THE EFFECTS OF TOEIC EDUCATION IN SOUTH KOREAN UNIVERSITIES

by

STEPHAN THOMSON

A dissertation submitted to the
College of Arts and Law
of the University of Birmingham
in part fulfilment of the requirements
for the degree of

Master of Arts
in
Teaching English as a Foreign or Second Language (TEFL/TESL)

This dissertation consists of approximately 13,000 words
Supervisor: Douglas Sewell

Centre for English Language Studies
College of Arts & Law
University of Birmingham
Edgbaston, Birmingham B15 2TT
United Kingdom

September 2012
ABSTRACT

This paper considers the effects studying the Test of English for International Communication (TOEIC) has on South Korean university students. It examines: the washback effects in the classroom; how studying TOEIC affects motivation to learn English and; the resulting validity issues. After presenting background information on language testing, this paper examines the issues involved through mixed-methods research. While more research is needed on the topic, this paper suggests that the concentration in South Korean universities on the Listening and Reading sections of the TOEIC lowers student motivation to learn English, and lowers the validity of the TOEIC.
ACKNOWLEDGEMENTS

I would like to dedicate this work to my beautiful wife Mina. This project would have been impossible without her support and calming influence.

I would also to thank Andy Shin for her help with translations, and all of my proofreaders for helping to minimize my language mistakes. I would also like to thank my course tutor Douglas Jarrell for getting me to this point, and my thesis advisor Douglas Sewell for his helpful advice and guidance.

Stephan Thomson

September 2012
**CONTENTS**

1. Introduction ................................................................. 1

2. Language Testing .......................................................... 3
   2.1. Types of Language Testing ............................................ 3
   2.2. Reliability ............................................................ 5
      2.2.1. Test Reliability .................................................. 6
      2.2.2. Test-taker Reliability .......................................... 8
   2.3. Validity .............................................................. 9
      2.3.1. Content Validity ................................................ 9
      2.3.2. Criterion-Related Validity ..................................... 11
      2.3.3. Construct Validity .............................................. 12
      2.3.4. Face Validity .................................................. 12
   2.4. Authenticity and Interactiveness ................................ 13
   2.5. Impact ............................................................... 14
      2.5.1. Test Impact ..................................................... 14
      2.5.2. Washback ....................................................... 15

3. Sociological Factors Affecting Language Testing in South Korea ......... 17

4. The TOEIC Test ............................................................ 19
   4.1. TOEIC Usage in South Korea ........................................ 19

5. Research Methodology ..................................................... 22
   5.1. Research Instruments ............................................... 22
      5.1.1. Quantitative Survey ............................................. 22
      5.1.2. Qualitative Survey ............................................. 23
      5.1.3. Classroom Observations ....................................... 24
   5.2. Participants .......................................................... 25
   5.3. Administration ....................................................... 26

6. Results ................................................................. 27
   6.1. Background on Student Knowledge of the TOEIC ................. 27
   6.2. Motivational Influences of Studying for the TOEIC ............ 29
   6.3. Washback Effects of the TOEIC in South Korean University Classrooms ....... 32
      6.3.1. Classroom Observations ....................................... 35

7. Analysis of Findings ...................................................... 37
   7.1. Findings on Motivation ............................................. 38
1. Introduction

An imperative of modern life is to be able to communicate effectively with those from other countries and cultures. One of the results this creates is the desire to learn other languages. Concomitant with this is language testing. This can be in the form of student achievement assessments; diagnostic tests to find out what students need to learn; aptitude tests to judge learner abilities to learn certain skills; and proficiency tests to analyze a person’s current language ability.

This paper is concerned with the Test of English for International Communications (TOEIC), which is one of the most widely used English language proficiency tests in Japan and South Korea, and is utilized to a lesser degree in a variety of countries worldwide (ETS 2005: 4). The TOEIC was designed to judge the proficiency of language learners in a business environment. It is claimed to be applied “as a benchmark for employment, promotion, overseas business travel, and transfer in workplaces” (In’nami and Koizumi 2011: 131).

Educational Testing Service (ETS) designed the TOEIC on commission from the Japanese government to aid decision-making about Japanese workers (Stoyoff 2009: 28). It is widely used there, and has become popular with South Korean companies, universities and other organizations. Due to its popularity, a range of test-preparatory education has grown in both countries.

In South Korea, numerous hagwons (private institutions) offer TOEIC test-preparation courses for businesspeople, students and graduates seeking to increase their employability. In addition TOEIC test-preparation or incorporation of TOEIC test scores is often a part of a university’s curriculum. Colleges and universities possibly include the TOEIC or TOEIC preparation classes for a multitude of reasons such as; advertising that the education provided takes globalization into account; as a tool for furthering comprehension of English; or as a method to improve student employability. Given the widespread use of the TOEIC in South Korea, this paper will examine two pressing concerns of how studying for the TOEIC in South Korean universities affects student motivation to learn English, and whether or not this test-preparation affects test validity.

In order to understand TOEIC impact on Korean university student performance, this paper begins with a review of current language testing research. It will then examine the
educational culture and sociological factors that affect language testing in South Korea. This will be followed by an overview of the TOEIC test.

Following the literature review is the research methodology conducted to judge the impact of the TOEIC. The purpose of the research was to determine if the TOEIC test has any washback effects in the South Korean university context. Mixed-methods research was conducted at a four year South Korean university. A quantitative survey was given to 104 students in their required first year of TOEIC study, together with 23 students studying an optional second year. In addition, a qualitative survey was administered to a separate group of 17 second year students who had completed their first year of TOEIC instruction.

Teachers play a large role in test-preparation classes. A series of class observations were carried out to establish variances in the ways in which individual teachers approach TOEIC instruction. This permitted an examination of the impact of the TOEIC test on their teaching habits. Finally, this paper will analyze the data, and the implications for the TOEIC test in the South Korean university context.
2. LANGUAGE TESTING

Language testing is complex, often making it difficult to create successful tests. This difficulty can result in some tests being viewed as better than others. However, good or bad tests do not exist “in the abstract, and… there is no such thing as the one ‘best’ test, even for a specific situation” (Bachman and Palmer 1996: 6, italics in original). It thus falls on test creators to consider a variety of factors while preparing a test.

Language testing is primarily used by institutions to fulfill educational or social policy objectives (McNamara 2000: 67). Tests are created by these establishments, and are used to determine any cognitive abilities the organisation feels are necessary for their goals. The following section will examine the various types of tests these institutions create.

2.1. TYPES OF LANGUAGE TESTING

Two major factors test designers must consider are validity and reliability of results (Hughes 2003: 59). Two major ways of doing this are ‘norm-referencing’ and ‘criterion-referencing’. Norm-referenced tests, such as the TOEIC, provide a numerical score showing students relations to one another. Criterion-referenced tests rate the specific tasks a test-taker is capable of doing, regardless of the performance of others. These two scoring types can be combined. People examining norm-referenced test scores can apply criteria they believe scores define. However, this does not make them criterion-referenced (Hughes 2003: 21).

In addition to being norm-referenced, the TOEIC test is standardized. The purpose of standardization is to ensure that “each individual experiences essentially the same test and context of administration” (Fulcher and Davidson 2007: 198). In each testing session, every test-taker is administered an identical test. Standardization should result in increased reliability, which will be discussed in the next section.

Another variable test designers must consider is whether to test directly or indirectly. Direct testing “requires the candidate to perform precisely the skill that we wish to measure” (Hughes 2003: 17). Therefore, a direct test designed to judge a candidate’s speaking ability will require them to speak.

Indirect testing measures “the abilities that underlie the skills in which we are interested” (Hughes 2003: 18). An example of indirect testing is using the TOEIC Listening and Reading (L/R) test to make judgments about a candidates speaking or writing abilities. Hughes (ibid) argues that:
“… the main problem with indirect tests is that the relationship between performance on them and performance of the skills in which we are usually more interested tends to be rather weak in strength and uncertain in nature.”

The decision to test directly or indirectly often depends on the type of test employed. As mentioned previously, four commonly used types of language tests are; achievement, proficiency, aptitude and diagnostic tests (Hughes 2003: 11-16). TOEIC is a norm-based proficiency test. Proficiency tests are “based on a specification of what candidates have to be able to do in the language in order to be considered proficient” (Hughes 2003: 11). These tests are usually classified as high-stakes, because the outcomes typically determine whether or not test-takers are given opportunities such as the chance to study in a foreign country or get a promotion at work. Since a high TOEIC score is necessary for achieving many goals in South Korea, as will be discussed further in section 3.1, this paper classifies it as high-stakes.

The three main types of language tests can be broken down based on the types of questions they employ. The most commonly used types of questions are:

1. Selected-response (multiple choice, true or false)
2. Constructed-response (short answer, cloze tests)
3. Personal-response (interviews, essays, portfolios)

(adapted from Brown and Hudson 1998: 653)

Each of these forms of assessment has advantages and disadvantages. Selected-response assessments are easy to administer, score, and are objectively assessed. These tests are often scored by computer. Selected-response questions are typically used for norm-referenced tests. The TOEIC L/R test is an example of a selected-response test.

There are two main disadvantages to selected-response tests. They are difficult to create, and they do not require test-takers to use any productive language skills (Brown and Hudson 1998: 658). A student could perform well on these tests, without having the ability to use the language in spoken or written communication. This was shown in a study performed by the Korean Ministry of Education, which found that test-takers with high scores on the TOEIC L/R test showed weakness in speaking and writing skills (Ko 2005). Selected-response test detractors often bring up this issue, because, as Brown and Hudson (1998: 659) point out: “Real-life language is not multiple-choice.” A final issue with this format is that test-takers
will get a percentage of answers correct by guesswork or luck. This percentage will naturally vary depending on the number of possible answers (ibid).

Unlike selected-response assessment, constructed-response tests require test-takers to use language in a limited format. This takes most of the guesswork out of answering questions, and forces students to interact with the language to a larger extent. These questions can be used in both norm-referenced and criterion-referenced tests.

Personal-response assessments, “require students to actually produce language [and]... allow each student’s responses to be quite different” (Brown and Hudson 1998: 663). Test-takers can answer questions in their own words instead of being forced to use provided constructs. The advantages and disadvantages of these assessments are the same as constructed response tests. Test-takers are more fully interacting with the language, and the subjectivity of grading and issues with inter- and intra-scorer reliability increase with the length of responses.

Educators and test administrators must determine the usefulness of a given test regardless of the type of test or questions employed. Bachman and Palmer (1996: 18) use the following equation to describe a test’s usefulness:

\[
\text{Usefulness} = \text{Reliability} + \text{Construct Validity} + \\
\text{Authenticity} + \text{Interactiveness} + \text{Impact} + \text{Practicality}
\]

Figure 2.1 – Test Usefulness

Unlike practicality, which describes how a test will be implemented, ‘usefulness’ is a combination of competing factors which need to be balanced according to the purpose of a given test (ibid). The following sections will examine factors affecting test usefulness. The first to be examined is reliability.

**2.2. RELIABILITY**

Bachman and Palmer (1996: 19) define reliability as the “consistency of measurement”. This can be seen as the extent to which the same candidate would achieve the same score on the same test at different testing opportunities (Hughes 2003: 36). Numerous factors affect test reliability. The goal is to create assessments that provide nearly identical scores for the test-takers, if they take multiple versions of the same test. “If a test does not measure something consistently, it follows that it cannot always be measuring accurately” (Alderson et.al 1995: 187). The facets of tests and test-takers that cause scores to vary are shown in Figure 2.2.
The most important factor for test-takers noted by Bachman is knowledge of the test and language being tested. This can include sociolinguistic competence, communicative language ability and knowledge of how to apply these skills on certain types of tests (Bachman 1990: 163). For the TOEIC L/R test, test-takers with knowledge of the layout of the test, and the types of questions asked, will have a greater knowledge of the test, and should perform better provided other factors are constant.

If a test is well constructed and administered, changes in ability should result in relevant changes in test scores. The goal of test reliability is to reduce the other issues in figure 2.2 affecting reliability to correctly gauge changes in the learner’s actual language ability (Bachman 1990: 166). The only factor that test-creators are able to control is ‘test method facets’. This will be discussed in the next section.

**2.2.1. Test Reliability**

Test reliability can be seen in figure 2.2 as the factors shaping ‘test method facets’. Test-creators often choose selected-response assessments such as the TOEIC L/R test because “scoring can be perfectly reliable” (Hughes 2003: 76). The test can be easily administered because students only need mark a test sheet to record an answer (ibid).

In constructed-response and personal-response assessments, the subjectivity of grading and rater-reliability can affect overall test reliability. The problem is that “if some raters rate more severely than others, then the ratings of different raters are not consistent, and the scores obtained could not be considered to be reliable.” (Bachman and Palmer 1996: 20) The test-creators must come up with grading strategies such as rubrics to improve reliability.
One way to quantify a test’s reliability is through the use of reliability coefficients (Hughes 2003: 38). This is a scale of 0 to 1, with 1 indicating that a test is completely reliable, and 0 being completely unreliable. A test with a reliability coefficient of 1 would result in test-takers obtaining the exact same score on a test regardless of when or where the test was taken. Conversely, a test with a reliability coefficient of 0 would have completely different results at different testing opportunities (Hughes 2003: 39). Research suggests that different reliability coefficients should be expected for different types of language tests. The ranges to be expected are shown in table 2.3.

<table>
<thead>
<tr>
<th>Type of test</th>
<th>Expected reliability coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading / vocabulary test</td>
<td>.90 - .99</td>
</tr>
<tr>
<td>Listening test</td>
<td>.80 - .89</td>
</tr>
<tr>
<td>Speaking / writing test</td>
<td>.70 - .79</td>
</tr>
</tbody>
</table>

Table 2.3 – expected reliability coefficients (Lado cited in Hughes 2003: 39)

Research performed by ETS on the TOEIC test showed that the reliability coefficient of both the listening and reading sections of the TOEIC test are in the range .92 - .93 (Liao, Hatrak and Yu 2010: 4).

In order to calculate a test’s reliability coefficient, there must first be two sets of scores to compare (Hughes 2003: 39). The three most commonly used methods to acquire two sets of scores are: 1. the test-retest method; 2. the alternate forms method and; 3. the split half method. The test-retest method administers the same test twice to a group of test-takers. The disadvantage of this method is that test-takers are likely to remember their responses to questions, artificially raising the reliability (ibid). The alternate forms method administers two different forms of the test to a group of test-takers, one on each testing occasion. However, alternate forms are not always available (Hughes 2003: 40). The most commonly used method of obtaining the required two scores is the split half method (ibid). For this, a single test is administered once, but scores are recorded for each half of the test. These two scores are then used as if the test was taken twice (ibid).

Once the reliability coefficient of a test has been calculated, further calculations can be performed to determine the standard error of measurement (SEM). The SEM “is a measure of the tendency of test-takers’ scores to vary because of random factors” (ibid). The SEM is inversely related to the reliability coefficient. The lower the reliability coefficient, the higher the SEM will be, and vice-versa. This is then used to predict a test-taker’s true score. True
score is the hypothetical average of test-taker scores if it were possible to take a test an infinite number of times (Hughes 2003: 40). The estimated SEM for the TOEIC L/R test is 25 score points per section (Liao, Hattrak and Yu 2010: 4). This allows us to make the following predictions through the application of statistics for a hypothetical test-taker that sits the TOEIC only once with a score of 300 on each section:

- The test-taker’s true score is 68% certain to be in the range of 275 – 325 (within one SEM) for each section, with a total score of 550 – 650.
- The test-taker’s true score is 95% certain to be in the range of 250 – 350 (within two SEM) for each section, with a total score of 500 – 700.
- The test-taker’s true score is 99.7% certain to be in the range of 225 – 375 (within three SEM) for each section, with a total score of 450 – 750.

(adapted from Hughes 2003: 41)

This means that it is reasonable to expect this hypothetical test-taker’s true score to be between 500 and 700 points. There is too much statistical possibility for error if the true score is only calculated within 1 SEM (ibid).

The remaining factors affecting reliability as noted in figure 2.2 are largely outside the control of the test-creators and administrators. The following section will deal with these aspects.

2.2.2. TEST-TAKER RELIABILITY

The random factors noted in figure 2.2 can have negative effects on test-taker scores. Issues such as poor health, fatigue, confidence, motivation, topic interest and stress can all impact test scores (Everson and Hines 2010: 2). Of these issues, stress is one of the most important. “Any learner who is obliged to do something under pressure will perform abnormally” (Alderson and Wall 1993: 117). This is especially true for high-stakes tests such as the TOEIC, given that results could impact their life goals and future plans. For example, at the South Korean chemical company Miwon Commercial Co. Ltd., a TOEIC L/R combined score of 900 is needed among other requirements for promotion from chemical researcher to team leader (Lee 2012). Being required to obtain such a score can cause high stress on test day, making it more difficult for test-takers to perform to their abilities.

Reliability is an important factor in language testing, but it must be balanced with validity. Alderson et.al (1995: 187) assert that test-creators “maximize one at the expense of the other.
Which you choose to maximize will depend upon the test’s purpose and individual consequences of gaining an inaccurate result.” For the TOEIC L/R test, research conducted by In’nami and Koizumi (2011: 148) found that the validity of the test is “relatively weak”. This finding may be a result of the TOEIC test-designers focus on reliability, having a negative effect on validity. The following section deals with validity and its relation to reliability.

2.3. VALIDITY

A language test can be said to be valid if it actually tests the skills it claims to examine. Valid tests are able to “ensure the defensibility and fairness of interpretations based on test performance” (McNamara 2000: 48). Validity is often forced to trade off with reliability in the form of questions it employs. The American Psychological Association argues that:

“no language test can use only one response format, or sample only one construct, if it is to be considered valid on even prima-facie grounds” (cited in Fulcher 1997: 114).

This supports the findings of In’nami and Koizumi noted above. The TOEIC L/R test uses selected-response questions in order to raise reliability, however since only one type of response is required of test-takers, the trade-off is a reduction in test validity.

Jafarpur (1987: 202) notes four basic ways of analysing test validity, being content validity; criterion-related validity; construct validity and face validity. The following sections will deal with each of these facets individually.

2.3.1. CONTENT VALIDITY

Test development begins with the development of specifications. A test is said to have content validity if it meets specifications such as language skills or structures that it is meant to be concerned with (Hughes 2003: 26). Content validity “involves gathering the judgment of ‘experts’: people whose judgment one is prepared to trust, even if it disagrees with one’s own” (Alderson et.al. 1995: 173). These opinions derive from current language acquisitions and testing theories and are used by test administrators and designers to analyse test content compared to relevant language skills (ibid). This analysis will determine whether or not the test is actually assessing the skills it claims.

Tests fail content validity examinations when areas noted in test specifications are under-represented (Hughes 2003: 27). An oral test designed to examine communicative
competence that only deals with questions on one topic may allow a test-taker to discuss that
topic and achieve a high score. However, inferences cannot be made about their abilities to
talk about any topic in the target language. In order to increase content validity, a test should
test widely and randomly from the test specifications (ibid). After test-creators have
produced a test considered to be content valid, it must be examined in relation to other forms
of assessment to judge its criterion-related validity.

2.3.2. CRITERION-RELATED VALIDITY

Criterion-related validity is “the degree to which the results on the test agree with those
provided by some independent and highly dependable assessment of the candidate’s ability”
(Hughes 2003: 27). This can be further broken down into concurrent validity and predictive
validity. Concurrent validity is when “the test and the criterion are administered at about the
same time” (ibid). For proficiency tests, concurrent validity is commonly carried out by
comparing the results to another commonly used proficiency test claiming to test the same
structures.

Two scores can be compared to create a correlation coefficient, also known as a validity
coefficient, which is similar in structure to the reliability coefficient discussed in section 2.2.1.
(ibid). Two tests that have a validity coefficient of 1 are in perfect agreement with each other,
and can be said to be testing the same thing. A validity coefficient of 0 indicates no
correlation, and that one or both of the tests are not examining the language processes they
claim to be testing. Research by Liao, Qu and Morgan (2010) examined the correlation
between the different sections of the TOEIC test. The lowest correlation of test-takers that
took all sections of the test within one month was .55 between the reading and speaking
sections. The highest correlation of .65 was between the listening and speaking sections.
This indicates that the different sections of the TOEIC “measure four different aspects of
English proficiency” (Liao, Qu and Morgan 2010: 5-6, emphasis added).

Predictive validity is concerned with “the degree to which a test can predict candidates’
future performance” (Hughes 2003: 29). For the TOEIC test, the predictive validity indicates
how well a test-taker would be able to communicate in an English language business
environment. However, since a number of other factors such as motivation, health and
subject knowledge have an effect on performance, a validity coefficient of .4 is as high as can
be expected (Hughes 2003: 30).
Validation involves factors other than test content. Test-creators must look at what results actually mean (Alderson et.al. 1995: 171). This is one part of construct validity, and will be discussed in the following section.

2.3.3. **Construct Validity**
A construct is defined by Bachman and Palmer (1996: 21) as:

“… the specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from the task.”

Tests may vary depending on how the constructs are characterized. This may have an impact on how the scores are interpreted.

Construct validity is defined by Dörnyei (2007: 51) as:

“… a property of the conclusions, interpretations or inferences that we draw from the assessment instruments and procedures, not the instruments and procedures themselves.”

In Dörnyei’s definition, construct validity can be seen as an all-encompassing term. This includes the interpretation of scores and “subsumes specific test validation operations” (Smith cited in Dörnyei 2007: 51).

Tests may fail construct validity examinations if they are used for purposes other than those for which they were originally designed. “A test that proves ideal for one purpose may be quite useless for another; a technique that may work very well in one situation can be entirely inappropriate in another” (Hughes 2003: 8). In the case of the TOEIC, since it was designed to judge proficiency in a business environment, it may be an unacceptable test to judge a test-taker’s ability to use English in an academic setting.

McNamara notes that, “test score inferences need to be revalidated for every major context of use” (cited in Dörnyei 2007: 52). If an organization wishes to use the TOEIC to judge proficiency in situations other than business, it will need to be revalidated for the new situation.

2.3.4. **Face Validity**
Face Validity “refers to the degree to which the test seems valid, in the eyes of those involved in taking or administering it.” (Küçük and Walters 2009: 332). A test must appear to be
testing what it claims to be. Otherwise, the results may be viewed as irrelevant to the test-takers. This may have an impact on motivation. In situations where test-takers view a test as being valid, it can be argued that “they are more likely to perform to the best of their ability on that test and to respond appropriately to items” (Alderson et.al. 1995: 173).

**2.4. Authenticity and Interactiveness**

Test authenticity can be defined as how well language test tasks can be related to actual language use (Bachman and Palmer 1996: 23). This is closely related to construct validity. While construct validity is concerned with the interpretations of the scores, authenticity is concerned with how a test is able to correspond “to language use in specific domains other than the language test itself” (ibid). As with validity, there is a trade-off between authenticity and reliability. In the TOEIC Listening section, as will be further explained in section 4, 60% of the questions require test-takers to listen to spoken conversations or announcements, and select answers based on these talks. These questions can be made highly reliable because all test-takers hear the same recording and have selected-response answers to mark. However, this “may result in response processes that deviate significantly from those predicted… in non-testing contexts” (Rupp et al. 2006: 469). The test-taker may be able to break the text into lexical chunks and choose a correct answer based on grammatical or lexical clues rather than understanding the entire situation (ibid). The authenticity of this type of question is lowered because it is unlikely that a test-taker will be in a situation listening to English speech without being required to respond more fully than marking a piece of paper.

Directly related to authenticity, interactiveness is the “extent and type of involvement of the test-taker’s individual characteristics in accomplishing a test task” (Bachman and Palmer 1996: 25). The higher the level of interaction in a test, the higher the authenticity should be. Interactiveness becomes an issue when language skills are being tested indirectly. For tests that only examine receptive skills, “if use is what we are interested in, that gap will mean that test scores are at best giving incomplete information” (Hughes 2003: 76). As noted in the research performed by Liao, Qu and Morgan in section 2.3.3, the four sections of the TOEIC test four different aspects of English (2010: 5-6). Therefore, indirectly testing speaking through the use of only the L/R test may lower validity in part because of the lack of interactiveness.

Neither interactiveness nor authenticity is finite. Bachman and Palmer (1996: 39) state that:
“We speak of ‘relatively more’ or ‘relatively less’ authentic or interactive, rather than ‘authentic’ and ‘inauthentic’, or ‘interactive’ and ‘non-interactive’”.

Although these factors cannot be judged absolutely, they still affect test impact. Impact will be discussed in the following section.

2.5. IMPACT
Impact is known by a variety of names in language testing literature. It has been referred to as; washback; backwash; test impact; measurement-driven instruction; curriculum alignment; and test feedback (Muñoz and Alvarez: 2010; Brown and Hudson: 1998; Bachman and Palmer: 1996; Fredericksen and Collins: 1989; Messick: 1996). This paper will divide validity into two discrete points; test impact, which is the influence of the test on society, and washback, which is the effect on teaching. The first to be discussed is test impact.

2.5.1. TEST IMPACT
Languages are closely tied with the cultures in which they are used. Learning a second language is often concomitant with learning about these cultures. In countries with limited contact to outside cultures, learning languages can lead to gradual cultural changes as new ideas are introduced through this learning.

Test impact can be defined as “the ripples or waves [a test] makes in the wider educational and social world” (McNamara 2000: 72). This impact comes from a multitude of sources. Test developers, textbook publishers, teachers, parents, school districts, ministries of education and federal governments can all play a role in a test’s impact (Hamp-Lyons 1997: 298). This is most prevalent for high-stakes tests. Government and educational institutions that put an emphasis on language education drive motivation to achieve high scores on high-stakes language tests. These tests can affect social standing, and the lengths test-takers will go to in order to achieve high scores.

Social standing in South Korea can be influenced by scores on the TOEIC exam. Park and Abelmann (2004: 646) note that knowledge of English in South Korea is a sign of educational opportunity and experience of foreign cultures. One way of expressing knowledge of English is through TOEIC scores. The desire to get high test scores can lead to learners pursuing high scores on the TOEIC rather than learning the language (Choi 2008: 58). The pursuit of high scores often leads test-takers to attend TOEIC test-preparation classes. One company in South Korea became popular with test-takers because they
advertised that their students were able to greatly improve their TOEIC test scores. Early in
2012, this company was indicted for illegally obtaining TOEIC test questions (Jung and Son
2012: 1). However, this did not lead to a drop in student enrollment, even though students
were notified of this development (ibid). The desire to get high test scores superseded the
desire to abide by the moral norm of fair play. This example is also an illustration of the
washback effects the TOEIC has on teaching, which is the topic for the next section.

2.5.2. WASHBACK

Washback is defined by McNamara (2000: 72) as “the influence that testing has on teaching.”
Washback can be positive or negative. Negative washback can occur from the fear of poor
results, leading to “‘teaching to the test’, which is an undesirable ‘narrowing of the
curriculum’” (Smith cited in Alderson and Wall 1993: 118). This occurs in testing of all
subjects. In a U.S. study, “60% of maths teachers and 63% of science teachers described
negative effects on student learning resulting from mandated testing programmes” (Madaus
cited in Hamp-Lyons 1997: 296). Teachers are pressured to help students achieve high test
scores, which results in a narrowing of the curriculum, and questions about the content
validity of these tests (ibid).

There are numerous ways for test developers to facilitate positive washback. In the previous
example, this can be done by sampling widely and unpredictably. Expanding rather than
narrowing the curriculum will support positive washback effects in all areas rather than those
currently included in the tests (Hughes 2003: 54).

In language testing, direct rather than indirect testing can promote the skills wished to be
tested. In this way, “practice for the test represents practice in those skills” (ibid). A test
judging spoken language ability that requires test-takers to talk to a test administrator may
promote positive washback. Teachers preparing students for this type of test must design
lessons that will allow the students to practice their spoken language skills.

Another method Hughes (2003: 55) recommends to promote positive washback is through the
use of criterion-referenced tests. In this way, students will understand that if they can
complete tasks “at the criteria level, then they will be successful on the test” (ibid). This may
promote students learning to perform tasks rather than competing to get a high score in
comparison to their peers.
This paper has so far examined overall factors affecting language testing. The next section will examine sociological factors affecting language testing specifically in South Korea.
3. **Sociological Factors Affecting Language Testing in South Korea**

South Korea has a long history of using high-stakes tests. Beginning in the Chosun Dynasty (1392-1897), government officials were required to take tests based on the Confucian model to gain employment (Choi 2008: 40-41). This style of test is still used widely in gaining government and private sector employment.

In public education, the tradition of high-stakes testing is best exemplified by the Korean Scholastic Aptitude Test (KSAT) which is used for admission to university. Since the KSAT “literally determines the future education and career of test-takers, it is considered the most important high-stakes test in Korean education” (Choi 2008: 55). Unlike the TOEIC, this test is only administered once a year and students that do not perform well often study for an additional year or longer to improve their score and chances of getting into one of the best universities in South Korea. This test is considered so important for students that on test day the government asks citizens not to drive in the morning to ensure that students will not be late to the exam because of traffic congestion.

The KSAT includes an English section which employs selected-response questions to test reading and listening comprehension. The lack of questions designed to test productive language skills may result in these skills being under-represented by high school teachers (Choi 2008: 41). As a result, university students may attend good universities based in part on their English language ability, yet are not able to use the language. This suggests that the test has low validity, as it is not testing the structures it claims to be testing.

English language education in South Korea is part of the country’s desire to become more globalised. Former President Kim Sang Min introduced his *segyehwa* (globalisation) project, which made English a mandatory subject for all students beginning in elementary school in 1995 (Park and Abelmann 2004: 649). Programs such as these may be influenced in part by the success of large South Korean corporations such as Samsung and Hyundai. Since these corporations have become international entities, there is a necessity for their employees to communicate with non-Korean speakers in all of the countries in which they operate.

Some South Koreans feel that public education does not provide adequate English training to be able to use the language enough to cope with the necessities of the modern global society. This has resulted in the creation of a large private language education industry. A survey performed by the Samsung Research Institute estimated that 15 trillion won (approximately
£8.3 billion) is spent yearly on private English education, and a further 700 billion won (approximately £390 million) is spent on standardised tests such as the TOEIC (Chosun Ilbo 2007). The emphasis on language testing is encouraged by large corporations, which often require applicants to submit results of English proficiency tests as part of the interview process.

With the emphasis placed on English through public and private education, and high-stakes tests, knowledge of English has become a class marker for South Korean citizens (Park and Abelmann 2004: 646). It is seen as a sign of a good education, as well as a sign of travel experience, or the ability to travel easily. Given the implications of English knowledge in South Korea, a large portion of the population desires to have good English ability.

To this point, this paper has dealt with overall theories dealing with language testing, and issues affecting English language education and testing in South Korea. The next chapter will move on to examine the TOEIC test and its usage in South Korea.
4. The TOEIC Test

Since its inception in 1979, the TOEIC has grown to be one of South Korea’s largest high-stakes proficiency tests. In 2010, more than six million TOEIC tests were administered, making it “the largest and most widely used English-language assessment for the workplace” (ETS 2011). Of those six million tests, over two million tests were administered in South Korea.

The TOEIC has generated large numbers of advocates and critics. Supporters of the TOEIC battery of tests claim that test-takers are required “to use multiple strategies and abilities to comprehend and connect information” (ETS cited in Stoy off 2009: 28). The use of multiple strategies means that interactivity is quite high, but this is in situations where all of the sections are taken.

TOEIC detractors argue that the TOEIC is formalised, and allows students to employ test taking strategies to get a high score without knowing how to use the language. The Chosun Ilbo (2005b) claims that the test has become so highly structured that one TOEIC preparation instructor found that if ‘South Korea’ appears in a question, the answer is always ‘strategic’.

The current TOEIC is made up of three tests; the L/R test, a speaking test and a writing test. Both the speaking and writing tests are optional. The L/R test is made up of 200 questions, as outlined in table 4.1.

<table>
<thead>
<tr>
<th>Listening Comprehension</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Part I</td>
<td>One picture, four spoken sentences</td>
</tr>
<tr>
<td>Part II</td>
<td>Spoken utterances, three spoken responses</td>
</tr>
<tr>
<td>Part III</td>
<td>Short Conversation, four printed answers</td>
</tr>
<tr>
<td>Part IV</td>
<td>Short talks, three printed questions and answers</td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading Comprehension</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Part V</td>
<td>Incomplete sentences</td>
</tr>
<tr>
<td>Part VI</td>
<td>Text Completion</td>
</tr>
<tr>
<td>Part VII</td>
<td>Reading comprehension - single passages</td>
</tr>
<tr>
<td></td>
<td>Reading comprehension - double passages</td>
</tr>
<tr>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1 –TOEIC L/R test format (ETS 2008: 2)

The scores are norm-referenced, and reported on separate scales of 5 to 495 with a maximum combined score of 990 (Woodford 1982: 3).

The TOEIC was designed for businesses, so:
“… content was based on what test developers expected examinees would encounter in workplace settings where English was used for general commercial purposes” (Stoynoff 2009: 28).

All of the questions to judge workplace English in the L/R test are selected-response. To insure that items meet test criteria, all questions are examined by 20 reviewers before they are used in a test (Powers 2010: 4).

The optional TOEIC speaking and writing tests are not taken at the same time as the L/R test. The speaking test is composed of 11 questions, and the writing test is made up of 8 questions. Both the speaking and writing test are scored from 0 to 200. The speaking and writing tests were first offered in December 2006 (Liao et al. 2010: 2). Both of these tests are computer-administered, and scored online. As the next section will show, the speaking and writing tests are administered far less than the L/R test in South Korea, and as a result, an in-depth analysis falls outside the scope of this paper.

4.1. TOEIC Usage in South Korea

As mentioned previously, many colleges and universities incorporate the TOEIC test into their language requirements. However, there is no government mandate for universities to incorporate TOEIC tests or preparation into the curriculum (Chun 2012). In 2011, the L/R test was administered more than 2.1 million times (ETS 2012a). However, this does not mean 2.1 million people took the test, as test-takers are able to take the test multiple times. By comparison, the Speaking test was administered fewer than 200,000 times (ETS 2012b). Data was not available to this author on the writing test, but it is likely that it is taken in similar frequency to the speaking test. This suggests that while the reading and listening sections were implemented to enhance the TOEIC, these are not yet required in South Korea to the extent that the L/R test is.

The TOEIC is not always used in an appropriate manner by South Korean corporations. Gilfert (1996: 2-3) notes that some employers “use the TOEIC incorrectly, by requiring their domestic employees (who do not use English on a regular basis) to obtain a certain score for promotion or raises.” A possible reason is that companies use the TOEIC as a proxy for employee work ethic, rather than the need to learn English.

The majority of TOEIC L/R test-takers in South Korea are of university age. Figure 3.2 shows the ages of L/R test-takers in South Korea. Not surprisingly given the age of most
South Korean test-takers, 43% of respondents stated they were taking the TOEIC L/R test to get a job (ETS 2012a: 14). A further 15% noted they were taking the TOEIC for graduation (ibid).

Figure 4.2 – TOEIC L/R South Korean test-takers’ age (ETS 2012: 14)

The next section will discuss research conducted at a South Korean university to gather more information on the impact of the TOEIC test.
5. **RESEARCH METHODOLOGY**

The purpose of this study is to determine the effects of the TOEIC test in South Korean university classrooms. It is concerned with the questions in table 5.1.

<table>
<thead>
<tr>
<th>Question 1</th>
<th>What are the effects of studying for the TOEIC on South Korean university student motivation to learn English?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypothesis 1</td>
<td>Student motivation to learn English is negatively impacted by studying for the TOEIC.</td>
</tr>
<tr>
<td>Question 2</td>
<td>Does the TOEIC influence language teaching in TOEIC preparatory courses at South Korean universities? If there is influence, to what degree is it positive or negative?</td>
</tr>
<tr>
<td>Hypothesis 2</td>
<td>The TOEIC test has negative washback on teaching by narrowing the curriculum that is taught.</td>
</tr>
</tbody>
</table>

| Table 5.1 – Research questions and hypotheses |

This research is being conducted because of the prevalence of the TOEIC in South Korea, and the need to understand the validity of its use in university education.

5.1. **RESEARCH INSTRUMENTS**

This study collected three types of data to determine the effects of the TOEIC and TOEIC training programs in a South Korean university. The TOEIC training consists of two 75 minute classes per week for 15 weeks. One class is taught listening comprehension by a native English speaking professor, and the second class is taught reading comprehension by a Korean professor. The two classes’ grades are then combined for one final score. Students are required to pass two semesters of this instruction to graduate.

5.1.1. **QUANTITATIVE SURVEY**

The initial research conducted was a student quantitative survey. This consisted of 17 selected-response questions. The survey was written in English, and then translated into Korean (see Appendices A and B) by a native Korean speaking teacher with near-native English speaking skills. These translations were later translated back to English by a second Korean speaking teacher with a similar skill set to ensure correct translation. The Korean version was administered in order to reduce confusion about the meaning of the questions or answers.
The questions on the quantitative survey were designed to cover several topics. Questions 1 through 3 looked at students’ knowledge of the TOEIC test, and their experience with it. Questions 5, 7 and 10 examine how well the students perceive studying for the TOEIC will improve different aspects of their English ability. The results from these questions will show whether the students believe studying for the TOEIC will improve their English, or merely improve their scores. Positive results could suggest that the TOEIC has positive washback, and negative results could suggest negative washback. Question 5 is used as a baseline, and it is expected that students will agree that studying for the TOEIC will improve their TOEIC score.

Question 6 was included to determine if excess stress could have an effect on overall test reliability. Question 8 examines one of the possible reasons why the TOEIC Speaking and Writing sections are not administered as often as the L/R sections. Question 14 examines the expectancy of students to continue to use English after taking the TOEIC, while question 17 is concerned with one of the social impacts of the TOEIC on Korean society.

The remaining seven questions examine the influence the TOEIC has on student motivation. Question 4 looks at the reason they are studying to take the TOEIC. Question 9 examines if the students feel they are improving quickly. If students agree with this, it will suggest improved motivation, and if they disagree, decreased motivation. Questions 11 through 13 are concerned with three sociological factors affecting motivation to get high TOEIC scores. Questions 15 and 16 examine if students prefer studying for the TOEIC to other forms of English instruction, and their interest in studying for the TOEIC. These concepts have a direct correlation with student motivation. Positive results may suggest positive motivation.

Quantitative data is able to provide an overview of these issues, and provide averages of opinions, however this may not note the underlying reasons for statistical similarities (Dörnyei 2007: 35). In order to get a better understanding of these issues, a qualitative survey was administered.

5.1.2. Qualitative Survey

In order to carry out mixed-methods research, a small scale qualitative survey consisting of 11 questions was administered to a separate group of students. The survey was written in English, translated into Korean, and the survey was administered to students with both the English and Korean questions written together (see Appendix C). Students mostly answered
in Korean. Their answers were translated into English using the same method employed for the survey questions.

Question 1 on the qualitative survey looked to expand upon the reasons why the Speaking and Writing tests are not taken as much in South Korea, corresponding with question 8 on the quantitative survey. Question 2 correlates with question 10 on the quantitative, examining if the students feel that studying for the TOEIC improves their English conversation abilities. Questions 3, 4 and 5 further expand on the perceived usefulness of studying for the TOEIC. Question 9 links with question 14 on the quantitative survey, examining how students plan to use their English knowledge, and if taking the TOEIC is appropriate or necessary for their career.

The remaining 5 questions are concerned with motivating factors in studying for the TOEIC. Questions 6 and 7 expand upon question 4 in the quantitative survey. These examine the reasons for taking the TOEIC, and student desire to study this subject. Question 8 deals with the students’ interest in studying the TOEIC, which corresponds with questions 15 and 16 on the quantitative survey. Questions 10 and 11 look at the correlation between TOEIC scores and motivation to continue studying English.

One of the main reasons for conducting qualitative research in addition to quantitative is that it is “useful for making sense of highly complex situations” (Dörnyei 2007: 39). The additional data provided allows participants to provide opinions which minimises, but does not eliminate possibly leading answers. The combination of the two allows for multi-level analysis, and improves the validity of the survey by providing evidence for “convergence and corroboration of the findings” (Dörnyei 2007: 45).

As previously mentioned, several of the questions dealt with the possible washback effects of the TOEIC. As discussed in section 2.5.2, this is the effect that a test has on teaching. This effect can vary “from teacher to teacher depending on the teacher’s beliefs and attitudes” (Muñoz and Alvarez 2010: 36). In order to better judge the results of this, a series of class observations were carried out to understand the extent and nature of washback.

### 5.1.3. Classroom Observations

The original plan for the classroom observations was to observe three foreign professors teaching TOEIC listening comprehension classes and three Korean professors teaching TOEIC reading comprehension classes. This could provide insight into how different
educators approach TOEIC instruction. This research allows information to be gathered on the methodologies used in the classroom as well as the skills that are emphasised by teachers. During the observations, the observer sat in the back of the classroom, and did not interact with the instructor or students in any way. It is probable that the students recognised the observer as another member of the faculty.

Sixteen Korean professors were contacted several times to request permission to observe their classes. All of the Korean professors denied permission. When explanations for refusing permission were given, they claimed a lack of experience, or excess nervousness at having someone observe their classes. As a result, this portion of the research could not be completed.

In order to gain some insight into how the Korean professors teach classes, small group conversations were held with students in two TOEIC listening comprehension classes. These students were enrolled in Korean professor taught TOEIC reading classes concurrently with the listening comprehension classes. Small groups were implemented in order to reduce student stress at having a one-on-one interview discussing other professors. This also allowed students to expand upon ideas presented by other students. These students were informed that any details they relayed would be used to make generalizations, and would not be referenced to specific instructors. This data is not direct observation, and hence conclusions are more difficult to reach.

5.2. Participants
A variety of students participated in the quantitative survey. The survey was taken by students in three TOEIC listening comprehension classes. Two of the classes were attended by first year students. These classes were the first of their two mandatory semesters of TOEIC instruction. A total of 104 students responded to the survey in these two classes. These students will be referred to as “group A”. The students in group A have a wide variety of majors. This was their first semester of exposure to university classes. In South Korean public schools, students are never held back or forced to repeat classes based on grades, so this was the first semester where students’ in-class performance had an effect on what they would study in following semesters. The failure rate in first semester TOEIC classes at this university is typically 30%.
The third class of students were in their second, third or fourth year, and were completing an optional second year of TOEIC instruction. However, this course can also be taken to fulfill their second English requirement. A total of 24 students responded to the survey in this group. This class will be referred to as “group B”. The male students in group B have all completed their mandatory military service, and are typically two to three years older than the females in the same year of education. Due to their military training and age, these male students are generally more serious about their studies than the male students in group A.

The qualitative survey was administered to a separate group of second year students. A total of 17 students participated in this survey. These students had completed the required two semesters of TOEIC listening and reading comprehension classes. They either chose not to take the optional further two semesters of TOEIC instruction, or had not yet started these classes.

As stated previously, three foreign teachers’ classrooms were observed. To collect data on the Korean professors’ classrooms, groups of five and seven students were asked to describe how their Korean teachers taught the reading comprehension classes. These were from the two classes making up group A. The groups of five and seven students will be referred to as groups C and D, respectively.

5.3. Administration
Research was conducted at a four year University in Seoul, South Korea. To conduct the surveys, classes were finished 15 minutes early, and volunteers were asked to complete them. Students not wishing to participate in the surveys were allowed to leave class early. For the group discussions about teaching methods, small groups of volunteers were asked to stay after class. The next section will examine the data collected through the research.
6. RESULTS
This section begins with the results of survey questions providing background on student understanding of the TOEIC. It then examines questions dealing with TOEIC impact on student motivation, and issues with washback, including the classroom observations.

6.1. BACKGROUND ON STUDENT KNOWLEDGE OF THE TOEIC
As mentioned previously, the three initial questions on the quantitative survey were designed to provide background information on student familiarity with the TOEIC test. These questions provided unexpected results. Question 1 asked the students how long they have studied for the TOEIC. In group A, 50% of students answered that they had not studied for the TOEIC. These students were ten weeks into their first semester of TOEIC preparatory classes. In group B, 17% of students answered that they had not studied for the TOEIC. These students were ten weeks into their third semester of TOEIC classes. Possible reasons for the disparity in these answers will be outlined in section 7.

In question 2, 13% of students in group A abstained from answering. All of these students answered in question 1 that they had not studied for the TOEIC. This leaves 37% of the students in group A that answered they had not studied for the TOEIC at all, and then in the next question chose at least one section they were currently studying for. In group B, all of the students marked that they were currently studying for at least one section of the TOEIC. One possibility for the differences between the answers in questions 1 and 2 may be the students not having an option in question 2 to note they were not studying for any sections of the TOEIC.

Question 8 on the quantitative survey and question 1 on the qualitative survey examined further information about why students do not study for the Speaking and Writing portions of the TOEIC as much as the L/R sections. In question 8, 74% of all students surveyed said they would be more likely to take the Speaking or Reading sections if the time and place were more convenient.

Students provided a variety of answers on the qualitative survey. These are listed in table 6.1.
Question 3 on the quantitative survey checked to establish student understanding of TOEIC testing objectives. The results from this question can be seen in figure 6.2. A much higher proportion of the older students in group B understood that the TOEIC tests business English compared to the younger students in group A.

<table>
<thead>
<tr>
<th>Answer</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>No intention - no need</td>
<td>5</td>
</tr>
<tr>
<td>No intention - too difficult</td>
<td>4</td>
</tr>
<tr>
<td>No intention - no reason specified</td>
<td>2</td>
</tr>
<tr>
<td>Plan to take one or both - curious about my level</td>
<td>3</td>
</tr>
<tr>
<td>Plan to take one or both - need it for work</td>
<td>2</td>
</tr>
<tr>
<td>Plan to take one or both - no reason specified</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.1 – Student intentions to take the TOEIC Listening or Reading sections

Question 6 on the quantitative survey asks the students about anxiety, as it relates to performance on the TOEIC. In addition to providing background, this question deals with one of the major factors affecting test-taker reliability discussed in section 2.2.2. In both groups, the majority of students answered that they feel anxious to perform well on the TOEIC. As well as affecting test-taker reliability, stress can have an effect on motivation, which is discussed in the following section.

Figure 6.2 – Student opinions on what the TOEIC examines
6.2. Motivational Influences of Studying for the TOEIC

The first question examining motivation on the quantitative survey asks the reasons why students are studying for the TOEIC exam. In group A, 75% of students responded they are studying for the TOEIC because it is a university requirement. Nine students wrote in answers. These are listed in table 6.3. As shown, one student noted they had not studied for the TOEIC. As mentioned, this survey was conducted in a TOEIC class.

<table>
<thead>
<tr>
<th>Reason</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Korean Augmentation to the United States Army (KATUSA)</td>
<td>3</td>
</tr>
<tr>
<td>To get a job</td>
<td>2</td>
</tr>
<tr>
<td>TOEIC is necessary for my future</td>
<td>2</td>
</tr>
<tr>
<td>TOEIC is important in Korean society</td>
<td>1</td>
</tr>
<tr>
<td>I did not study for the TOEIC</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.3 – Quantitative survey alternate reasons for studying for the TOEIC

In group B, 48% of students stated their reason for TOEIC preparation was to meet university requirements. This was followed by 30% who stated they were studying of their own volition, and 30% who marked other. All of the students who marked “other” in group B stated they were studying TOEIC to get a job.

Questions 6 and 7 on the qualitative survey expanded upon the reasons why students studied TOEIC. Question 6 asked the students to list three reasons why they studied TOEIC. A wide variety of answers were given. These are shown in table 6.4.

<table>
<thead>
<tr>
<th>Reason</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>To get a job</td>
<td>12</td>
</tr>
<tr>
<td>To raise my English level</td>
<td>7</td>
</tr>
<tr>
<td>To meet requirements for graduation</td>
<td>6</td>
</tr>
<tr>
<td>To raise my TOEIC score</td>
<td>4</td>
</tr>
<tr>
<td>To be able to talk with foreigners</td>
<td>1</td>
</tr>
<tr>
<td>To get a scholarship</td>
<td>1</td>
</tr>
<tr>
<td>To get certification</td>
<td>1</td>
</tr>
<tr>
<td>For satisfaction in getting a high score</td>
<td>1</td>
</tr>
<tr>
<td>To improve listening skills</td>
<td>1</td>
</tr>
<tr>
<td>To improve grammar</td>
<td>1</td>
</tr>
<tr>
<td>To improve vocabulary</td>
<td>1</td>
</tr>
<tr>
<td>I didn't study for the TOEIC</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.4 – Qualitative survey reasons for studying for the TOEIC
The results from this question support the findings of the quantitative survey. Many students study TOEIC because of its importance in getting a job. The other major motivating factor was to complete the course for graduation. Question 7 asked if they would still study TOEIC if it was not required for their job or graduation. 12 of 17 students answered they would not study TOEIC, without offering additional information. Three students answered they would not study TOEIC, but would continue to study English. The final two students said they would study TOEIC because they enjoy learning English.

As further back-up for this data, questions 11 through 13 on the quantitative survey asked if getting a high score on the TOEIC is important for the test-takers’ future employment, education and personal life. A strong majority of students in both groups agreed with all three of these questions. These results indicate the students’ understanding that the TOEIC is a high-stakes test with strong implications for their future.

Question 9 on the quantitative survey asked if students noted changes in their English ability while studying for the TOEIC. The results are shown in figure 6.5.

![Figure 6.5 – Student opinions on improvement of their English level](image)

Questions 10 and 11 on the qualitative survey asked if getting high or low scores would result in increased or decreased motivation to continue studying English. Their answers are provided in tables 6.6 and 6.7.
Questions 15 and 16 on the quantitative survey asked about student preference in studying TOEIC to conversation, and if studying TOEIC increased their interest in learning English. In both groups, over 70% of students answered that they preferred studying conversation to studying TOEIC, and studying TOEIC did not increase their interest in learning English.

Questions 15 and 16 were expanded upon in question 8 of the qualitative survey. The results from this question were mixed. Seven of the 17 respondents gave noncommittal answers, and one student did not answer. Of the remaining nine respondents, three stated studying TOEIC decreased their interest because it was only about getting a score, talking about test tips, and repeating the same structures. The reasons stated for a raise in interest level were; curiosity about their English abilities and; wanting to raise conversation abilities.

The reasons the students gave for changes in motivation are largely dependent on the methods used to teach their TOEIC classes. This is a result of test washback, which is discussed in the next section.
6.3. Washback Effects of the TOEIC in South Korean University Classrooms

As noted in section 2.5.2, washback is “the influence that testing has on teaching” (McNamara 2000: 72). A number of questions on both surveys as well as the classroom observations dealt with this topic. Question 5 on the quantitative survey found that, as expected, students felt that studying for the TOEIC will improve their TOEIC score. This result was quite different from question 10, which asked if studying for the TOEIC improves communication skills. The results from question 10 are shown in figure 6.8.

![Figure 6.8 – Opinions on studying for the TOEIC improving communication skills](image)

Question 7 on the quantitative survey expanded on student opinions on the benefits of studying TOEIC. It asked if students felt they could complete more tasks as a result of TOEIC preparatory studies. In group B, 74% of students disagreed, and the remaining 26% only slightly agreed. These results correspond with those found in group A. 69% of students in group A disagreed that studying for the TOEIC helped them perform more daily activities in English.

The usefulness of studying TOEIC in order to improve English skills was expanded upon in the qualitative survey questions 2 through 5. 11 of the 17 students said that studying TOEIC did help their English communication. Seven of the students answered that studying listening and grammar are the basis of conversation, and that any study is helpful. Two students noted that listening is hard for Koreans, so studying the listening portion aids their communication. One student answered that it is helpful because it reduces fear, although it is not clear if they...
meant fear of communicating in English, or fear related to performance on the test. The final student did not specify a reason. Four students stated that studying TOEIC L/R sections is not helpful because it is only about how to perform better on the test. One student stated that, in order to improve conversation skills, the students must practice conversation.

Questions 3 – 5 on the qualitative survey asked the students further questions about the usefulness of studying TOEIC. In question 3, 10 students felt that it is useful, and gave answers such as grammar, listening and continued practice as their reasons. The seven students that disagreed again mentioned studying test skills only, and a lack of a speaking section. Questions 4 and 5 expanded on the aspects found useful or not useful. The results are shown in tables 6.9 and 6.10.

<table>
<thead>
<tr>
<th>How is studying for the TOEIC useful?</th>
<th>Answer</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improves grammar</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Improves listening comprehension</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Improves reading comprehension</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Improves vocabulary</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Useful for business</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>No answer</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.9 – Aspects of studying for the TOEIC students found useful (some respondents provided more than one answer)

<table>
<thead>
<tr>
<th>How is studying for the TOEIC not useful?</th>
<th>Answer</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Doesn’t teach how to use language</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>How to get a higher score</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>KSAT</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Teaches same materials over and over</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Homonyms</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>It’s just memorising</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>No answer</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>I didn’t study</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.10 – Aspects of studying for the TOEIC students did not find useful

In question 5, it is unclear if the students that answered ‘speaking’ mean that studying speaking is not useful because they are not taking a speaking test, or that a lack of speaking makes the study not useful.
As mentioned in section 4.1, achieving a certain TOEIC score is often required for South Korean employees, regardless of the necessity in their business to use English. Question 14 on the quantitative survey asked the students if they expected to use English regularly in their future job. The students overwhelmingly agreed in both groups that they will need to use English. This corresponded with question 9 on the qualitative survey which asked the students how they planned to use English after they took the TOEIC test. The results from this question are shown in table 6.1.

<table>
<thead>
<tr>
<th>Expected use of English after taking the TOEIC</th>
<th># of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Just continue practising</td>
<td>4</td>
</tr>
<tr>
<td>Speaking (with whom unspecified)</td>
<td>3</td>
</tr>
<tr>
<td>Travel abroad</td>
<td>2</td>
</tr>
<tr>
<td>Continue studying TOEIC</td>
<td>2</td>
</tr>
<tr>
<td>Simple reading</td>
<td>1</td>
</tr>
<tr>
<td>Speaking to foreigners</td>
<td>1</td>
</tr>
<tr>
<td>Reading English novels</td>
<td>1</td>
</tr>
<tr>
<td>Watching television and movies in English</td>
<td>1</td>
</tr>
<tr>
<td>Won't use English</td>
<td>1</td>
</tr>
<tr>
<td>No answer</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.11 – Student expected use of English after taking the TOEIC

Question 17 on the quantitative survey dealt with one aspect of test impact. It asked the students if getting a high score on the TOEIC was more important than learning English. Their answers are shown in figure 6.12.

![Figure 6.12 – Student opinions on the importance of TOEIC scores compared to learning English](image)
As mentioned previously, some of the students’ answers on both the quantitative and qualitative surveys may be a result of teaching techniques employed in the classroom. The next section will discuss the classroom observations conducted and the various teaching methods employed.

6.3.1. Classroom Observations

In order to understand how the TOEIC affects language teaching, three TOEIC classrooms were observed (see appendix E). Each class was 75 minutes long, and designed to teach TOEIC listening comprehension skills. Class C was composed of third and fourth year students, while classes A and B were made up of first year students. All three classes started with a quiz. The professors all used audio recordings to have the students do various practice sections of the TOEIC listening comprehension test. After completing the practice sections, all professors went over the questions again, and explained each answer.

The classrooms had a number of differences. One main factor was the level of interaction. Class sizes were large (see appendix E) compared to typical conversation classes, so professors A and B used a microphone connected to speakers to ensure all of the students could hear. Professor C used a microphone connected via headphones. The use of headphones appeared to reduce student interaction.

Apart from the technology used, or perhaps as a result of it, professor C used different teaching methods in the class. Class C was teacher-centred, the largest portion of time being used to explain TOEIC test taking strategies. The only time students interacted with the professor was to indicate a question they had difficulty with. Some students did this in the L1 instead of the L2, meaning they failed to use productive language skills in the class.

Unlike class C, classes A and B required students to use productive language skills in working through specific tasks. This was largely carried out in pair work exercises where students described pictures to their partners, or read words aloud to check for pronunciation and listening comprehension. Professor B explained TOEIC test-taking strategies, although not as in depth as professor C. Professor A did not explain any test-taking strategies.

Overall, professors A and B used task based instruction and student-centred learning throughout the lessons. This expanded on the materials taught in the book. Professor C designed a class centred on the TOEIC listening comprehension test, and did not include materials outside the scope of what the students needed to know in order to improve their
TOEIC score. In classes A and B, students that were not interested in class sat in the back of the classroom talking in the L1 and not following the tasks laid out by the professors. In class C, students not interested in class fell asleep. The classes all had similar percentages of students not paying attention.

In the TOEIC reading classes, student discussions from groups C and D revealed that non-native professors employed different teaching techniques from those of native English speaking professors. Both groups relayed that the majority of their classes were taught in the L1. This included explanations of grammar and vocabulary as well as translation of large portions of the text. The students also relayed that test-taking strategies were taught each week. The students stated that the only time English was used in the classroom was in reading the passages, or when professors corrected student pronunciation while reading.

Students from both groups were asked if there were any speaking or writing exercises employed in class dealing with the vocabulary or commenting on the reading. Both groups said that no tasks of this sort were used.

The classroom differences relayed by the students were minor. Group A stated that no homework was given in class. The professor went through the chapter and explained several tips relating to specific questions. In group B, each student was given five questions from the next chapter to prepare for the following class. The professor would then call upon students to answer the questions they were assigned in class and base explanations around issues the students had with the homework.

The next chapter of this paper will discuss the findings of the research.
7. Analysis of Findings

The research produced a number of interesting findings. An unexpected result of the research was students stated they had never studied for the TOEIC. Several reasons may be posited for this. Some may have thought the question was asking about studying outside of the university, or previous to that semester. Some students may not have understood that they were studying TOEIC despite attending lectures on the TOEIC. This could be a result of students hearing the word TOEIC regularly, but not associating it with the English they are being taught. If this is the case, it would suggest that for these students, TOEIC learning may have become completely disassociated from English learning. Another possibility is that the students considered the class very low priority compared to their other subjects, so were not paying attention in class, and therefore not actually studying.

Students were asked about sections of the TOEIC being studied to discover the ratio of those studying for the speaking and writing sections in addition to the L/R test. Due to the ambiguity of student responses, this could not be determined. South Korean data from ETS (2012a, 2012b) shows the L/R test taken at a ratio of over 10:1 compared to the speaking test. However, this does not take into consideration individuals repeatedly taking certain sections. This research was unable to expand upon the ETS data due to the inconsistencies in student replies.

The most common answer students gave for not studying TOEIC speaking or writing was because there was no need for it. As will be discussed in the next section, students understand that TOEIC scores are important in South Korea. The results of the survey indicate this does not yet include the speaking or writing tests. The lack of interest in sections not directly required for social or professional reasons indicate that students “wind up getting disoriented in pursuing the true goal of learning English and keep only one goal in mind… obtaining as high an EFL score as possible” (Choi 2008: 58). Since students view studying TOEIC speaking and writing as unnecessary, they may be more likely to ignore practice of productive language skills. This may be the cause of the findings of Choi (2008: 44) that “those who obtained a perfect score on the TOEIC were found to be seriously deficient in demonstrating communicative skills.”

Being overly concerned with scores can raise a test-taker’s anxiety to perform well on a given test. This was supported by the findings in question 6 of the quantitative survey. The majority of students responded that they feel anxious to perform well on the TOEIC. As
discussed in section 2.2.2, when test-takers complete tasks under excessive pressure, they are more likely to perform abnormally (Alderson and Wall 1993: 117). The stress placed on TOEIC test-takers in South Korea can result in lowered test reliability since the test-takers may not be performing as they would in a less stressful situation.

Another issue the research uncovered was students not understanding the purpose of their studies. Less than a third of the students in group A understand that the TOEIC tests business English. However, as noted, a number of these students did not know they were studying TOEIC, which likely had an impact on these results. The difference between the two age groups may also be attributed to group B having already studied the TOEIC for two semesters, and should therefore be more knowledgeable about the test and its contents. In addition, group B will need TOEIC scores to get jobs in the near future, making it more important for them to have a better understanding of the test, which should result in a rise in motivation to know more about the TOEIC. Motivation will be covered more fully in the next section.

7.1. FINDINGS ON MOTIVATION

One goal of this paper was to determine the effects of studying TOEIC on South Korean university student motivation. The hypothesis was that student motivation to study English is negatively impacted by TOEIC instruction. The results of the surveys showed a number of trends. First year students were primarily motivated to study TOEIC because of university requirements. In older students, the primary motivating factor was to get a job. These results indicate that the students are primarily extrinsically motivated. Only a small percentage of students noted that they are motivated to study TOEIC primarily because of their interest in learning English. Lightbown and Spada (2001: 33) argue that in situations where external pressure is the main motivating factor, “internal motivation may be minimal and general attitudes towards learning may be negative”. If, as suggested in the previous section; students do not correlate studying TOEIC with studying English; negative attitudes towards studying TOEIC are likely to have little impact on attitudes towards studying English. However, students for whom this is not the case may experience a decrease in motivation, which supports this paper’s hypothesis.

For a number of students in the observed classes, the extrinsic motivating factors for studying TOEIC did not appear sufficient to encourage language uptake. Numerous students were not interested in class, and either did not participate in tasks, or fell asleep. This trend held true
for all classes; regardless of age or teaching methods employed. The only difference was in how students did not participate. The older students in class C listening to the professor through headphones fell asleep. The younger students in classes A and B listening through speakers discussed unrelated issues in the L1 with classmates. Dörnyei (1994: 276) notes that “students will lose their intrinsic interest in an activity if they have to do it to meet some extrinsic requirement”. It is not known if the observed students were originally interested and lost motivation, or never felt motivated. However, it is possible that the students in these classes currently motivated to learn English may lose interest over time because of the extrinsic factors.

Alderson and Wall (1993: 119) suggest that a person’s need for achievement is balanced by “the motivation to succeed and the motivation to avoid failure”. Since the majority of first year students are motivated to meet university requirements, it is possible that many of them are motivated primarily to avoid failure. They are likely not particularly interested in the course or materials, and as a result do the minimum possible amount of work. It is unlikely that doing enough to avoid failure will translate to success in either English or TOEIC scores.

One unexpected result of the research is most students answered that both high and low TOEIC scores increase motivation to study. 13 students said they would continue studying if they received low scores, and 11 said they would continue studying if they received high scores. In contrast, two students said they would be less likely to continue studying English in either circumstance. This appears to contradict this paper’s hypothesis. However, this may also suggest that scores themselves, not specifically TOEIC scores are good motivating tools. Well designed English programs should include regular testing to increase this form of motivation.

Another motivating factor students gave for studying TOEIC was to raise their English level. Being a proficiency test, the TOEIC is designed to judge test-taker abilities. It is not designed to be used as a study aid. The majority of students also indicated that they did not feel they were improving quickly. Lightbown and Spada (2001: 42) argue that if students do not notice a positive change in skills over time, they are likely to become less motivated to continue. Since this TOEIC program has not helped students note improvements in their skills, their motivation is likely to decrease. These results are to be expected for the most part because studying a language for 2.5 hours per week as these students do in this class is
unlikely to quickly improve skills, regardless of the methods employed. A more wide-scale research project is required to see if these results hold true in other TOEIC programs.

Student preferences can have an effect on motivation. Lightbown and Spada (2001: 35) argue that, “virtually all learners, particularly older learners, have strong beliefs and opinions about how their instruction should be delivered”. The research showed that the majority of students preferred conversation classes to TOEIC classes. This suggests students would be more motivated in a conversation class rather than a TOEIC class. However, this result is not certain in the context of this research since many of the students indicated they did not know they were in a TOEIC class.

The majority of students noted the correlation between success on the TOEIC and success in education, business and their social lives. This is a reflection of the importance the TOEIC and other standardised tests in South Korea. The goals that motivate students are often to perform well on high-stakes proficiency tests such as the TOEIC. Since the students feel the TOEIC is important for many aspects of their lives, they may concentrate more on test-taking strategies than on language skills, as noted in the previous section. This may have an effect on the validity of the TOEIC, which will be discussed further in section 7.3.

Overall, the research showed there is little indication that studying for the TOEIC increases motivation. There is evidence to indicate that studying for the TOEIC may decrease student motivation to learn English, but further research into the complex relationship between motivation and learning is needed.

As noted in section 4.1, universities have a number of reasons to wish to include TOEIC instruction as part of the curriculum. The TOEIC adds an international component to programs, and can improve student employability. Both of these can raise university rankings. However, by mandating TOEIC classes, and not offering them as an option forces students unmotivated to study TOEIC into these classes. This can negatively affect student perceptions of learning English, and create unwanted washback effects on teaching. The next section will examine the findings on washback.

7.2. Findings on Washback
The second hypothesis postulated by this paper is that the TOEIC promotes negative washback through a narrowing of the curriculum in South Korean universities. In any discussion of washback, it is important to note Alderson and Wall’s claim that a test “cannot
determine how teachers teach, however much it might influence what they teach” (1993: 127 emphasis in original). The classroom observations appear to support this argument. The three observed professors each used their own style of teaching. It is unlikely they changed their teachings habits because of the subject they were teaching. The two professors that used task-based exercises likely use similar types of tasks in non-TOEIC based classrooms. The effect of the TOEIC was in the subjects they taught.

The clearest washback effect the TOEIC had on the observed classes was the teaching of test-taking tactics. This is a clear example of the teachers doing “things they would not otherwise do because of the test” (Alderson and Wall 1993: 117 emphasis in original). The effect for the students is they will be able to pick up clues on how to get a higher score on the TOEIC. This will help students to raise their scores. However, this does not mean their English proficiency is improving, only their test-taking skills.

In the classrooms taught by the non-native English instructors, it appears as though the washback effects are more pronounced. The students in groups C and D did not mention any activities or tasks undertaken in class that went outside of information required for the TOEIC. The non-native professors appear from these answers to have narrowed their curriculums to concentrate only on language involved in taking the TOEIC reading test. This was also suggested from student responses on the qualitative survey.

In addition, it appears as though the majority of classes were taught in the L1. Some use of L1 in the classroom can be beneficial to students struggling to understand certain concepts, but overuse will likely lead to a reduced uptake of target language. It is possible that the limited use of English in the classrooms was one of the causes of the professors not wanting an observer in their classrooms. It may be that the non-native professors understand that they are giving instruction that is useful primarily to raise scores, and not to improve fluency. This may account for their replies stating that they were too inexperienced to be observed. However, further research is required to make further claims.

Students in both native and non-native classes did not find TOEIC instruction to be useful for anything outside the scope of the L/R test. A number of students remarked that studying test-taking tips was the least useful aspect of studying for the TOEIC. The areas they noted as not useful suggest they have been in classes where the curriculum has been narrowed in order to help them achieve higher scores. These results support the hypothesis put forward in section 5.
Teachers may try to teach a subject as normal, but “find ‘teaching to the test’ almost unavoidable” (Bachman and Palmer 1996: 33). One reason that teaching to the test is difficult to avoid was demonstrated in question 17 of the quantitative survey. Over 40% of students agreed that TOEIC scores are more important than learning English. This result shows how important the TOEIC and other standardised tests are in South Korean society. With such a large percentage of students more concerned with improving scores instead of learning the language, there is pressure placed on instructors to help students achieve these results. The result of this pressure is the undesirable ‘narrowing of the curriculum’ noted in section 2.5.2.

Further evidence of the curriculum being narrowed can be found in the answers to questions 7 and 10 of the quantitative survey. The majority of students answered that they are not able to do daily activities in English more effectively as a result of their TOEIC study. In addition, 65% of the students in group B did not feel their communication skills were improving. The most common answer in group B was completely disagreeing that studying for the TOEIC improves communication skills. This may be because the students in group B have more experience in studying for the TOEIC, or that they have a better understanding of how it has, or has not, benefitted them than the students with less experience in group A. One would expect improvement in these two areas in English classes not designed for test preparation.

Washback is not always negative. As outlined in section 2.5.2, there are many ways tests can be designed to facilitate positive washback. However, the research did not discover any areas in which the TOEIC L/R test facilitated positive washback. Since evidence for positive washback was not found, the TOEIC’s use as a teaching tool must be questioned, regardless of the importance of getting high TOEIC scores in South Korean society.

So far this paper has examined effects the TOEIC has on student motivation to learn English and Washback effects of the test. The next section will look at these results and their implications for the validity of the TOEIC.

### 7.3. Implications for TOEIC Validity

As this paper has shown, the TOEIC L/R test is a high-stakes test in South Korea. It is therefore necessary for it to be reliable. As discussed previously, reliability can have an inverse relationship with validity, interactivity and authenticity. One reason for TOEIC
developers to create the Speaking and Writing tests was to include parts of the test with reduced reliability, but increased validity.

One issue surrounding TOEIC use in South Korea is that the L/R test is taken much more often than the Speaking or Writing sections. Research at Samsung has found that as a result of this usage, employees with TOEIC scores over 860 points had “the ability to speak with a proficiency sufficient for business… only 10 percent of the time” (Choi 2008: 44). It is likely that those with high scores have taken courses similar to those outlined in this paper, with the negative washback effects outlined previously.

The inclusion of TOEIC test preparation at universities brings up another issue with the TOEIC. The TOEIC is a proficiency test, designed to judge a test-taker’s ability at the time of the test. Proficiency tests are not designed to be used as study material. Due to the widespread studying of TOEIC in South Korea, it becomes less of a proficiency test, and more of an achievement test. Students such as those involved in the research for this paper study for the TOEIC, and may use their score to judge how much has been learned. However, as discussed, this is often knowledge of the test rather than language knowledge. In the research, some students responded that they plan to continue to study the TOEIC after taking the test. This will likely serve to exacerbate this problem.

The research found some conflicting data on how students expected to use their knowledge gained in their TOEIC course. Over 90% of students that participated in the quantitative survey answered that they expected to use English frequently in their future jobs. In contrast, none of the respondents in the qualitative survey mentioned using English in a business environment. It may be that students answering on the quantitative survey expect to have different careers than those answering the qualitative. Alternately, they could be thinking about the importance of English in South Korean society mentioned in section 3, and as a result, assume that English will play a role in their employment. If the students in group C are correct, and they will be using their knowledge outside of business, then studying for the TOEIC may be unnecessary, or counterproductive to their English language goals.

The application of the TOEIC L/R test as an overall proficiency test indirectly tests speaking and writing skills. This is one of the reasons ETS states that “none of the TOEIC tests are regarded as an appropriate substitute for any one of the other TOEIC tests” (Powers 2010: 2). This means that the TOEIC L/R test is used in South Korea for purposes it was not intended. Choi (2008: 44) argues:
“It is essential that the test be utilized to serve the purpose intended by the test developer, if it is to be considered a valid test.”

Since the TOEIC L/R test is used for purposes other than those it was created for in South Korea, it may lower both the content and construct validity of the test. This argument is supported by Bachman (cited in Choi 2008: 44), who states:

“Utilizing the TOEIC as a multi-purpose test may compromise the validity of the test, if validity is defined as the extent to which test results are interpreted appropriately and meaningfully.”

Since the TOEIC L/R test is used to indirectly test productive language skills, the narrowing of the curriculum reduces productive language. Use of direct testing, as is done in the TOEIC Speaking and Writing sections reduce this effect. Hughes (2003: 54) states that “if we test directly the skills that we are interested in fostering, then practice for the test represents practice in those skills.” It appears that one of the easiest ways to solve some of the issues put forward in this paper would be to require all sections of the TOEIC to be administered to each test-taker. In this way, test-takers would not be able to focus solely on selected-response test taking tips in order to achieve a high score. Alternately, it is possible that the Speaking and Writing Sections could replace the L/R test. In this way, the students would be involved in direct testing, and practicing their productive language skills in the classroom. This may also bring the positive washback effects of direct testing into TOEIC classrooms. Further research is needed to discover the advantages and disadvantages of integrating this system.
8. Conclusion

Given the wide use, and misuse, of the TOEIC L/R test in South Korea, this paper has attempted to shed light on its impact. It is hoped that this will add to the literature on the use of high-stakes tests in order to facilitate positive change.

The research found that, when students have are required to study for the TOEIC L/R test, this can lead to poor impressions of English language learning as a whole. When offered as electives, these negative impacts may be ameliorated.

It was confirmed that the TOEIC is important for finding employment after graduation. Stress encountered by students could be reduced by restricting it to those intending to use English on a regular basis.

The findings on washback revealed that it exist in TOEIC preparatory classrooms. Its impact was found to be negative, largely because the TOEIC L/R test is used as an indirect test of speaking and writing skills. Increasing the range of study to include education into the speaking and writing sections of the TOEIC would promote positive washback through the use of productive language skills.

It is hoped, through research such as this, that educational institutions will increasingly recognise English language teaching in general and the TOEIC in particular, as involving more than helping students to pass tests.
APPENDIX A

Student Quantitative Survey (English)

1. How long have you studied for the TOEIC test?
   a. None  d. Up to 6 months
   b. Up to 1 month  e. Up to 1 year
   c. Up to 3 months  f. More than 1 year

2. What sections of the TOEIC test are you currently studying for? (circle all that apply)
   a. Listening  c. Speaking
   b. Reading  d. Writing

3. What do you think the TOEIC test examines?
   a. Conversation English  c. Basic English
   b. Business English  d. Academic English

4. Why are you currently studying for the TOEIC test? (circle all that apply)
   a. Encouraged by private tutor  e. Encouraged by employer
   b. Encouraged by school teacher  f. Decided on my own
   c. Encouraged by parents  g. Required by university
   d. Encouraged by friends  h. Other (please specify)

5. I feel that studying for this test will improve my TOEIC score.
   a. Completely agree  d. Slightly disagree
   b. Agree  e. Disagree
   c. Slightly agree  f. Completely disagree

6. I feel anxious to perform well on the TOEIC test.
   a. Completely agree  d. Slightly disagree
   b. Agree  e. Disagree
   c. Slightly agree  f. Completely disagree

7. I feel I can perform more daily activities in English because of studying for the TOEIC test.
   a. Completely agree  d. Slightly disagree
   b. Agree  e. Disagree
   c. Slightly agree  f. Completely disagree

8. I would be more likely to take the TOEIC Speaking or Writing tests if the places and times were more convenient.
   a. Completely agree  d. Slightly disagree
   b. Agree  e. Disagree
   c. Slightly agree  f. Completely disagree
9. I do not feel my English is improving quickly while studying for the TOEIC test.
   a. Completely agree  d. Slightly disagree
   b. Agree  e. Disagree
   c. Slightly agree  f. Completely disagree

10. I feel that studying for this test will improve my English communication skills.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

11. I feel that getting a high score on the TOEIC is important for my future employment.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

12. I feel that getting a high score on the TOEIC is important for my future education.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

13. I feel that getting a high score on the TOEIC is important for my personal life.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

15. I prefer studying TOEIC to studying conversation.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

16. Studying TOEIC makes me interested in learning English.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree

17. I feel that getting a high score on the TOEIC is more important than learning the language.
    a. Completely agree  d. Slightly disagree
    b. Agree  e. Disagree
    c. Slightly agree  f. Completely disagree
APPENDIX B

Student Quantitative Survey (Korean)

1. TOEIC 공부를 한 지 얼마나 되셨습니까?
   a. 전혀 안 함
   b. 1 개월
   c. 3 개월
   d. 6 개월
   e. 1 년
   f. 1 년 이상

2. 현재 TOEIC의 어떤 영역을 공부하고 있습니까? (모두 선택)
   a. 듣기
   b. 읽기
   c. 말하기
   d. 쓰기

3. TOEIC이 어떤 부분을 테스트한다고 생각하십니까?
   a. 영어 회화
   b. 비즈니스 영어
   c. 기본적인 영어
   d. 학문적인 영어

4. 현재 TOEIC 공부를 하는 이유는 무엇입니까? (모두 선택)
   a. 과외 선생님의 권유
   b. 학교 선생님이 권유
   c. 부모님이 권유
   d. 친구들이 권유
   e. 고용주가 권유
   f. 스스로 결심
   g. 대학교 필수 과정
   h. 기타 (직접 기술해 주세요)

5. TOEIC 공부를 하면 TOEIC 점수가 더 높게 나올 것이라고 생각한다.
   a. 매우 그렇다
   b. 그렇다
   c. 조금 그렇다
   d. 다소 그렇지 않다
   e. 그렇지 않다
   f. 매우 그렇지 않다

6. TOEIC 시험을 잘 볼 수 있을지 걱정된다.
   a. 매우 그렇다
   b. 그렇다
   c. 조금 그렇다
   d. 다소 그렇지 않다
   e. 그렇지 않다
   f. 매우 그렇지 않다

7. TOEIC 공부를 하고 난 후 영어로 더 쉽게 일상적인 대화를 할 수 있게 되었다.
   a. 매우 그렇다
   b. 그렇다
   c. 조금 그렇다
   d. 다소 그렇지 않다
   e. 그렇지 않다
   f. 매우 그렇지 않다

8. TOEIC 말하기 및 쓰기 시험이 더 자주, 더 가까이 있었다면 시험을 못 생각이 있다.
   a. 매우 그렇다
   b. 그렇다
   c. 조금 그렇다
   d. 다소 그렇지 않다
   e. 그렇지 않다
   f. 매우 그렇지 않다
9. TOEIC 공부를 해도 영어 설력이 빠르게 높고 있는 것 같지 않다.
   a. 매우 그렇다  d. 다소 그렇지 않다
   b. 그렇다       e. 그렇지 않다
   c. 조금 그렇다  f. 매우 그렇지 않다

10. TOEIC 공부를 통해 영어 회화 설력이 향상될 것이라고 생각한다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

11. 높은 TOEIC 점수를 받는 것이 구직 활동에 크게 도움이 될 것이라고 생각한다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

12. 높은 TOEIC 점수를 받는 것이 학업 활동에 크게 도움이 될 것이라고 생각한다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

13. 높은 TOEIC 점수를 받는 것이 개인적으로 크게 도움이 될 것이라고 생각한다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

14. 향후 내가 가질 직장에서 영어가 종종 필요할 것이다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

15. 영어 회화 공부보다 TOEIC 공부가 더 좋다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

16. TOEIC 공부를 통해 영어에 흥미를 갖게 되었다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다

17. 영어 자체를 배우는 것보다 TOEIC 점수를 잘 받는 것이 더 중요하다.
    a. 매우 그렇다  d. 다소 그렇지 않다
    b. 그렇다       e. 그렇지 않다
    c. 조금 그렇다  f. 매우 그렇지 않다
APPENDIX C
Student Qualitative Survey

1. When you take the TOEIC, are you planning on taking the Speaking and Writing sections? Why or why not?
TOEIC 시험을 볼 때 말하기 및 쓰기 영역을 볼 생각이십니까? 이유를 설명해 주세요.

2. Do you think studying for the TOEIC Listening and Reading sections will improve your English conversation skills? Why or why not?
TOEIC의 듣기 및 읽기 영역을 공부하면 영어로 대화하는 능력이 향상될 거라고 생각하십니까? 이유를 설명해 주세요.

3. Do you find the TOEIC test useful for studying English? Why or why not?
TOEIC 시험이 영어를 공부하는 데 도움이 된다고 생각하십니까? 이유를 설명해 주세요.

4. What aspects of studying for the TOEIC, if any, did you find the most useful? Why?
영어 실력을 높이는 데 TOEIC 시험 공부가 어떤 면에서 가장 많이 도움이 되었습니까? 이유를 설명해 주세요.

5. What aspects of studying for the TOEIC, if any, did you find the least useful? Why?
영어 실력을 높이는 데 TOEIC 시험 공부가 어떤 면에서 가장 적게 도움이 되었습니까? 이유를 설명해 주세요.
6. What are the three main reasons you are studying for the TOEIC test?
   TOEIC 시험 공부를 하는 이유 세 가지를 기술해 주세요.

7. If studying for the TOEIC was not required by the university or future employers, would you still do it? Why or why not?
   대학이나 구직에 TOEIC 점수가 필요 없어도 TOEIC 공부를 계속 하시겠습니까?

8. Does studying for the TOEIC make you more or less interested in studying English? Why?
   TOEIC 공부를 함으로써 영어 공부에 흥미가 늘거나 줄었습니까? 이유를 설명해 주세요.

9. How do you plan to continue using English, if at all, after you take the TOEIC test?
   TOEIC 시험을 본 이후에도 영어를 계속 사용할 계획이라면 어떻게 사용할 계획입니까?

10. If you get a low score on the TOEIC test, would you be more likely to stop studying English or study more? Why?
    TOEIC 점수가 낮게 나온다면 영어 공부를 중단할 것 같습니다, 아니면 더 열심히 공부할 것 같습니까? 이유를 설명해 주세요.

11. If you get a high score on the TOEIC test, would you be more likely to stop studying English or study more? Why?
    TOEIC 점수가 높게 나온다면 영어 공부를 중단할 것 같습니다, 아니면 더 열심히 공부할 것 같습니까? 이유를 설명해 주세요.
## APPENDIX D

Student Quantitative Survey Results

<table>
<thead>
<tr>
<th>#</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>26</td>
<td>19</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>60</td>
<td>12</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>32</td>
<td>31</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>32</td>
<td>78</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
<td>38</td>
<td>15</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>43</td>
<td>32</td>
<td>11</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>9</td>
<td>15</td>
<td>22</td>
<td>33</td>
<td>17</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>30</td>
<td>18</td>
<td>16</td>
<td>18</td>
<td>5</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>39</td>
<td>19</td>
<td>14</td>
<td>11</td>
<td>2</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>36</td>
<td>21</td>
<td>11</td>
<td>22</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>35</td>
<td>37</td>
<td>12</td>
<td>6</td>
<td>11</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>39</td>
<td>42</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>37</td>
<td>40</td>
<td>16</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>55</td>
<td>33</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>10</td>
<td>16</td>
<td>23</td>
<td>30</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>32</td>
<td>22</td>
<td>31</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>20</td>
<td>17</td>
<td>13</td>
<td>15</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>#</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>17</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>5</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>9</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>10</td>
<td>5</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td></td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>7</td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group A</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
</tbody>
</table>
**APPENDIX E**

Classroom Observations

Professor A  
Class size: 42 students  
Class duration: 75 minutes  
Students’ Year: 1st

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time</th>
<th>Details</th>
</tr>
</thead>
</table>
| Introduction     | 3 min| Opening announcements  
Student directions (open book, no speaking Korean, etc.) |
| Quiz             | 20 min| Dictation (12 sentences)  
All materials from previous week’s lesson |
| Worksheet        | 5 min| Practising vowel sounds (ex. pull pole, pool, pail)  
Class reading in unison |
|                  | 5 min| Pair work  
Reading aloud and listening for comprehension / recognition |
| Class book       | 10 min| Reading vocabulary from the chapter in unison  
Explaining difficult pronunciation  
Explaining meaning in English  
(Korean definition in book) |
|                  | 10 min| Listening comprehension  
Practice section from TOEIC test |
| PowerPoint       | 10 min| Explaining answers to practice test  
Showing script while listening to playback |
| Pair work        | 7 min| Using vocabulary to make new sentences |
| Homework         | 5 min| Explaining homework for next week |
| Student Reactions|      | All students stayed awake  
Wide use of L1 during tasks  
Generally, students in the back of the classroom did not pay attention or participate in tasks |
Professor B  
Class size: 46 students  
Class duration: 75 minutes  
Students’ Year: 1st

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2 min</td>
<td>Opening announcements</td>
</tr>
</tbody>
</table>
| Group quiz       | 15 min| Dictation (single words)  
Students answer verbally  
Must be first student to answer correctly  
Students answer in L1 or L2 depending on question  
All materials from previous week’s lesson |
| Lecture          | 5 min| Description of TOEIC listening test section 1  
Explaining test taking strategies |
| Class book       | 5 min| Describing pictures on first page as a class  
Students provide short descriptions at teacher prompting |
|                  | 10 min| Describing pictures on second and third pages with a partner |
|                  | 5 min| Describing pictures on second and third pages as a class |
|                  | 10 min| Listening comprehension  
Practice section from TOEIC test |
| PowerPoint       | 10 min| Explaining answers to practice test |
| Quiz             | 10 min| Written vocabulary quiz |
| Homework         | 3 min| Explaining homework for next week |
| Student Reactions|      | All students stayed awake  
Wide use of L1 during tasks  
Generally, students in the back of the classroom did not pay attention or participate in tasks |
Professor C  
Class size: 34 students  
Class duration: 75 minutes  
Students’ Years: 3rd and 4th

<table>
<thead>
<tr>
<th>Activity</th>
<th>Time</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3 min</td>
<td>Opening announcements</td>
</tr>
</tbody>
</table>
| Quiz      | 10 min | Dictation (20 sentences)  
All materials from previous 2 weeks’ lessons |
| PowerPoint | 25 min | Explaining test taking strategies |
| Class book | 8 min  | Listening comprehension  
Practice section from TOEIC test |
| PowerPoint | 8 min  | Explaining answers to practice test |
|           | 6 min  | Dictation review |
| Discussion | 2 min  | Discuss with partner to find out with which questions each had difficulty |
| Lecture   | 13 min | Further explain questions with which students had difficulty |
| Student Reactions | 12 students slept through some or all of class  
Students listened to class through headphones, so no use of L1 in class outside of 2 min discussion section  
No use of L2 speaking or writing |
REFERENCES


ETS (2011). Global use of the TOEIC Tests Continues to Increase as a Record-Breaking Six Million Tests were Administered in 2010. ETS. [online] available from:


