English Language & Literature Teaching, Vol.27, No.3 Fall 2021
https://doi.org/10.35828/etak.2021.27.3.47

Using TOEIC as an English Assessment Tool: A Retrospective Case Study of 'A' University

Seung Won Jun
(R.O.K Air Force Academy)

Jun, Seung Won (2021). Using TOEIC as an English assessment tool: A retrospective case study of 'A' university. *English Language & Literature Teaching, 27*(3), 47-64.

The study investigated the long-term use of TOEIC by one university to measure its students' English proficiency, examining (a) the changes in the students' TOEIC scores across years, (b) the effectiveness of the intensive TOEIC programs run by the university to help the students prepare for TOEIC, and (c) the degree in which TOEIC was able to measure the students' communicative skills in English. The results showed that the students' TOEIC scores had a tendency to gradually increase each year before 2018, after which the scores started to decrease. The decrease may be due to the removal of English Conversation courses from the curriculum in 2018, which resulted in the absence of English courses for the students after their freshman year. Also, the intensive TOEIC programs had positive short-term effects on improving the students' score after the completion of the programs. However, students were able to demonstrate similar increases in their scores after a 12-month period both with and without the intensive programs. Finally, TOEIC was not able to reflect the students' communicative skills well, nor did the students perceive TOEIC as doing so. Based on the results, pedagogical implications for English instruction and assessment are also discussed.

[effects of TOEIC/intensive TOEIC program/language testing/validity]

I. INTRODUCTION

Many universities in Korea require their students to meet a certain level of English proficiency to fulfill a course requirement, graduate, apply for scholarship, enter a graduate program, etc. Students normally prove their English proficiency level by submitting an official score report of Test of English for International Communications (TOEIC) or other similar high-stakes tests, such as Test of English as a Foreign Language (TOEFL). Some

universities offer incentives to students with high test scores such as by providing exemptions from enrolling in certain English courses.

To help their students prepare for and achieve high scores in the test, universities also include TOEIC courses as part of their curriculum. In fact, many universities have actually started to replace traditional English courses with the more practical TOEIC-related courses to equip their students with the necessary skills as potential job applicants (K-R. Kim, 2014; E. Lee & C. Kim, 2014; H-S. Park & Choi, 2013; Shin & Oh, 2012). In addition, universities may run special intensive TOEIC programs for several weeks to provide an environment in which students can focus solely on preparing for TOEIC.

With such importance placed on attaining necessary English skills, language tests such as TOEIC become extremely high-stakes tests for students and naturally produce far-reaching washback effects at various levels for English teaching and learning (Cheng & Watanabe, 2004; Hughes, 2003; McNamara, 2001). Hence, a close examination of issues related to TOEIC test-taking and curricular development will be important for both educators and policymakers. More specifically, the following questions need to be addressed in consideration of the current situation of universities just mentioned above.

First, encouraging students to take TOEIC periodically by requiring them to meet a certain level of English proficiency will have an influence on the students' TOEIC performance. Therefore, analyzing the changes in students' TOEIC scores from a long-term perspective should be meaningful in itself. Moreover, examining whether any notable changes in the scores can be observed together with curricular changes seems necessary since students in Korea are exposed to English primarily in classrooms.

Another important point that needs to be considered is how effective intensive TOEIC programs can really be. Not surprisingly, students may be able to increase their test scores after a few weeks of intensive learning and studying for the test. However, students who do not participate in these intensive programs may also be able to demonstrate comparable increases in their test scores within a longer-term period.

Finally, it is important to take into account how much TOEIC can actually measure the communicative skills of students. Although TOEIC is widely used to measure the general English proficiency level of test-takers, alternative tests, such as TOEIC Speaking and Writing, now exist for measuring the communicative aspect of language skills.

This study attempted to examine the above three issues more closely, looking back on the long-term use of TOEIC by one university in Korea (referred to as 'A' University hereafter) to measure the English proficiency of its students. 'A' University has used TOEIC as its English proficiency assessment tool for the past 20 years, having recently decided to replace the test with TOEIC Speaking in order to measure the incoming students' communicative skills more accurately. Therefore, it seemed timely and appropriate to embark on this endeavour and draw out implications that could help educators and policymakers make

better informed decisions in the future.

More specifically, the three research questions for the current study were:

1. How have the students' TOEIC scores changed across the years from 2012 to 2020?
2. How effective were the intensive TOEIC programs for increasing the students' TOEIC scores in the years 2013 and 2014?
3. How much was TOEIC able to measure the communicative skills of the students?

II. PREVIOUS STUDIES

According to the Educational Testing Service, TOEIC was developed in 1979 and has been used worldwide by more than 14,000 organizations in more than 160 countries to measure the English proficiency of non-native speakers required to work in an international environment. The test was first introduced to Korea in 1982 by YBM and is being administered to approximately seven million test-takers every year, who have to complete two sections (i.e., the Listening and Reading Comprehension sections) of the test in 120 minutes.

Considering the continued popularity of the test in Korea, not many studies in TOEIC have been reported. Chun and Ko (2020) provide a summary of TOEIC-related research published in Korean academic journals since 2005. According to their report, most studies have been quantitative research published within the last decade with a sample size of less than 1,000, which consists solely of university students. Of the 55 studies they reviewed, 35 of them focus on the teaching and learning of TOEIC, examining the effectiveness of online learning, certain types of teaching method, intensive TOEIC programs, etc.; 20 of them focus on the learner, examining learner strategies that are used or learners' attitude, perception, motivation, and self-regulated learning with regard to TOEIC.

Studies that report on the changes in students' TOEIC scores from a long-term perspective (i.e., related to research question 1) are lacking in the literature; most have examined changes in students' TOEIC scores after relatively short-term interventions that lasted, for example, a semester or two. In general, interventions in the form of TOEIC instruction had positive effects on improving students' TOEIC scores (e.g., K-R. Kim, 2014; I-Q. Lee, 2014; Shin & Oh, 2012). It is worth noting that TOEIC instruction can be more effective when students are enrolled in courses that promote the enhancement of listening skills, such as English conversation, at the same time (Shin & Oh, 2012).

In addition, studies show that interventions that are more focused on improving either TOEIC listening or reading scores can be effective. More specifically, having students perform a dictogloss listening activity regularly (i.e., an activity in which students take notes

individually as they listen to a text and then reconstruct the text in small groups) (Joo, 2014) or read aloud listening texts that they have difficulty comprehending (Heo, 2016) can help the students improve their TOEIC listening scores. Similarly, engaging students in extensive reading (Jeon, 2015) or providing students with instruction that allow them to focus either on vocabulary or grammar to understand the readings texts better (M-K. Lee, 2011) can help them improve their TOEIC reading scores.

Studies about the effectiveness of intensive TOEIC programs (i.e., related to research question 2) all report positive short-term effects of the programs on increasing students' TOEIC scores. The increase in the students' scores after the completion of a program in comparison to their scores before the program varies (from the most to least increase): 148.3 after eight weeks of the program (Ha, 2014), 140 after three weeks (G. Kim, 2014), 120.5 after five weeks (E. Lee & C. Kim, 2014), 96.39 after eight weeks (Nam & J-Y. Kim, 2014), 70.11 after four weeks (T. K. Kim, 2018), 73 after four weeks (T. K. Kim & B. L. Kim, 2017), and 39 after three weeks (H-S. Park & Choi, 2013).

The programs in these studies ran from as short as three weeks to eight weeks during the summer or winter vacation, usually including more than one class based on the students' proficiency level. Two studies (T. K. Kim, 2018; E. Lee & C. Kim, 2014) provided result that were based on more than one program, each of which were run similarly at different times. Moreover, two of the four programs reported in E. Lee & C. Kim's (2014) study had classes that consisted of students with mixed proficiency levels. Most studies included an analysis of the students' TOEIC scores based on their proficiency level, which showed a general tendency of students with lower proficiency levels (i.e., normally below a score of 500) to demonstrate the greatest increase in their scores after the programs.

Studies also looked into other factors that can have an influence on TOEIC scores, such as students' satisfaction with the program (Ha, 2014; T. K. Kim, 2018), attendance rate and GPA (G. Kim, 2014), as well as grammar and vocabulary knowledge (T. K. Kim & B. L. Kim, 2017). Both Ha (2014) and T. K. Kim (2018) found no meaningful relationship between students' satisfaction level of the programs and their increase in scores; G. Kim (2014) found that both attendance rate and GPA had a significant effect on the increase in scores; and T. K. Kim and B. L. Kim (2017) discovered that vocabulary knowledge had a higher correlation with both the listening and reading comprehension section scores than grammar knowledge.

To examine the validity issues regarding TOEIC (i.e., related to research question 3), studies have calculated correlations between TOEIC and other productive skills-based tests. For example, I-S. Lee (2006) found the correlation between TOEIC and Spoken English Proficiency Test (SEPT) to be 0.839. In S. J. Lee's (1999) study, the correlation between TOEIC and SEPT was 0.74, and 0.88 for TOEIC and Test of Written English (TWE). Moreover, H-W. Lee and G-P. Park (2003) reported the correlations between TOEIC and

TOEFL, Multimedia Assisted Test of English (MATE), and the English test from KSAT to be 0.79, 0.55, and 0.6, respectively. In Y. Lee's (2020) study, the correlation between TOEIC and the students' scores from a course taught by native English speakers was 0.45. Finally, the correlations between TOEIC and the students' performance in English conversation and composition courses were 0.28 and 0.21, respectively, in H. Lee and S. Lee's (2004) study.

Studies have also investigated the validity of TOEIC based on students' perception of the test. They show that students do not perceive the increase in TOEIC scores as indicating an improvement in their actual English proficiency (K-R. Kim, 2014); that although students perceive TOEIC as a tool capable of enhancing their English skills, they do not expect that preparation in TOEIC alone will be able to do so (D. Lee, K. H. Hong, & K. Hwang, 2013); that as students advance in their university years, they regard TOEIC more as a social requirement than as a tool for developing their English proficiency (D-M. Lee, K-H. Hong, & K-H. Hwang, 2013); and that students do not believe TOEIC to be an objective tool for measuring English proficiency, even though they consider it to play a crucial role in obtaining employment (Ryu, 2017).

Overall, the studies in the literature show that there were normally positive changes in students' TOEIC scores after interventions in the form of TOEIC instruction; that intensive TOEIC programs were effective; and that while correlations between TOEIC and other productive skills-based tests varied across studies, students did not have a positive perception of the test regarding its validity.

III. METHODS

1. Background Information of the Study

1) English Curriculum of 'A' University

Before 2018, the students at 'A' University normally took one two-hour English course each semester, starting with College English 1 and 2 in their freshman year and continuing with English Conversation 1 to 4 in their sophomore and junior years. The senior students took an English for Specific Purpose (ESP) course titled Military English for one semester in their final year. These English courses were not necessarily related to TOEIC, and the students were encouraged to study for TOEIC on their own. Only the students who did not meet the university's minimum proficiency level took an extra non-credit TOEIC course in their senior year.

College English 1 and 2 courses were focused on developing the students' overall English proficiency. More specifically, they were designed to help the students enhance their reading

comprehension skills and acquire the lexical and grammatical knowledge of English through analyses of several reading articles, about eight to ten paragraphs long each, on various topics. In addition, speaking and writing tasks (e.g., a writing assignment followed by an oral presentation) were included to help the students with their productive skills.

English Conversation courses, on the other hand, had more emphasis on helping the students hone their communicative skills. From performing simple language functions (e.g., asking questions, requesting, inviting, etc.) to engaging in connected discourse (e.g., producing narratives, supporting opinions, debating, etc.), the students were able to become familiar with various real world language tasks as they progressed through their sophomore and junior years. Four native English-speaking professors taught the courses, each class size being small – normally, less than ten students – to maximize each student's opportunity for speaking. At the end of each course, one-on-one interviews with the students were conducted by the four professors individually.

In 2018, all English Conversation courses were removed from the curriculum. Hence, the students currently take only College English 1 and 2 in their freshman year and one ESP course in their senior year. Instead of taking mandatory English Conversation courses in their sophomore and junior years, they are now encouraged to use their free hours during the day to visit the native English-speaking professors and have short one-on-one conversation sessions at their own disposal. These sessions are not necessarily structured, and the students can use the time to get help on any aspect of English depending on whatever their needs are. Since participation in these sessions is completely voluntary, the students with high motivation to study English mostly tend to take advantage of the sessions.

2) TOEIC and Intensive TOEIC Programs

'A' University has used TOEIC to assess the English proficiency of its students since 2000, as well as set a minimum proficiency level for the students to achieve before graduation. Over the years, both the official TOEIC and mock TOEIC had been used, and the targeted minimum score was adjusted several times. The most recent minimum score was set to 730 in 2013, only the official scores being accepted by the university since 2011. The students have been required to take TOEIC - or other similar English proficiency tests such as TOEFL, Test of English Proficiency developed by Seoul National University (TEPS), etc. — at least once a year. With the emphasis on enhancing the students' actual communicative skills of English, the university changed its assessment tool from TOEIC to TOEIC Speaking in 2020. However, the students took the regular TOEIC in 2020 due to unforeseen circumstances arising from the COVID-19 pandemic.

In addition, 'A' University has offered TOEIC courses as part of the Winter English Program to help the students attain the minimum proficiency level set by the university as

early as possible. More specifically, the university ran the Winter English Program from 2001 to 2014 for approximately a month during the winter for the students each year before the Spring semester began. The courses for the program and the group of students they were offered to somewhat varied over the years, some example courses being English Conversation, Composition, Speech, Debate, and TOEIC.

In the 2013 and 2014 programs, only TOEIC courses were offered to the students who had not yet met the university's minimum proficiency level (referred to as under-achieving students hereafter). For these programs, the students (N = 208 and 148 in 2013 and 2014, respectively) were divided into different classes based on their TOEIC scores (four and six classes in 2013 and 2014, respectively), and professional TOEIC instructors recruited outside the university were responsible for teaching TOEIC six hours a day for four weeks. The students were required to take an official TOEIC shortly after the programs ended.

2. Data Collection and Analyses

The bulk of the data used for the current study consists of TOEIC scores of students at 'A' university, which were collected from 2012 to 2020. After undergoing a competitive selection process, approximately 200 students enter the university each year and pursue a major in one of eight academic disciplines. Regarding their English proficiency level upon entering the university, the average TOEIC score of the students in their freshman year from 2012 to 2020 was 748.8. Only the scores of the freshman, sophomore, and junior students who were enrolled at the university during this period were used for the study. The scores of the senior students were not analyzed because the sample size for them turned out to be too small most of the years. Although all the students were required to take TOEIC at least once a year, most students met the university's minimum proficiency level before their senior year, and only a small portion of the senior students resulted in taking the test.

To answer research question 1, the overall average scores of all the students who took TOEIC each year from 2012 to 2020 were first examined. The sample size for each year was 433, 339, 388, 407, 412, 476, 491, 512 and 454 respectively, adding up to a total of 3,912. Next, the average scores of these students according to their class (i.e., enrollment period) and grade (i.e., freshman, sophomore, and junior) in each year were examined. With regard to the students' class, only the scores from the class of students who were enrolled at the university for the full three years during the 2012 to 2020 period were analyzed.

To answer research question 2, the scores of TOEIC taken before and after the intensive TOEIC programs in 2013 and 2014 by the under-achieving students were analyzed. The sample size for the two years was 185 and 143, respectively. Furthermore, the changes in the test scores from the years when the university ran the programs were compared to those from the other years when the university did not run the programs.

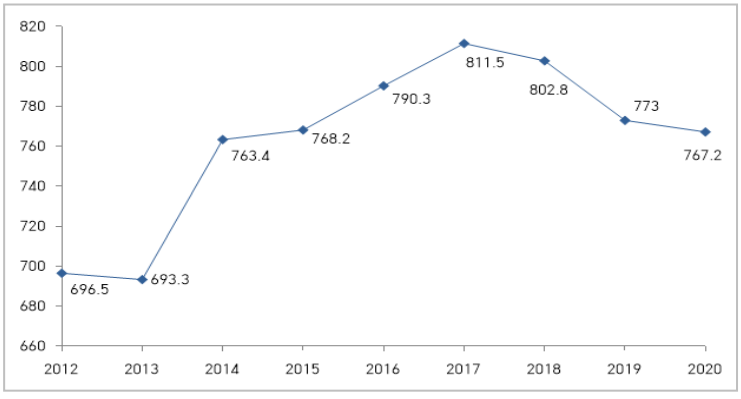
To answer research question 3, correlations between the interview scores from the English Conversation courses that three different classes of students took either in the Fall Semester of 2015 or 2016 and their TOEIC scores from the same years were calculated. The sample size for each class of students was 97, 95, and 117, respectively. The interviews were conducted one-on-one with the students by the four native English-speaking professors individually and lasted approximately 10 minutes each. Moreover, the questions for the interviews were tailored to and different for each student to elicit answers that required the students to apply what they had learned in class. Therefore, item and inter-rater reliabilities for the interviews were not calculable. In addition to the correlations, two open-ended questions (i.e., "How much does TOEIC reflect your actual English proficiency?" and "How much is studying for TOEIC helpful for studying English?") from the questionnaire used in Jun and Rah's (2015) study were analyzed. For these questions, similar responses from the students were grouped together, and their frequencies were tallied.

IV. RESULTS

1. Changes in the Students' TOEIC Scores

A graphic representation of the total average scores across the years is shown in Figure 1 (refer to appendix for a detailed summary of the students' average TOEIC scores from 2012 to 2020).

FIGURE 1
Total Averages Scores Across the Years



Moreover, the average scores across the years according to the students' class and grade are shown in Figures 2 and 3, respectively. In Figure 2, 'class1215' refers to the class of

students who were enrolled at the university from 2012 (freshman year) to 2015 (senior year), and so on. Therefore, Figure 2 illustrates the change in scores for the same group of students throughout their three years in university. On the other hand, Figure 3 shows the average scores of students at each grade level for a given year. Several observations can be made here.

FIGURE 2
Averages Scores According to Students' Class

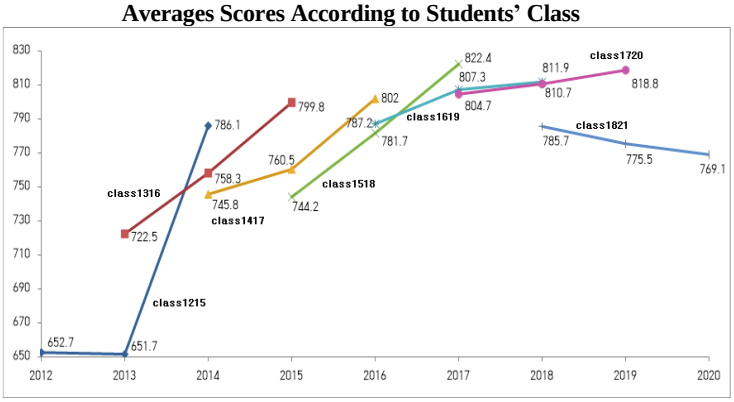
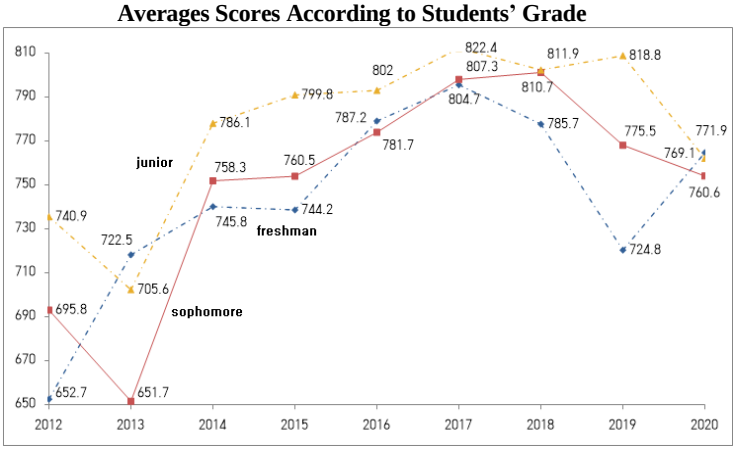


FIGURE 3
Averages Scores According to Students' Grade



First, Figure 1 illustrates that there was a gradual increase in the average scores across the years before 2018, with the exception of 2013 when there was a slight decrease of 3.2 points in comparison to the average score in 2012. On the contrary, the average scores showed a gradual decrease after 2018.

Second, examining the changes in scores according to the students' class (i.e., Figure 2),

we can see that there was a general tendency for each class of students to improve their scores each consecutive year until their junior years, with the exception of Class 1821, whose scores showed a gradual decrease across the three years. In addition, each class of students entering the university was shown to start out with a higher average score in their freshman year when compared to their previous classes, with the exception of Class 1518 and 1821.

Third, according to Figure 3, the average scores of the students for each grade level showed a tendency to increase across time before 2018. The increase seemed more consistent and apparent from 2013 to 2017; in 2013, the average scores of the sophomore and junior students dropped notably from the previous year. With the exception of the juniors in 2019 and of the freshmen in 2020, the average scores tended to gradually decrease after 2018.

2. The Effectiveness of the Intensive TOEIC Programs

Table 1 shows the increase of TOEIC scores for the under-achieving students after the intensive TOEIC programs in 2013 and 2014 and at the end of each year from 2013 to 2020. After four weeks of the programs in 2013 and 2014, the scores increased by 57.6 and 61.3 points, respectively. Furthermore, the total increase of the scores for all under-achieving students in the years 2013 and 2014 (in a period of 12 months inclusive of the programs) was 67.9 and 107.5 points, respectively. In comparison, the increase of scores for all under-achieving students from 2015 to 2020, when there were no intensive TOEIC programs to help them improve their scores, was 80, 96.4, 103.3, 64.4, 63.5, and 69.9 points, respectively. All the instances of increase in the scores were statistically significant ($p < 0.001$) based on paired *t*-tests.

TABLE 1
Increase of Average Scores for Under-achieving Students

	2013	2014	2015	2016	2017	2018	2019	2020
Increase of average score (after IEP)	57.6	61.3	-	-	-	-	-	-
N	185	143	-	-	-	-	-	-
Increase of average score (end of the year)	67.9	107.5	80	96.4	103.3	64.4	63.5	69.9
N	181	180	103	105	90	66	70	109

The results show that four weeks of the intensive TOEIC programs in 2013 and 2014 were indeed effective for helping the under-achieving students improve their TOEIC scores. Further improvements by the end of the respective years after the programs were also notable. However, there were comparable increases in the scores for all under-achieving students after a 12-month period without the programs in the years from 2015 to 2020. In other words,

although the intensive TOEIC programs were shown to be effective in helping the under-achieving students improve their scores in a relatively short time, they were also able to demonstrate a similar amount of improvement in a 12-month period, irrespective of the years with or without the programs.

3. The Use of TOEIC for Measuring Communicative Skills

The correlations between the students' TOEIC and their English interview scores are shown in Table 2. The correlations — which were calculated for three different classes of students — were 0.44, 0.32, and 0.27, respectively, with an overall correlation of 0.33. Although the correlations were all statistically significant, these values cannot be considered to be high.

TABLE 2
Correlations Between TOEIC and Interview Scores

Class	*N*	Semester	Correlation
Class 1316	97	2015 Fall Semester	0.44 ($p < 0.001$)
Class 1417	95	2016 Fall Semester	0.32 ($p < 0.01$)
Class 1518	117	2016 Fall Semester	0.27 ($p < 0.01$)
Total	309		0.33 ($p < 0.001$)

The students' responses to the two open-ended questions from the questionnaire used in Jun and Rah's (2015) study are shown in Table 3. A total of 225 students completed the questionnaire. Only the students' responses with a frequency of over five are listed.

TABLE 3
Students' Responses to the Open-ended Questions

How much does TOEIC reflect your actual English proficiency?		How much is studying for TOEIC helpful for studying English?	
reflects listening skills well	40	helpful for listening skills	34
reflects reading skills well	39	helpful for reading skills	32
doesn't reflect speaking skills well	43	helpful for vocabulary	15
doesn't reflect writing skills well	33	helpful for grammar	10
Performance is influenced by your skills in answering the test items.	14	helps enhance basic proficiency	10
It's just a test.	6	motivates you to study English	23
		isn't helpful for speaking skills	15
		isn't helpful for writing skills	8
		forces you to study for a test	8
		is unrelated to everyday English	7

Note. Numbers indicate the frequency of students' responses.

With respect to the first question, many students perceived that TOEIC reflects only the

two skills that the test directly measures (i.e., listening and reading skills) well (40 and 39 responses, respectively). Some students responded that TOEIC is just a test (6 responses) and test performance is influenced by your skills in answering the test items (14 responses). Regarding the second question, many students again perceived that studying for TOEIC is helpful for only listening and reading skills (34 and 32 responses, respectively). Some responded that the test itself motivates you to study English (23 responses), while others answered that it forces you to study for a test (8 responses) and is unrelated to everyday English (7 responses).

V. DISCUSSION

This study looked back on the long-term use of TOEIC by 'A' University and examined the changes in the student's TOEIC scores from 2012 to 2020 (i.e., research question 1), the effectiveness of the intensive TOEIC programs run by the university to help the students achieve higher scores (i.e., research question 2), and the degree in which TOEIC was able to measure the students' communicative skills in English (i.e., research question 3).

First, with regard to research question 1, the students' test scores in general were seen to gradually increase across years before 2018. The increase may have been partially due to the fact that the students entering the university each year were scoring higher than the students from the previous years. Regardless, the analyses also showed that there was a tendency for each class of students to increase their test scores throughout their three years in university.

On the contrary, the students' test scores showed a tendency to gradually decrease after 2018. Surprisingly, the students who entered 'A' University in 2018 scored lower each following year. The other classes of students before had displayed the opposite tendency. Moreover, the increase of scores each year for the under-achieving students after 2018 was relatively low when compared to the previous years.

These results have significant meaning in light of the fact that all English Conversation courses at the university were removed from the curriculum in 2018. Before, the conversation courses were obligatory for four semesters during the students' sophomore and junior years. Only the students with a TOEIC score of 800 or higher were exempt from one or more of the conversation courses (maximum three) depending on their scores. Although other factors could have influenced the students' test scores, the removal of the English Conversation courses in 2018 is the only notable difference regarding the university's English education policy or curriculum.

Therefore, the absence of regular English courses after the freshman year seems most likely to be accountable for the distinctive decrease in the students' TOEIC scores after 2018. With the English Conversation courses, the students had regular exposure to English

throughout the semesters, through which they had the opportunity to develop their English skills. After the removal of the courses, only a small portion of the students who were eager enough to visit the native English-speaking professors regularly during their free hours had exposure to English. The lack of opportunity to develop their English – particularly, listening and speaking – skills continuously over the years apparently seems to account for the gradual decrease in the students' TOEIC scores. Interestingly, Shin and Oh (2012) found that the increase in their students' TOEIC scores was primarily due to the score increase in the listening comprehension section of the test, claiming that providing students with courses related to enhancing listening skills, such as English Conversation, together with TOEIC instruction can produce better results for them.

Second, concerning the university's efforts to help the under-achieving students by running the intensive TOEIC programs (i.e., research question 2), the results showed that the programs were indeed effective for helping the students improve their test scores at the end of the four-week programs. However, from a longer-term perspective, the under-achieving students who were enrolled during the years when the university did not run the programs were able to achieve a similar amount of improvement in a 12-month period in comparison to those who were enrolled during the years when the university ran the programs.

The finding that the increase of the under-achieving students' TOEIC scores over a 12-month period was similar regardless of the years with or without the intensive TOEIC program is quite surprising. On the one hand, the programs were definitely able to help the students improve their scores in a relatively short period. As described in the Previous Studies section, all studies on intensive TOEIC programs were shown to be effective for improving students' TOEIC scores. On the other hand, the long-term effect of the programs in the current study was less notable. The university offered the programs to only the under-achieving students for a short time. Programs that last for a longer period with different proficiency levels of students may have yielded different results.

Finally, regarding research question 3, the overall correlation between the students' TOEIC scores and their English interview scores was 0.33, implying that TOEIC was not able to genuinely reflect the communicative skills of the students. This value is much lower than the correlations reported in other studies. Furthermore, many students in the current study also perceived that TOEIC reflects and is helpful for only the two skills that the test directly measures (i.e., listening and reading skills). There were several students who perceived that although TOEIC can motivate one to study English, it is just a test that doesn't reflect everyday English well and that its performance can be influenced by one's skills in answering the test items.

In Jun and Rah's (2015) study, the correlation between their students' TOEIC and English interview scores was a slightly higher value of 0.48. Moreover, the average scores of the students' responses to the questionnaire questions about how much the TOEIC reflected their

English proficiency and how much the TOEIC was helpful for studying English were 4.89 and 5.40 (10 maximum), respectively. Together, the two studies show that the validity of TOEIC for measuring communicative skills was somewhat questionable.

VI. CONCLUSION

Overall, the current study showed that (a) the students' TOEIC scores showed a general tendency to improve gradually from their freshman to junior years, but started to decrease with the removal of the English Conversation courses in 2018, (b) the university's efforts to help the under-achieving students by running the intensive TOEIC programs had positive short-term effects, and (c) TOEIC, nonetheless, was not able to reflect the communicative skills of the students well. Based on these results, several implications can be drawn from the study.

First, although the long-term use of TOEIC by 'A' University cannot be seen as a direct cause of the improvement in the students' TOEIC scores, this seems likely to have motivated the students to study English, since the university's policy on regular TOEIC test-taking and establishment of specific goals for the students to achieve show the university's emphasis on the importance of acquiring a certain level of language skills for its students. Actually, several students commented in their responses to the questionnaire that TOEIC motivated them to study English.

Second, this study does not in any way imply that intensive TOEIC instruction, such as in the form of intensive programs, is not effective. When increasing TOEIC scores in a relatively short time is the goal, intensive TOEIC programs can be definitely effective, possibly being one of the few options available. Furthermore, these types of programs can be helpful especially for students who are new to TOEIC and those who have difficulty preparing for the test on their own. Rather, this study implies that students may be able to demonstrate a similar increase in test scores with enough time and motivation. Motivation to study for TOEIC and learn English in the present study could have been possible through the university's policy regarding TOEIC, its efforts to help students meet the minimum proficiency level, and the regular English courses offered as part of the curriculum.

Third, what should not be ignored from the present study is that regular English instruction throughout the semesters in which students are continuously exposed to an English learning environment can be paramount for developing their English proficiency. The removal of English Conversation courses at 'A' University left the students with no remaining English courses after their freshman year. The ESP course in the senior year does not necessarily focus on developing the students' general English skills. Therefore, it seems that entrusting the students to study English on their own, even with the university's policy on regular

TOEIC test-taking, was not enough to help the students maintain, much less improve, their English proficiency throughout their university years.

Finally, because of the washback effects that a language test can have on the teaching and learning of a language at various levels, educators and policymakers should be prudent about the language tests they adopt to use as an assessment tool. The current study showed that TOEIC was not able to reflect the communicative skills of the students well. Even more, the students' perception regarding the validity and helpfulness of TOEIC was found to be somewhat unconvincing. Although the four skills in English are strongly related, tests such as TOEIC will have limitations in measuring all the skills. Therefore, testing the four skills together will be able to provide a more accurate and complete picture of one's English proficiency (Liu & Costanzo, 2013; Powers, 2010).

Nonetheless, I believe that TOEIC served its purpose at 'A' University by helping the students develop their receptive skills in English and vocabulary knowledge when English Conversation courses mainly comprised the English curriculum before 2018, through which the students had been able to develop their productive skills (Jun & Rah, 2015). Needless to say, a well-balanced English education that focuses on all four skills is important. Since the university had been focusing on the productive skills through its curriculum, a listening- and reading-based test had been able to motivate the students not to neglect the receptive skills.

There are several limitations of the study that need to be mentioned. First, the results of the study were based on data collected from students at one university in Korea. Therefore, the study lacks generalizability, and future studies with other students from various universities would help to interpret the findings of this study better. Second, only the under-achieving students participated in the 2013 and 2014 intensive TOEIC programs. The effectiveness of these programs could have been different for the other students. With the literature overwhelmingly showing the positive short-term effects of intensive TOEIC programs, it would be interesting for future studies to examine the effectiveness of such programs from a longer-term perspective. Finally, the English interviews were conducted one-on-one by four different native speakers of English. Although they used the same rubric to rate the students, the interviews were conducted individually. Including a second or third rater in the interviews could have allowed more consistent rating, which could have produced higher correlations with the TOEIC scores.

REFERENCES

Cheng, L., & Watanabe, Y. (Eds.). (2004). *Washback in language testing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Chun, E., & Ko, M. (2020). An analysis of the articles related to TOEIC teaching or studying in Korean academic journals. *The Journal of Humanities and Social Science, 11*(2), 257-270.

Ha, M. (2014). Management effects of ability grouping based intensive TOEIC program. *The Jungang Journal of English Language and Literature, 56*(4), 609-630.

Heo, S-Y. (2016). The effect of reading aloud on listening comprehension: Focusing on listening text type and learners' proficiency level. *Studies in Humanities, 51,* 61-77.

Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.

Jeon, J. (2015). The impact of extensive reading in a formal English reading class on the improvement of TOEIC reading scores. *Journal of Korea Air Force Academy, 66*(2), 1-16.

Joo, M. (2014). The effects of dictogloss listening activity on each part of TOEIC LC and learning attitude: Focusing on low proficiency level learners. *English21, 27*(2), 273-291.

Jun, S., & Rah, Y. (2015). A study on the validity of TOEIC and its contribution to the learning of English from cadets' perception. *Journal of Korea Air Force Academy, 66*(1), 27-43.

Kim, G. (2014). The effects of an intensive TOEIC program for college students focusing on factors of participants' progress. *English Language & Literature Teaching, 20*(1), 311-327.

Kim, K-R. (2014). College student perceptions on the effects and necessity of TOEIC classes as liberal arts education courses. *The Journal of Linguistic Science, 69,* 63-86.

Kim, T. K. (2018). A case study of the short-term TOEIC program. *Journal of Language Sciences, 25*(2), 17-39.

Kim, T. K., & Kim, B. L. (2017). The effects of a short-term TOEIC camp for college students on the improvement of TOEIC score: Case study of Y'sU. *Journal of Language Sciences, 24*(4), 15-32.

Lee, D., Hong, K. H., & Hwang, K. (2013). A study on Korean college students' perception of TOEIC. *The New Korean Journal of English Language & Literature, 55*(2), 233-248.

Lee, D-M., Hong, K-H., & Hwang, K-H. (2013). A study on Korean college students' attitude toward TOEIC and their TOEIC score. *The Journal of the Humanities, 68,* 107-124.

Lee, E., & Kim, C. (2014). Educational effects of intensive English TOEIC program for university students. *Studies in English Language & Literature, 40*(2), 215-230.

Lee, H., & Lee, S. (2004). Validity of TOEFL, TOEIC and TEPS as a measure of communicative competence. *English Language & Literature Teaching, 10*(3), 191-210.

Lee, H-W., & Park, G-P. (2003). The reliability and validity of TOEIC. *English Teaching, 58*(4), 27-42.

Lee, I-Q. (2014). The effects of ability grouping on academic achievement in university TOEIC classes. *Journal of Language Sciences, 21*(4), 149-166.

Lee, I-S. (2006). The effectiveness of TOEIC scores on English oral proficiency. *Modern English Education, 7*(1), 33-52.

Lee, M-K. (2011). The effects of grammar-focused and vocabulary-focused teaching in TOEIC reading classes. *Modern Studies in English Language & Literature, 55*(2), 155-177.

Lee, S. J. (1999). A validation study on TOEIC. *Korean Journal of Psychology: Social Issues, 5*(2), 49-58.

Lee, Y. (2020). Analysis of the construct validity of the TOEIC test when used as an English proficiency evaluation tool. *English Language Assessment, 15*(2), 149-167.

Liu, J., & Costanzo, K. (2013). *The relationship among TOEIC listening, reading, speaking, and writing skills*. Retrieved March 24, 2021, from the World Wide Web: https://www.ets.org /Media/Research/pdf/TC2-02.pdf

McNamara, T. (2001). *Language testing*. Oxford: Oxford University Press.

Nam, J-M., & Kim, J-Y. (2014). A case study of short-term intensive English program. *Journal of Language Sciences, 21*(2), 69-88.

Park, H-S., & Choi, M-O. (2013). The effects of intensive English Camp for college students on the improvement of TOEIC score. *Studies in Linguistics, 30,* 149-166.

Powers, D. E. (2010). *The case for a comprehensive, four-skills assessment of English language proficiency*. Retrieved March 24, 2021, from the World Wide Web: https://www.ets.org /Media/Research/pdf/TC-10-12.pdf

Ryu, R. (2017). A comparison of perceptions toward TOEIC between learners and teachers: Focusing on the differences by learners' proficiency level. *The Journal of Studies in Language, 32*(4), 733-754.

Shin, K-I., & Oh, C-S. (2012). The effects of TOEIC teaching as a compulsory course in a college. *Studies in Linguistics, 24,* 125-137.

APPENDIX

Average TOEIC scores of the students from 2012 to 2020

Year	Grade	Class	*N*	Score	*SD*	Year	Grade	Class	*N*	Score	*SD*
2012	Freshman	1215	158	652.7	135.3	2017	Freshman	1720	173	804.7	129.4
	Sophomore	1114	148	695.8	103.3		Sophomore	1619	154	807.3	113.6
	Junior	1013	127	740.9	94.1		Junior	1518	149	822.4	87.9
	Total		433	696.5	118.9		Total		476	811.5	112.7

2013	Freshman	1316	157	722.5	128.0	2018	Freshman	1821	189	785.7	96.7
	Sophomore	1215	108	651.7	102.5		Sophomore	1720	148	810.7	104.0
	Junior	1114	74	705.6	92.5		Junior	1619	154	811.9	105.1
	Total		339	693.3	117.1		Total		491	802.8	102.1
2014	Freshman	1417	143	745.8	134.3	2019	Freshman	1922	196	724.8	152.7
	Sophomore	1316	114	758.3	101.3		Sophomore	1821	153	775.5	122.0
	Junior	1215	131	786.1	73.7		Junior	1720	163	818.8	96.7
	Total		388	763.4	108.3		Total		512	773.0	133.6
2015	Freshman	1518	156	744.2	141.0	2020	Freshman	2023	177	771.9	126.6
	Sophomore	1417	122	760.5	129.5		Sophomore	1922	181	760.6	146.9
	Junior	1316	129	799.8	106.9		Junior	1821	96	769.1	126.6
	Total		407	768.2	129.4		Total		454	767.2	134.9
2016	Freshman	1619	145	787.2	114.5	-					
	Sophomore	1518	141	781.7	105.8						
	Junior	1417	126	802.0	109.9						
	Total		412	790.3	110.2						

Note. 'Class1215' refers to the class of students who were enrolled at the university from 2012 (freshman year) to 2015 (senior year), and so on.

Examples in: English
Applicable Languages: English
Applicable Levels: Tertiary

Seung Won Jun(Professor)
Dept. of Languages and Literature
R.O.K. Air Force Academy
635 Danjae Ro, 335-1, Sangdang Ku, Cheongju, South Korea
Email: seungjun@mnd.go.kr

Received: July 15, 2021
Reviewed: August 15, 2021
Revised version received: September 4, 2021