

LEARNERS' PERSPECTIVES OF FACTORS INFLUENCING GAINS IN STANDARDIZED ENGLISH TEST SCORES

Yi-Ching Pan

(huangpan63@yahoo.com.tw)

National Pingtung University

4-18 Minsheng Rd., Pingtung City, Pingtung County 90003, Taiwan, R.O.C.

Abstract: There has been an increased level of attention devoted to the consequences of test use in recent years; however, the majority of washback studies focused on teaching. In fact, little research has addressed learners' perspectives to analyze possible determinants of test results. To address this issue, this study first compared the pre-and-post standardized English tests of two groups of Taiwanese university students, and six students from each of the groups were interviewed at a later date to investigate the possible factors that may influence their score gains. A control group of 140 Taiwanese university students at a school without any English proficiency certificate exit requirement was compared to a contrast group of 136 similar students at a school that required the students to pass an English certification test in order to graduate. The major finding indicates that the amount of time spent on language learning plays an essential role in determining the degree to which student scores improve. This study subsequently offers pedagogical implications for the instruction of English at institutes of tertiary education.

Keywords: standardized tests, test results, score gains, graduation requirement

DOI: <http://dx.doi.org/10.15639/teflinjournal.v27i1/63-81>

There has been an increased level of attention devoted to the consequences of test use in recent years; however, the majority of washback studies focused on

teaching (Cheng, 2008, 2014). For example, a number of studies on test consequences have taken an educational perspective and focused on teachers in various contexts, such as the US (Alderson & Hamp-Lyons, 1996; Stecher, Chun, & Barron, 2004); Australia (Burrows, 1998, 2001, 2004); Europe (Oerke, Merki, Holmeier, & Jager, 2011; Smyth & Banks, 2012; Wall & Horak, 2006); China (Cheng, 2004, 2005); Japan (Watanabe, 1996, 2004), and Taiwan (Pan, 2011). In fact, a far more limited number of studies have addressed learners' perspectives to analyze possible factors that may contribute to their test results. Wall (2000, p. 502) contends, "What is missing... are analyses of test results which indicate whether students have learnt more or learned better because they have studied for a particular test." Wall and Horak (2006) also claim that much of the research regarding washback or impact focuses less on investigating the test effects on products than teaching practices, that is, the outcome or quality of student learning. Since learners are the major test stakeholders, this study aims to investigate whether English certification exit requirements in Taiwan have significantly enhanced their listening and reading scores of the most common test of EFL proficiency in Taiwan, the Elementary Level General English Proficiency Test. In particular, it intends to explore from learner perspectives regarding the possible factors influencing the gains in their test scores.

Mixed findings have been found from the limited number of studies that investigated test effects on learning outcomes. These findings are discussed below.

Hughes (1988) reported that at a Turkish university, students' performance on the Michigan Test (a measure of English proficiency) increased after the introduction of a new test, along with additional summer courses in English. Based on a need analysis of what language skills students in their first year of undergraduate study required, the new performance test was designed to assess students' English proficiency to decide whether after a year of study at the university, they could continue to study. If they failed the test, they had to leave the university because it was presumed that their unsatisfactory English proficiency would make it difficult for them to understand the lectures, which were primarily conducted in English. The research revealed that before the introduction of the test, usually less than 50% of students in the Foreign Languages School had scores on the Michigan Test that qualified them to enroll in certain subjects. However, after the introduction of the test, 72% of the students

reached the requirement, and after the summer school, the figure went up to 83%.

The factors attributed to the score gain, according to Hughes, are the changes made for the new test such as the new syllabus, textbooks and pedagogy, which were designed to match the test content. These changes suggest that test scores can be improved if what is taught is what is tested, but it would be naïve and simplistic to attribute score gain mainly to teaching factors.

Andrews, Fullilove, and Wong (2002) compared the scores of three groups of students on the use of English oral exam in Hong Kong from 1993 to 1995. The first and third groups had test-driven instruction in their first and second years respectively. The second group received no test-preparation instruction. The scores of the students in the first and third groups tended to increase, but not to statistically significant levels. These researchers claimed that students' improved proficiency might be attributed to their "familiarization with the exam format, the rote-learning of exam-specific strategies and formulaic phrases" (p. 220).

Elder and O'Loughlin (2003) examined the relationship between intensive English language study and band score gains on the IELTS after 10-12 weeks of instruction. They found significant gains in listening, but no significant progress in reading skills. In their study, a range of factors were linked to improving scores on tests, such as personality, motivation, confidence, and exposure, but initial score level was the strongest predictor. Students with low scores at the onset exhibited more score gains than students with high scores at the onset. The fact that there was a test preparation course was not influential.

Green (2007a, 2007b) investigated whether test preparation classes helped students improve their IELTS writing scores. He found "no clear advantage for focused test preparation" (2007b, p. 75) in terms of grammar/vocabulary test performance among different groups in a 4-14 week intensive IELTS preparation course. However, score gains were found primarily among two groups of learners: those who planned to take the test again and those who had low initial writing test scores. These findings have two implications: first, as indicated by Green (2007a), test-driven instruction does not necessarily raise students' scores; second, students' motivation plays an important role in increasing their test scores.

Both Elder and O'Loughlin's (2003) study, as well as Green's (2007a, 2007b), suggest that students' original proficiency plays a more important role in affecting score gain than the time they spend in test-preparatory courses. It

appears that the lower the proficiency of the students, the more likely it is that they will receive a score gain.

Test effects on learning outcomes are an intricate issue, as is evident from the foregoing discussion. Due to a lack of research in this area, one may be tempted to question whether 1) test-related instruction and practice or 2) other factors such as their original proficiency, personality, motivation, and the length of exposure to the target language are greater determinants of learning outcomes.

In order to motivate their students to enhance their English proficiency, 90% of four-year technical universities/colleges in Taiwan have established an English certification exit requirement policy for non-English majors (Pan & Newfields, 2012). According to this policy, students must choose from an array of external English proficiency tests such as the international tests -TOEIC[®], TOEFL[®], IETLS[™], or the local tests - the General English Proficiency Test (GEPT), and the College Student English Proficiency Test (CSEPT) and reach a certain level or score in order to graduate. Since 2002, the GEPT has existed at five different levels: Elementary, Intermediate, High Intermediate, Advanced, and Superior. It is designed to test all four language skills: listening, speaking, reading, and writing. The Elementary Level, which is required by the two recruited four-year technical universities, is thought to correspond to the Council of Europe's A2 Waystage level (Taiwan Ministry of Education, 2004). Given the 4.3 million GEPT test takers (Wu, 2012), the most common test in Taiwan for university EFL learners was used. For simplicity's sake, English certification exit requirements are abbreviated as ECER in this study. This study compares two groups of schools: the ECER school that has established English certification exit requirements, and the non-ECER school that has not.

Since students' levels of English proficiency in these two schools were not similar (as seen in Table 1), an investigation of which group of students received better scores would not be able to determine whether test requirements promoted learning outcomes. Instead, a comparison of students' score gains by group can generate an understanding of whether such a test-driven policy has resulted in improved student English proficiency. Furthermore, few studies have been made in regard to the factors that may influence learning outcomes from learners' point of view in the field of language testing. Three research questions, therefore, were made in order to understand test effects on learning outcomes, and the possible determinants of learning outcomes, that is: (1) Have English certification exit requirements promoted statistically significant differ-

ences in the score gains?; (2) What are the possible factors contributing to students' scores at the ECER school?; and (3) What are the possible factors contributing to students' scores at the non-ECER school?

METHOD

A comparative approach is usually conducted to investigate test effects (Hayes & Read, 2003, 2004). For examples, studies by Alderson and Hamp-Lyons (1996) and Green (2007a, 2007b) compared test-preparation and regular classes to determine whether there are differences in regard to teaching practices and student learning. If differences exist, they can be taken as evidence of the existence of washback effects. Utilizing a mixed method which involved the collection and analysis of both quantitative and qualitative data, this study, therefore, made a comparison of the differences between the ECER school and non-ECER school in regard to their pre-and-post GEPT test scores.

Subjects

One hundred and thirty six first-year students at an ECER school and 140 first-year students at a non-ECER school took the 1st stage (listening and reading sections) of two authorized GEPT tests.

Six participants at an ECER school and six at a non-ECER school were interviewed. As seen in Table 1, the majority of the participants (83%) are females; two are males. Three have already successfully met their school's exit requirements; none of their counterparts passed the English certification tests. 67% of the interviewees made progress on the listening section, while 58% improved on the reading section. The bio-data of these 12 interviewees is categorized in Table 1, in regard to type of school (with exit or without exit requirements), gender, possession of certificates, and whether they progressed or regressed on their post-test scores.

Table 1. Profiles of the 12 Interviewees for Pre- and Post-Tests

ID	ECER/non-ECER	Gender	English certificates	Progress(+) Regress (-) in listening	Progress (+) Regress (-) in reading
S206-2	ECER	M	No	+	+
S206-1	ECER	F	No	+	+
S206-3	ECER	F	No	+	-

ID	ECER/non-ECER	Gender	English certificates	Progress(+) Regress (-) in listening	Progress (+) Regress (-) in reading
S207-2	ECER	F	Yes	+	-
S206-4	ECER	F	Yes	+	no change
S207-1	ECER	F	Yes	+	-
S208-1	Non-ECER	M	No	-	+
S212	Non-ECER	F	No	-	-
S210	Non-ECER	F	No	-	+
S209-1	Non-ECER	F	No	+	+
S209-2	Non-ECER	F	No	+	+
S208-2	Non-ECER	F	No	-	+

Instruments

To investigate test effects on learning outcomes, two authorized GEPT elementary level, first-stage tests (listening and reading parts), were given to the students as pre- and post-tests at ECER and non-ECER schools to determine whether students with the exit requirement progressed more than those without, and if they did, in which area of language skills (i.e. listening, reading) student scores increased the most. Only the first stage of the test was administered because of limited access to trained raters to assess the writing and speaking sections. Nevertheless, this would still demonstrate whether student learning outcomes improved more at the ECER school than those of the non-ECER school. This research incorporated authorized authentic GEPT elementary level tests because participants would be more likely to take the tests seriously if they were told the tests were real tests and had been used in the past. Therefore, the validity of the findings would be stronger.

The primary goal of the structured interviews with 12 students was to obtain the possible factors contributing to their score gains. The interview questions for these 12 students addressed 1) their time spent on English study, 2) the length of their English classes, and 3) the language skills focused in their English classes.

Data Collection and Analysis

The pre- and post- authorized elementary GEPT tests were held four months apart; one at the beginning and the other at the end of the semester.

Semi-structured interviews were conducted to 6 students from the ECER school and another 6 students from the non-ECER school. Students were recruited based on their willingness to take the pre-and-post tests, and be interviewed. Interviews were conducted in Mandarin for better communication and to avoid confusion. Each interview lasted about 10-30 minutes. Several follow-up phone interviews took place if incomplete or confusing responses occurred.

For the pre- and post-tests, descriptive statistics were calculated as well as independent t-tests to see if a larger score gain was found among the ECER students, and whether this was statistically significant at $p < .05$.

All the electronically recorded interview data was first translated from Chinese to English. Following the five analytical-strategy steps for structured interviews proposed by Schmidt (2004), the transcripts were analyzed. First, "intensive and repeated reading" (p. 254) of the transcribed interviews were done. Second, analytical main and sub-categories were constructed as a fundamental guide. Third, with the assistance of the qualitative software packet NVivo 8, all the transcribed data were sorted according to the analytical categories, with the goal of providing examples for the research questions. Fourth, quantifying the results where possible was applied for a preliminary overview of the data. Finally, detailed explanations were presented to draw inferences for the research questions.

FINDINGS AND DISCUSSION

This section presents the findings and discussion for the three research questions.

Question 1: Have English certification exit requirements promoted English learning outcomes?

A total of 276 first-year students from an ECER school and a non-ECER school took the listening and reading sections of two authorized Elementary GEPT tests. Since students' levels of English proficiency in these two schools were not similar, an investigation of which group of students received better scores would not be able to determine whether test requirements promoted learning outcomes. Instead, a comparison of students' score gains by group can generate an understanding of whether such a test-driven policy has resulted in improved student English proficiency.

Table 2 shows the pre- and post- listening and reading scores of 136 students at an ECER school and 140 students at a non-ECER school. Students' original proficiency levels at the ECER school were a little lower than those of students at the non-ECER school.

Table 3 shows the score gain the two groups of students achieved for each section. Both groups made greater progress on the listening section than on the reading section. Students at the ECER school showed a slightly higher score gain on the listening section (1.6 out of 120 points), while their counterparts showed a slightly higher score gain on the reading section (3.3 out of 120 points). Although students at the ECER school showed greater improvement than their counterparts in listening skills, they did not exhibit comparable increased progress in reading.

These findings are not fully consistent with Elder and O'Loughlin's (2003) and Green's (2007a, 2007b) studies, where learners at lower levels of proficiency made more rapid progress. Students' original proficiency does not seem to be a sole predictor of their score gains in this study.

The consistency among these studies is that students' score gains for listening were higher than those for reading. As Elder and O'Loughlin (2003) suggest, progress on listening may be easier to detect within a specific timeframe due to the fact that reading abilities involve knowledge of vocabulary, grammar, phrases, and reading comprehension. However, according to Table 4, an independent sample t-test reveals no statistically significant score gain difference on either the reading or listening sections between the two groups.

Table 2. Pre- and Post- GEPT Test Mean Scores and Standard Deviations of 276 Students in Two Groups of Schools

ECER (Listening test)		Non-ECER (Listening test)		ECER (Reading test)		Non-ECER (Reading test)	
Pre-test N=136	Post-test N=136	Pre-test N=140	Post-test N=140	Pre-test N=136	Post-test N=136	Pre-test N=140	Post-test N=140
m=73.41 SD=19.34	m=80.72 SD=21.59	m=74.94 SD=19.41	m=80.69 SD=22.53	m=67.82 SD=26.11	m=69.69 SD=26.09	m=75.50 SD=14.60	m=80.66 SD=18.20

Table 3. A Comparison of GEPT Test Score Gain for 276 Students in Two Groups of Schools

Score gain (Listening) (post-test – pre-test)		Score gain (Reading) (post-test – pre-test)	
ECER N=136	Non-ECER N=140	ECER N=136	Non-ECER N=140
7.31	5.75	1.87	5.16
SD= 13.25	SD= 15.07	SD=15.36	SD= 15.52

Table 4. An Independent Sample T-Test for Score Gain of 276 Students in Two Groups of Students

	t	df	Significance (two-tailed)
Score gain for listening	0.912	274	0.36
Score gain for reading	1.77	274	0.08

The minimal difference in score gains between the two groups suggests that such a test requirement policy plays a minimal role in improving student proficiency. In addition, the insignificant score gains suggest that the exit requirement policy does not affect students' learning behaviors to a large extent.

Question 2: What are the possible factors contributing to students' scores at the ECER school?

All 6 interviewees at the ECER school stated that they did not make any special efforts to study for the GEPT tests, apart from attending regular English classes. As seen in Table 3, all these 6 interviewees made progress on their listening scores. As for their reading scores, two made progress, three did worse, and one remained the same. Three out of six indicated that they also attended extra test-preparation classes offered by the school after class. The required English classes offered at this school are 2 hours of aural/oral classes and 2 hours of reading classes each week for first-year students, and 2 hours of aural/oral class for one semester for second-year students. In other words, 10 hours of English classes in total (6 hours of aural/oral classes and 4 hours of reading classes) are offered to students in their first two years.

Possible factors that led to a greater score gain on listening than on reading are: 1) more hours of aural/oral classes offered at the school with exit requirements, 2) test preparation classes offered for test-taking strategies and test-oriented practice, and 3) additional time required/encouraged to spend at a self-access center for listening practice. These are discussed below.

1. Aural/oral classes are offered in addition to 2 hours of reading classes.

Four out of the 6 interviewees stated that their improved listening scores may be linked to the aural/oral classes, where they were given opportunities to do listening exercises offered in the text, and to do oral practice such as chorus repetition or dialog practice with the teacher or a partner. For example, S206-4 claimed that she did not make more effort to prepare for the test, but only attended classes; however, in aural/oral classes, they were given opportunities to practice speaking.

In regular English reading classes, half of these 6 interviewees stated that their teacher included some test-preparation materials, while the remaining students indicated that little test-preparation instruction occurred. S207-2 stated that although vocabulary and grammar quizzes were given on a regular basis, there were fewer than they were given in high school, and that in high school, they were pushed more to study, but at university, it all depended on whether students wanted to study. Interestingly, when asked about the factors that resulted in their reading score gain, students were uncertain, even though test-preparation material and instruction were included in regular English classes. S206-4 said that the inclusion of test-preparation material and instruction began only a month prior to the post-test, and she wished that this type of instruction could have happened earlier. Another student, (S207-1) whose post-test reading score regressed, mentioned that the test-related instruction did not help her to gain a better understanding of the GEPT because she did not have the opportunity to take the entire set of mock tests; however, she found this type of instruction very beneficial because the teacher started with basic grammar, which helped her to build a solid foundation.

2. Test-preparation classes, where teachers told students about test-taking strategies, are offered.

Of the six interviewees, three students stated that in the TOEIC test-preparation classes, they were given some listening practice. S206-2 said,

I did not really study harder after class, but I paid attention in class... As for listening, we were required to listen to the test questions, and the teacher told us some listening strategies. She said something like the first sentence usually was the most important one.

S206-3 claimed that test-preparation classes pushed her to study more and improve her scores because in regular English classes, students were not given quizzes on a regular basis, so she did not really study for them. However, she liked regular English classes more because there was less pressure there.

I feel I'm studying at a university in regular English classes, while in test-preparation class, I felt like I was taking classes at a cram school... In test-preparation classes, we were usually forced to study, which I don't like very much because I don't like to study for tests... We were required to be there at a certain time and finish a certain number of mock tests before we were allowed to go home.

3. English teachers recommended or required their students to spend time on test-related practice at the self-access center, so students engaged in more listening practice there.

Four out of these 6 interviewees stated that their English teachers recommended that they spend a certain amount of time doing test-related practice at the self-access center, and those who finished the task were given extra credits as a reward. S206-2 said, "I would have felt like I had lost face if everyone else except me hadn't gone there." S206-3 said, "I went there for 10 hours and did listening practice." Likewise, S207-1 stated that her teacher encouraged them to do test-related practice at the self-access center, and "I go there when I have time to do listening practice."

Question 3: What are the possible factors contributing to students' scores at the non-ECER school?

The non-ECER school requires that first-year students take three hours of English classes. S208-2 stated that she attended lessons for 3 hours a week after class, which were offered by the school to help students to earn English tour certificates. S212 was preparing for an exam herself to transfer to another university. The remaining students did not attend classes, other than the 3-hour regular required English classes.

Factors that may possibly have contributed to students' lower listening score gain but higher reading score gain than their counterparts are: 1) English classes focused on reading rather than aural/oral skills, 2) self-study or after-class English lessons, where practice on reading skills is the focal point, and 3) additional listening practice after class is not required. These are discussed below.

1. English classes focused on reading skills more than on aural/oral skills.

All 6 interviewees claimed that the focus of their English classes was reading skills, not listening skills. They mainly read the text, reviewed the vocabulary, or did grammar-related assignments.

S208-1 said that since he entered university, he had never been given much practice on listening. The teacher focused on reading in class, and he therefore spent a lot of time memorizing vocabulary, reading the text and practicing making sentences, but he spent little time on listening. Later, he continued, saying that the English classes had helped to increase his vocabulary knowledge base, and his reading comprehension has improved because of his greater knowledge of vocabulary and expressions. "However, there seems to be little improvement on my listening." He offered this explanation:

In high school, listening was assessed on the mid-term and final exams, so I listened to the audio version of the textbook to practice my listening skills. At university, since listening is not assessed, I spent more time memorizing vocabulary, sentences and reading the text to prepare for the exams; therefore, I spent little time on listening.

S209-2 stated that her reading scores improved because the English teacher assigned her class a grammar report, where she needed to compile grammar rules and make sentences for each rule. In order to finish this assignment, she reviewed the grammatical knowledge she had learned at high school, and by reviewing it, she contended that she seemed to have improved her English.

S210 said that her teacher did not instruct on listening, but she assumed her reading had improved because she had read more English articles in class, and her vocabulary bank had enlarged accordingly.

When asked if they made any efforts to practice their listening skills, they replied "nothing in particular." Three students said that they listened to English occasionally when their teacher played the video version of the textbook, or

when he read the text, but that did not happen often. S209-2, whose listening and reading scores both improved a lot, said that she did not have time to do any listening practice because there were too many reports to do for other subjects, but she said her listening was satisfactory because she used to work in a restaurant that many foreigners visited, so she had opportunities to practice her English. Moreover, she did not find the second test more difficult than the first one.

When asked if oral skills were practiced in class, their answer was “very little.” S208-1 said this was because it was a reading class. S209-2 said that the teacher occasionally read the text, played its audio version, and had them read the text; that was all. S210 said, in regular English classes, reading seemed to be the only skill focused on, but there was no other choice because there were only three hours of class time each week, and it was impossible to cover all skills in such a short time. Usually the class time was over by the time the teacher had finished instructing about the article, and there was not much time left to practice other skills.

2. Self-study or after-class English lessons, where practice on reading skills is the focal point

Two out of the 6 interviewees said that they studied vocabulary, grammar, and English articles after class by themselves either to prepare for the transferring entrance exam or to enhance their own English proficiency. Neither of them stated they did any listening practice themselves after class. S208-1 said that the entrance exam was a written test, so she focused on the vocabulary and grammar. S212 claimed that although her scores on the GEPT test did not reflect her improvement, she found she improved a little by reviewing vocabulary and grammar herself. She said,

I understood some grammatical points that had confused me better. I found that my reading comprehension became better when I did some reading.

S208-2 attributed her score gain on reading to the fact that she attended the after school Travel English classes offered by the school. She said her progress on reading might have something to do with the Travel English class, where “we are required to do a lot of mock tests, which are similar to what is tested on the GEPT to some extent. All the questions are multiple-choice ques-

tions, consisting of vocabulary, grammar, and reading comprehension. The teacher reviewed and explained the questions after we finished them... After class, we are assigned to finish the mock tests we did not complete in class.”

3. Students are neither required to nor rewarded for engaging in listening practice on the school website.

When asked if they practiced their listening skills on the school website by doing test-related practice, or whether they attended test-preparation classes offered by school, the students either said they rarely did, or that they did not know where to go for this practice or class. S208-1 said,

I did not know about these (the test-preparation classes) until you told me just now... To be honest, I do not know what resources (for test-preparation) have been provided by the school, unlike in high school, where the teacher had us do mock test practice in addition to regular English instruction.

S209-1 said, “I do not know where we can do such practice.” However, S210 said that she plans to do the online test-related practice offered by the school when her English proficiency reaches a certain level. She said that she knew the school had offered this website for them to do test-related practice, because the teacher had told them about it, and she had visited it once before.

Jez and Wassmer (2015) contend that “the more time spent to produce something (holding the other inputs into the production constant) the greater should be the quantity and/or quality of the output produced” (p. 287). The findings derived from this study are consistent with their proposal and imply that the amount of time, both the extension of class time, and after-class learning time, spent on language learning plays an essential role in determining a student’s degree of improvement. Previous studies also consistently indicate that the more time learners spend in learning, the better academic outcomes they receive (Borg, 1980; Brown & Saks, 1986; Cotton & Savard, 1981). Similarly, the most recent study conducted by Jez and Wassmer (2015) found that “fifteen more minutes of school a day at a school site (or about an additional week of classes over an academic year) relates to an increase in average overall academic achievement of about 1%, and about a 1.5% increase in average achievement for disadvantaged students” (p. 284). Furthermore, as found in this study, autonomous learning by encouraging students to make use of the materials on the website or in the self-access center is considered beneficial in

boosting learners' test scores. Similar results were also discovered in the studies of McMurry, Tanne and Anderson (2009) and Thanasoulas (2002). English teachers should encourage more autonomous use of center materials by providing students with an appropriate website and database to help them find materials that are beneficial for their English learning.

CONCLUSIONS AND SUGGESTIONS

Utilizing two authorized GEPT tests as pre- and post-tests, and conducting interviews with 6 each from the ECER and non-ECER schools, this study intended to explore 1) whether English certification exit requirements have promoted learning outcomes, and 2) the possible contributing factors to the score gains for each group of students.

With regard to learning outcomes, 136 ECER students had a 1.6-point higher listening score gain than 140 non-ECER students, who in turn had a slightly higher level of reading score improvement (3.3 out of 120 points). Neither score gains were statistically significant.

Findings from the interviews with the 12 students confirmed that the slightly higher score gain on listening for the students at the ECER school is likely due to the larger amount of time engaged in practice of listening skills, which occurred in regular aural/oral English classes, test-preparation classes, or their school's self-access center or website. Additional hours of English classes (both regular English aural/oral classes or test-preparation classes) and the frequent use of the school's resources (self-access center or online test-related school website) have been established to help students to enhance their English proficiency and earn the English certificates, and they seem to have effects on improving students' listening scores, to some extent.

A focus on reading skills at the non-ECER school seemed to explain the slightly higher score gain on reading for its students. One more hour of reading class is offered at the non-ECER school than at the ECER school, suggesting that additional hours of English classes are beneficial to boosting students' scores. This is also proven in the case of the students at the ECER school, where the score gain on listening is a little higher than for their counterparts, partly because of the increased amount of time they spend practicing listening skills. In addition, since reading was mainly covered in both the mid-terms and finals in this school without exit requirements, students studied the text, vocab-

ulary, and phrases to prepare for the exam, and did little practice on listening after class.

These findings imply that the amount of time spent on language learning plays an essential role in determining a student's degree of improvement. It is therefore suggested that (1) additional hours of English classes, and (2) an additional number of test-preparation classes, (3) the extension of English classes from the first to the final year be offered to students to provide them with more opportunities and exposure to English practice. Furthermore, since students are the major stakeholders of this policy, of course, they need to work hard on their own after class. Autonomous learning must be encouraged. Even if teachers have done their jobs well, if students do not devote their own time and effort to studying English and preparing for tests, their learning outcomes will still be insignificant. To encourage autonomous learning, students should be provided with sufficient resources (e.g. self-access center and online test-preparation website) to increase the time they study both in and after class.

Although this study offers practical insights into the test effects of the English exit requirement policy on learning outcomes and highlights the possible factors contributing to test score gains, it has some limitations that indicate directions for further research. First, this was a case study of first-year students and 12 interviewed students only from two universities. In order to ascertain the relevance of the findings in other contexts, replication studies would need to be conducted. Second, while interviews were used in the study partly to triangulate what the test results indicated, other types of data could usefully have been employed in order to enhance the trustworthiness of the study, including student questionnaires, teacher interviews, and classroom observations.

REFERENCES

- Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280-297.
- Andrews, S., Fullilove, J., & Wong, Y. (2002). Targeting washback – A case study. *System, 30*(2), 207-223.
- Borg, W. R. (1980). Time and school learning. In C. Denham & A. Lieberman (Eds.), *Time to learn* (pp. 33-72). Washington, D. C.: National Institute of Education.

- Brown, B. W., & Saks, D. H. (1986). Measuring the effects of instructional time on student learning: Evidence from the beginning teacher evaluation study. *American Journal of Education*, 94(4), 480-500.
- Burrows, C. (1998). *Searching for washback: An investigation into the impact on teachers of the implementation into the adult migrant English program of the certificate in spoken and written English*. (Unpublished doctoral dissertation, Macquarie University, Sydney, Australia).
- Burrows, C. (2001). Searching for washback: The impact of assessment in the certificate in spoken and written English. In G. Brindley & C. Burrows (Eds.), *Studies in immigrant English language assessment*. (Vol. 2, pp. 95-187). Sydney, Australia: National Center for English Language Teaching and Research.
- Burrows, C. (2004). Washback in classroom-based assessment: A study of the washback effect in the Australian adult migrant English program. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 113-128). Mahwah, NJ: Lawrence Erlbaum.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 147-170). Mahwah, New Jersey: Lawrence Erlbaum.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge: Cambridge University Press.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed., Vol. 7, pp. 349-364). New York; London: Springer.
- Cheng, L. (2014). Consequences, impact, and washback. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1130-1146). New York: John Wiley and Sons, Inc.
- Cotton, K., & Savard, W. G. (1981). *Time factors in learning*. Portland, OR: Northwest Regional Educational Laboratory.

- Elder, C., & O'Loughlin, K. (2003). Investigating the relationship between intensive English language study and band score gain on IELTS. In R. Tolloh (Ed.), *IELTS Research Report 4* (pp. 207-254). Canberra: IELTS Australia Pty Limited
- Green, A. (2007a). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge: Cambridge University Press
- Green, A. (2007b). Washback to learning outcomes: A comparative study of IELTS preparation and university pre-sessional language course. *Assessment in Education, 14*(1), 75-97.
- Hayes, B., & Read, J. (2003). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In R. Tolloh (Ed.), *IELTS Research Report 4* (pp. 153-206). Canberra: IELTS Australia Pty Limited.
- Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: Preparing students for the IELTS academic module. In L. W. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 97-111). Mahwah, New Jersey: Lawrence Erlbaum.
- Hughes, A. (1988). Introducing a needs-based test of English language proficiency into an English-medium university in Turkey. In A. Hughes (Ed.), *Testing English for university study* (pp.134-153). Hong Kong: Modern English Publications and The British Council.
- Jez, S., & Wassmer, R. W. (2015). The impact of learning time on academic achievement. *Education and Urban Society, 47*(3), 284-306.
- McMurry, B. L., Tanne, M. T., & Anderson, N. J. (2009). Self-Access centers: Maximizing learners -- Access to center resources. *TESL-EJ, 12*(4), 1-13. Retrieved from <http://www.cc.kyoto-su.ac.jp/information/tesl-ej/ej48/a2.html>
- Oerke, B., Merki, K. M., Holmeier, M., Jager, D. J. (2011). Changes in student attributions due to the implementation of central exit exams. *Educational Assessment, Evaluation, and Accountability, 23*(3), 223-241.

- Pan, Y. (2011). Teacher washback from English certification exit requirements in Taiwan. *Asian Journal of English Language Teaching*, 21, 23-42.
- Pan, Y., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching (e-FLT)*, 9(1), 108-122.
- Schmidt, C. (2004). The analysis of semi-structured interviews. In U. Flick, E. V. Kardoff, & I. Steinke (Eds.), *A companion to qualitative research* (pp. 253–258). London: Sage.
- Smyth, E., & Banks, J. (2012). High stakes testing and student perspectives on teaching and learning in the Republic of Ireland. *Educational Assessment, Evaluation, and Accountability*, 24(4), 283-306.
- Stecher, B., Chun, T., & Barron, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington State. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 53-71). Mahwah, New Jersey: Lawrence Erlbaum.
- Thanasoulas, D. (2002). What is learner autonomy and how can it be fostered? *The Internet TESL Journal*, 6(11). Retrieved from <http://www.iteslj.org/Articles/Thanasoulas-autonomy.html>
- Wall, D. (2000). The impact of high-stakes testing on teaching and learning: Can this be predicted or controlled? *System*, 28(4), 499-509.
- Wall, D., & Horak, T. (2006). *The impact of changes in the TOEFL examination on teaching and learning in central and eastern Europe: Phrase 1, the baseline study*. Princeton: Educational Testing Service.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from classroom-based research. *Language Testing*, 13(3), 318-333.
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*, (pp. 129-146). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wu, J. (2012). GEPT and English Language Teaching and Testing in Taiwan, *Language Assessment Quarterly*, 9(1), 11-25.