studies require an overall score between 57-67. This is equivalent to 6.0-7.0 in IELTS and 85-98 in TOEFL iBT.</p> <p>In addition, PRE Academic has been aligned to the Common European Framework Level (CEFL), which is which is recognized as a standard across Europe. (PTE, 2012a)</p>												
Statistical distribution of scores	<p>NA (Though I requested the data, they told they are unable to provide the sensitive data.)</p>												
Standard error of measurement	<p>The SEM for PTE Academic is 2.32. This figure is based on test data from 30,000 test takers (PTE, 2012a).</p>												
Evidence of reliability	<p>It is noted that PTE Academic has high reliability (0.90 or higher). The table below is the reliability estimated for scores in the range of 53-79, which is the most relevant range for admission decisions.</p> <table border="1" data-bbox="410 1367 1414 1444"> <thead> <tr> <th>Score</th> <th>Overall</th> <th>Listening</th> <th>Reading</th> <th>Speaking</th> <th>Writing</th> </tr> </thead> <tbody> <tr> <td>Reliability</td> <td>0.97</td> <td>0.92</td> <td>0.91</td> <td>0.91</td> <td>0.91</td> </tr> </tbody> </table> <p style="text-align: right;">(PTE, 2012a)</p> <p>In addition, PTE Academic adopts an automated scoring system. According to the research about the overall reliability for both human marks and for machine marks, it shows 0.97, which is a very high coefficient and demonstrates the accuracy of both human and machine marking. (PTE, 2012c)</p>	Score	Overall	Listening	Reading	Speaking	Writing	Reliability	0.97	0.92	0.91	0.91	0.91
Score	Overall	Listening	Reading	Speaking	Writing								
Reliability	0.97	0.92	0.91	0.91	0.91								
Evidence of validity	<p>It is reported that PTE Academic has construct validity because the result based on the test specification, various item types, content of stimuli, task types, required competence, and response format corresponds with the test takers' future language use in an academic setting (Zheng &amp; De Jong, 2011).</p> <p>In addition, PTE Academic shows high correlation with the TOEFL iBT and IELTS based on the result about validity check and the data were collected from 364 test takers in 2009. All correlations are significant at <math>p &lt; .01^3</math>, and the</p>												

	number means that this concordance makes fairly accurate predictions. (PTE,2012a)
--	--



The context I consider is current and future English teachers in Korea who are preparing to take certified English proficiency tests, in order to study in TEFL/TESL MA programs in countries where English is used as the main instruction language. The majority of people who are current or future English teachers are highly motivated to study TEFL/TESL abroad to enhance their English proficiency, as well as teaching methods, if the circumstances permit.

It is assumed that the ages range from 23 to 45 years old. The general applicants' educational backgrounds are above undergraduate school, so they have at least a bachelor's degree from Korea. Generally, they have considerable English proficiency and those who have majored in English or English education have in-depth knowledge of English as well. However, since the Korean English education is test-oriented and focuses on reading, it is assumed that they have a relatively higher level of English in reading than in listening, speaking and writing. In addition, if they are in their twenties, they might have taken English proficiency tests such as the TOEFL iBT or IELTS, because certified English proficiency test scores were required for that age to enter a university or get a job as a supplementary document in Korea.

All three tests seem to have fairly high reliability with enough test items and validity in terms of interpretations and uses of assessment results. The cost of testing and availability of equivalent or comparable forms makes little difference. However, considering the target population and the results of the three test reviews, I think IELTS is the most appropriate for assessing their English proficiency to get admission to in TEFL/TESL MA programs provided in English speaking countries.

To begin with, though PTE Academic has the advantage that test takers will be notified of their results within 5 days and it shows high reliability, it is taken only in Seoul, so it is not easy to get access to the test in terms of usability. In addition, the popularity and recognition is

still less for test takers in Korea than for the TOEFL iBT and IELTS. Despite the increasing popularity of PTE Academic, the majority of MA TEFL/TESL programs still prefer the TOEFL iBT and IELTS. Above all, PTE Academic shows relatively insufficient research compared to the other two tests. It needs more research about the test itself and the results should be open to the public.

Both the TOEFL iBT and IELTS show fairly high reliability and validity with a variety of test items and there is a lot of research and information about these two tests. What is better, both of them provided enough information with researchers as well as test takers online. Therefore, anybody can get access to necessary data they want.

It is believed that the TOEFL iBT is one of the most popular English proficiency tests in Korea. However, despite the popularity of the TOEFL iBT, I think it has some shortcomings. First, the administration might be a big obstacle to the test takers who are not accustomed to the computer. To be specific, the people who are not good at typing English characters have difficulty in typing their essay in a writing section leading to poor scores compared to their true writing ability. In the case of reading, the test takers are required to pay more attention to view previous and next pages on the monitor because they have to use a mouse to move where they want to see. In contrast, IELTS is a paper-based test, so it is easier for test takers to see and read the text without any special manipulation.

Another big problem occurs in the speaking section. The test takers should interact with a computer in TOEFL and do not interact with an examiner as in IELTS. It is not authentic to respond to someone for 30 seconds without any interaction. Besides, the speaking topics in each part are not related to each other, so test takers should respond to the computer without any previous knowledge. In contrast, test takers have time to introduce themselves to the examiner,

and the speaking topics in each part are related to each other in IELTS. Therefore, I think IELTS can assess the test takers' speaking ability better than the TOEFL iBT.

In terms of content, the TOEFL iBT seems to focus too much on academic texts intended for a specialist audience. The texts often include professional and special topics, so those who have no topical knowledge about the given topic should struggle with it all the way through. For example, I had to struggle with the text about supernova, which I have not encountered in my actual academic setting. It can be an issue from the point of the construct validity view. The majority of the potential test takers for admission to TEFL/TESL MA programs might share my view. On the other hand, IELTS covers daily life English as well as academic English. As I wrote in the structure of IELTS, the texts include academic topics of general interest. Therefore, it can be a more useful and authentic reflection of real life for test takers. To be specific, international students should face a variety of situations such as reading a map, buying car insurance, as well as listening to lectures and taking notes in class.

Lastly, I would like to point out English as an International English. Unlike the TOEFL iBT, which mainly focuses on North American English, IELTS seems to include a variety of English native-speaker accents in the listening test. English is not restricted to North American English. Therefore, it is valuable to include various accents in certified English proficiency tests. In sum, I would conclude that IELTS is the most appropriate English proficiency test for those who are preparing for English proficiency test wishing to study abroad for a MA in TEFL/TESL.

## References

- Alderson, J.C. (2009). Test of English as a foreign language™: Internet-based Test (TOEFL iBT®). *Language Testing*, 26(4). 621-631.
- Cahlhoub-Deville, M. & Turner, C.E. (2000). What to look for in ESL admission tests: Cambridge certificate exams, IELTS, and TOEFL. *System* 28. 523-539.
- ETS. (2005). *Setting the final scores*. Retrieved from [http://www.ets.org/Media/Tests/TOEFL/pdf/setting\\_final\\_scores.pdf](http://www.ets.org/Media/Tests/TOEFL/pdf/setting_final_scores.pdf)
- ETS. (2012a). *Test and score data summary for TOEFL iBT® tests and TOEFL® PBT tests*. Retrieved from [http://www.ets.org/s/toefl/pdf/94227\\_unlweb.pdf](http://www.ets.org/s/toefl/pdf/94227_unlweb.pdf)
- ETS. (2012b). *The TOEFL iBT®*. Retrieved from <http://www.ets.org/toefl>
- IELTS. (2011). Test Performance. Retrieved from [http://www.ielts.org/researchers/analysis\\_of\\_test\\_data/test\\_performance\\_2011.aspx](http://www.ielts.org/researchers/analysis_of_test_data/test_performance_2011.aspx)
- IELTS (2012a). History of IELTS. Retrieved from [http://www.ielts.org/researchers/history\\_of\\_ielts.aspx](http://www.ielts.org/researchers/history_of_ielts.aspx)
- IELTS (2012b). Information for candidates. Retrieved from [http://www.ielts.org/pdf/Information\\_for\\_Candidates\\_booklet.pdf](http://www.ielts.org/pdf/Information_for_Candidates_booklet.pdf)
- IELTS (2013). IELTS guide for institutions and organization. Retrieved from [http://www.ielts.org/institutions/about\\_ielts/what\\_is\\_ielts.aspx](http://www.ielts.org/institutions/about_ielts/what_is_ielts.aspx)
- Lawrence, I. (2008). Validity evidence supporting the interpretation and use of TOEFL iBT™ Scores. *TOEFL iBT Research Insight*, 4, 1-16. Retrieved from [http://www.ets.org/s/toefl/pdf/toefl\\_ibt\\_insight\\_s1v4.pdf](http://www.ets.org/s/toefl/pdf/toefl_ibt_insight_s1v4.pdf)
- Lawrence, I. (2011). Reliability and comparability of TOEFL iBT™ scores. *TOEFL iBT Research Insight*, 3, 1-7. Retrieved from [http://www.ets.org/s/toefl/pdf/toefl\\_ibt\\_research\\_s1v3.pdf](http://www.ets.org/s/toefl/pdf/toefl_ibt_research_s1v3.pdf)

Miller, M. D., Linn, R., & Gronlund, N. (2008). *Measurement and assessment in teaching* (10th ed.). Upper Saddle River, N.J.: Prentice Hall.

O'Sullivan, B. (2005). International English Language System (IELTS). In S.Stoynoff & C.Chapelle, *ESOL Tests and Testing* (pp.73-38). Alexandria, VA;TESOL.

PTE (2012a). *PTE Academic score guide*. Retrieved from

[http://pearsonpte.com/PTEAcademic/scores/Documents/PTEA\\_Score\\_Guide.pdf](http://pearsonpte.com/PTEAcademic/scores/Documents/PTEA_Score_Guide.pdf)

PTE (2012b). *PTE Academic overview*. Retrieved from

<http://pearsonpte.com/PTEAcademic/Pages/TestOverview.aspx>

PTE (2012c). *Validity and reliability in PTE Academic*. Retrieved from

<http://pearsonpte.com/research/Pages/ValidityandReliability.aspx>

Woodgrow, L. (2006). Academic success of international postgraduate education students and the role of English proficiency. *University of Sydney papers in TESOL, 1*, 51-70.

Zheng, Y. & De Jong, J,H,A,L. (2011). Establishing construct and concurrent validity of Pearson Test of English Academic. *Research note*. Retrieved from

[http://pearsonpte.com/research/Documents/RN\\_EstablishingConstructAndConcurrentValidityOfPTEAcademic\\_2011.pdf](http://pearsonpte.com/research/Documents/RN_EstablishingConstructAndConcurrentValidityOfPTEAcademic_2011.pdf)