



# An Investigation of the Washback Effects of an Intensive Test-Preparation Program on EFL Learning

**I-Chun Vera Hsiao**

([918083@stu.chsc.tw](mailto:918083@stu.chsc.tw))

National Taiwan University of Science and Technology, Taiwan

**Ying Stella Lin**

([stella242630@gmail.com](mailto:stella242630@gmail.com))

National Taiwan University of Science and Technology, Taiwan

**Shao-Ting Alan Hung**

([alanhung123@gmail.com](mailto:alanhung123@gmail.com))

National Taiwan University of Science and Technology, Taiwan

---

## Abstract

An increasing number of universities in Taiwan have emphasized the importance of English proficiency tests to demonstrate students' English ability. Universities have initiated Intensive Test-Prep Programs (ITPPs), hoping to enhance students' English proficiency. Hence, this study aimed to explore the washback effects of an ITPP on non-English majors. Specifically, the study intended to find out how an ITPP affected non-English majors' English learning motivation and learning strategies, and the relationship between students' test performance and washback effects. An experimental group of 52 and a control group of 60 non-English majors were recruited from a national university. Students from the experimental group attended a five-week ITPP, which emphasized promoting students' reading and listening skills, while the participants from the control group prepared for the proficiency test on their own. Three instruments were employed to collect the data, including questionnaires, semi-structured interviews, and students' scores from TOEIC. The results disclosed that the two groups differed statistically in motivation. Furthermore, results showed that three out of sixteen learning strategies differed significantly between the two groups: taking mock tests, reading newspapers, articles and magazines, and learning through English websites. However, the relationship between students' test performance and washback effects was not significant. Finally, the study offers implications for non-English majors, teachers, researchers, and curriculum developers.

---

## 1 Introduction

### 1.1 Background of the study

In Taiwan, English proficiency is regarded as an important skill, and has become a crucial assessment criterion for institutions to evaluate their potential students or employees. Standardized English proficiency examinations meet the need by providing a credible evaluation system covering listening, speaking, reading, and writing skills. Since the majority of Taiwanese believe the score on such an exam corresponds to a person's English proficiency, professors and employers can easily

determine whether their candidates meet their minimum requirements. It leads to the phenomenon of more and more university students putting emphasis on external English proficiency tests, for example, the *Test of English for International Communication* (TOEIC), and the *General English Proficiency Test* (GEPT). Since 2003, universities in Taiwan have been encouraged to set an English graduation benchmark for non-English majors (Pan, 2014). Among all the English proficiency tests, TOEIC and the GEPT are adopted the most by students. In order to enhance students' English ability, a number of universities have initiated intensive test-preparation programs (ITPPs). These programs have become more prevalent in recent years. Although these kinds of programs are not included in the universities' official curriculums and do not provide any credit, they feature test-oriented lectures and practices, which help students grasp more effective test-taking skills and are believed to provide suitable access to passing the exit requirement.

### ***1.2 The definition of washback***

*Washback* is a term commonly used in the field of applied linguistics, language testing, and language literature. It has generally been defined as the influences that tests exert on teaching and learning (Alderson & Wall, 1993). To be more precise, they indicated that tests were regarded as "powerful determiners" of classrooms, for example, classroom activities, course syllabi, classroom behaviors, and so forth. They then proposed *The Washback Hypothesis* which assumed that "teachers and learners do things they would not necessarily otherwise do because of tests" (p. 5).

Bachman and Palmer (1996) considered that "washback has potential for affecting not only individuals, but the educational system as well" (p. 31). More specifically, they pointed out that test takers and teachers were those individuals who received the most direct impact from tests. For test takers, the testing procedure influenced them in the following respect (Bachman & Palmer, 1996, p. 31): "(1) the experience of taking, and, in some cases, of preparing for the test, (2) the feedback they receive about their performance on the test, and (3) the decisions that may be made about them on the basis of their test scores." As for teachers, they might be well aware of the impact of tests on their teaching, which could sometimes contradict their teaching values and objectives. Secondly, tests, especially high-stakes tests, involved a large number of individuals for such specific purposes as university entrance exams and immigration grant; people's perception of test results was also one of the washback effects that could exert either positive or negative influences on the educational system, or even the society.

Bailey (1999) summarized the definition of washback when he explained that "some take a narrow focus on teachers and learners in classroom settings, while others include reference to tests' influences on educational systems and even on society in general" (p. 3). This study intended to investigate the washback effects on such aspects as learning strategies and motivation.

### ***1.3 The gaps in the washback literature***

Synthesizing the washback literature, three research gaps have been identified. First, though there have been many studies emphasizing washback effects in the past 30 years, most of them were conducted from teacher perspectives, such as those by Chen (2002), Cheng (2004, 2005), Ferman (2004), Hawkey (2006), Qi (2005), and Wu (2008). Comparatively little research had focused on the effects of tests on students' learning. Hence, Cheng (2008), Spratt (2005), and Watanabe (2004) highlighted the need to explore washback effects on learners, who are directly influenced by tests. As Cheng (2008) further specified, "It is also important to investigate the impacts of the test constructs, test methods and the function of the tests on students and on their learning process (including test-taking processes) and learning outcomes (test scores)" (p. 360). Hence, there is a need to conduct a follow-up study to examine washback effects on learning. Second, little research has examined washback effects in relation to learners' behaviors. In other words, individual differences such as motivation, learning strategy use, and test performance remained unexplored in the washback studies. Third, Pan (2012, 2014) pointed out that information such as self-reported student data is

easily prone to expectancy bias (Nisbett & Wilson, 1977; Yu, 2010) and suggested that subsequent research should include participant interview data with a view to presenting a clearer picture of washback patterns on learners. Apart from questionnaires, interviews can provide more concrete data, and abundant data sources can also increase a study's validity. Thus, this research intended to collect more holistic data by conducting semi-structured interviews in addition to the use of questionnaires and scores.

#### **1.4 Purpose of the study**

To fill the above-mentioned gaps in the washback literature, the purpose of this study was to explore the washback effects of an ITPP on students' English learning. To be specific, the study intended to examine (1) English as a foreign language (EFL) learners' motivation and learning strategies in the ITPP and (2) the relationship between washback effects and test performance. Three research questions were proposed as follows.

- (1) In what ways does an ITPP affect non-English majors' English learning motivation?
- (2) How does an ITPP affect non-English majors' English learning strategies?
- (3) What is the relationship between students' washback effects and test performance?

## **2 Literature review**

### **2.1 The concept of intensive test-prep programs**

According to Collins, Halter, Lightbown, and Spada (1999), the different models of intensive English programs can be divided into three types: (1) distributed, (2) massed, and (3) massed-plus programs. In the distributed program, students received 8 hours per week for 10 months, taking nearly 300 hours of English courses; in the massed program, English courses lasted for 18 to 20 hours per week for 5 months, for approximately 400 hours in total; in the massed-plus program, students received almost the same amount of exposure to English courses as those in the massed program, but they were encouraged to continue using English outside the class additionally.

The distributed, massed, and massed-plus programs were long-term programs, but the intensive test-preparation program in this study lasted only 5 weeks, which is a short-term one, and is more concentrated than the models that Collins et al. (1999) proposed. Although the various kinds of intensive English courses were analyzed, there were few exact studies investigating whether intensive or more distributed courses were more effective at enhancing students' English abilities (Serrano, 2010). Most of the studies focused only on either intensive or distributed programs, but there were few studies combining and comparing these two types of programs. For instance, the results of Carroll's (1967) and Stern's (1985) analyses showed that the more time students spent on learning English, the better performance they had at the proficiency exam. However, other research (Serrano & Muñoz, 2007) showed that students participating in the intensive programs made more progress than those in the regular programs. Serrano (2010) also suggested that students in intensive programs outperformed their peers in less concentrated courses. These studies implied that both intensive and more distributed courses enhance students' English abilities, but no comparison between the two types of courses has been made.

### **2.2 Empirical washback studies**

Over the past 3 decades, there were relatively fewer studies focusing on the washback effects tests exerted on learners and their learning compared to studies conducted from teachers' perspectives (e.g. Alderson & Wall, 1993; Chen, 2002; Cheng, 2005; Ferman, 2004; Hawkey, 2006; Qi, 2005; Watanabe, 2004; Wu, 2008). Cheng (2008), Spratt (2005), and Watanabe (2004), therefore, called for more exploration of tests' washback effects on learners, who are directly influenced by tests. Thus, two claims have been made:

(1) The exit requirement may motivate a small amount of learning through test-preparation methods; however, there were no obvious changes in students' learning strategies. Stoneman's (2006) study revealed that for learning English or preparing for the test, students' learning activities are not influenced by tests, as found in the studies by Cheng (1998), Pintrich (1999), Pan (2012), and Xie (2013). Similarly, as Pan (2014) concluded, "There was no statistically significant difference between exit and non-exit students in regard to their language-skill building activities." (p. 13). Traditional strategies of preparing for the tests such as reading textbooks, memorizing vocabulary, learning grammar, and doing exercises on paper have still been most frequently adopted by students.

(2) Such factors as proficiency level, learning attitude, learning motivation, and perception toward tests play an important role in students' test performance and their devotion to preparing for the test. According to Xie (2013), the students volunteering to participate in the research may have had higher motivation, and this may have contributed to their test performance. In the same vein, in Chu's (2009) studies, low-proficiency students tended to be more anxious about the test than high-proficiency students. These two studies were all consistent with the implications of other various studies (e.g. Cheng, 2005; Green, 2007; Pan, 2012; Watanabe, 2004).

### ***2.3 The effects of test-oriented activities***

Whether or not the test-oriented activities have more effect on students' test scores than general English language courses remains controversial. Two entirely different claims have been made by several studies; for example, Brown (1998) found that a preparation course improved scores more than an EAP (English for Academic Purpose) course, but Robb and Ercanbrack (1999) established that test-prep courses did not outperform general English courses. Despite this contradiction, Green (2007) suggested that test-oriented instruction could exert positive effects on learners' confidence during the test preparation.

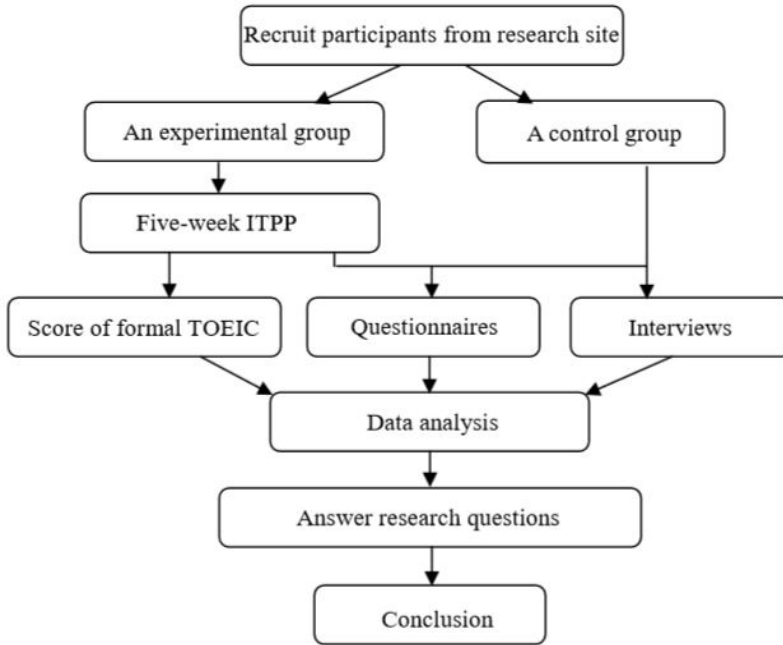
Messick (1982) categorized test-oriented activities into four types: test preparation improves test scores through (1) enhancing the intended constructs of the tests, (2) reducing construct-irrelevant difficulties, such as test anxiety, and unfamiliarity with test formats, (3) enhancing test-taking skills that are irrelevant to the construct, and (4) using test materials, following a test-based curriculum, using similar or identical test items, or focusing exactly on what the test measures. Whereas Type 1 activities focus on the development of a broad range of skills entailed by the target domain, Type 4 activities focus on a narrow range of skills sampled by the test. (p. 198)

Among these types of preparation, strengthening test-taking skills has been the most frequently used strategy for test preparation (Xie, 2013). However, the reputation of test-skills practice is quite low. Xie (2013) mentioned that it will threaten the extrapolation validity of a test, whereas Jin (2006) and Miller (2003) have raised concerns about narrowing curriculum content, excessive instruction in test-taking strategies, and repeated practice with test papers. Since the design of test-oriented activities influences learners' preparation strategies, improvement of test scores, and even study results, it plays an important role in test takers' learning.

## **3 Methods**

### ***3.1 Research design***

According to Mukundan, Mahvelati and Nimehchisalem (2012), "In experimental research, a control group can determine if the effect of the treatment has, in fact, resulted from the treatment rather than other possible factors" (p. 6). Hence, in order to exclude unexpected variables other than the implementation of the ITPP, the current research adopted an experimental design to examine the washback effects of an ITPP on students' learning. The design of the current research is shown in Figure 1.



**Fig. 1. The process of the research design**

### 3.2 Participants

A total of 112 non-English majors from a national university in northern Taiwan were recruited to participate in the study, including 52 students in the experimental group and 60 students in the control group. In the experimental group, eight students who were absent in any session, failed to fill out all the items in the questionnaires, and failed to submit the transcripts of the formal TOEIC were excluded. Hence, the final remaining number of students in the experimental group was 52. Due to the time and class size, 52 students were recruited from two consecutive semesters. However, these two ITTP sessions in two semesters were identical in terms of teaching assistants, course design, materials, hours and mock tests. Demographic profiles of both groups are presented in Table 1. The students came from engineering, electrical engineering, management, applied sciences, and design-related departments. All participants provided their latest TOEIC scores beforehand, and only those whose scores ranged from 500 to 750 were eligible to participate in the ITTP initiated by the Language Center of the university. These students were classified as pre-intermediate to intermediate level (B1-B2) English learners according to the Common European Framework of Reference for Language (CEFR). Students whose TOEIC scores were below 500 were advised to participate in other English classes, which were not considered as intensive programs, and therefore were not recruited for the current study.

For the interview phase of the current research, 16 students (five interviewees from the control group and 11 from the experimental group) were recruited. A demographic profile of the interviewees is shown in Table 2.

**Table 1. Questionnaire respondent profiles**

		The experimental group		The control group	
		Respondents (n)	Percentage (%)	Respondents (n)	Percentage (%)
Gender	Male	24	46	40	67
	Female	28	54	20	33
Program	Undergraduate	15	29	51	85
	Graduate	37	71	9	15
College	Engineering	12	23	15	25
	Electrical engineering	19	37	26	43
	Management	15	29	14	23
	Applied sciences	4	7	1	2
	Design-related	2	4	4	7

**Table 2. Interviewee respondent profiles**

		Respondents (n)	Percentage (%)
Gender	Male	10	63
	Female	6	37
Grade	Undergraduate	7	44
	Graduate	9	56
Group	The experimental group	11	69
	The control group	5	31

### 3.3 *Intensive Test-Prep Prog*

Aiming to promote EFL students' test-taking skills and improve their scores on standardized proficiency tests, the ITPP lasted 5 weeks, with 40 hours in total. Two main goals were expected to be achieved during the ITPP: (1) to help participants make progress in the listening and reading sections of the TOEIC; and (2) to evaluate the washback effects on such aspects as learning strategies and motivation. The curriculum design, which included TOEIC mock tests for the listening and reading sections, is shown in Table 3. Since the ITPP was a supplementary course to the university's formal English curriculum, it was scheduled in the evening to avoid an overlap with the courses in the formal curriculum, which were usually scheduled in the day. Furthermore, teaching assistants were one of the crucial elements of the ITPP. The ITPP teaching assistants were responsible for the following tasks: (1) teaching test-taking strategies, such as distinguishing different English accents and getting the gist from the reading passages; (2) conducting lessons on reading and listening; (3) explaining difficult test items; and (4) administering the mock tests and vocabulary tests. The three teaching assistants, who were responsible for one day in a week, respectively, were all graduate students in the Master's Program of Teaching English to Speakers of Other Languages (TESOL) from the same school as the participants. They gained almost full scores in the formal TOEIC, and have been teaching test-preparation program for over one year. In addition, they had several teaching experiences outside the school, including teaching in the cram schools, and participating in teaching practicum, and so forth. The researchers had discussed the purposes of the research with the three teaching assistants individually before the ITPP started and obtained their consent.

**Table 3. Curriculum design**

Time	Monday	Wednesday	Friday
6:30-7:30 (p.m.)	TOEIC mock test	Explaining difficult test items on reading	Explaining difficult test items on listening
7:30-9:00 (p.m.)	TOEIC mock test (including checking the answer & recording the results)	Teaching test-taking strategies & conducting test-prep lessons	teaching test-taking strategies & conducting test-prep lessons

### 3.4 Instruments

**Questionnaire:** Participants from the experimental group and control group filled out the questionnaire. The questionnaire was adapted from Cheng (2005) and translated into Chinese by researchers. The items were modified, for example, adding more questions about participants' background, and replacing the 5-point Likert scale with a 6-point Likert scale. According to Chomeya (2010), the 6-point Likert scale tends to give discrimination and reliability values which are higher than the 5-point Likert scale. In order to ensure content validity, the questionnaire items were given to an experienced researcher to check for any ambiguity or misleading statements. The experienced researcher had extensive research experiences, published numerous journal articles in international journals, and have been a TESOL professor for several years. The questionnaires was used to help the researchers understand and evaluate the participants' experiences, curricular and extracurricular activities, and strategies and motivation for learning English. A few additional questions designed for the experimental group were specifically added, for example, how do you agree with the following statements after you participated in intense TOEIC-oriented program? (see Item 13 in Appendix A). These items were designed by researchers in the hope of gaining a better insight into students' perceptions toward the ITPP. Containing a total of 42 items, the finalized questionnaire measured three major constructs: background information of participants, motivation for English learning, and English learning strategies (for complete questionnaire, see Appendix A). Considering that the program had a tight schedule and the questionnaire had 42 items, participants were given as much time as they wanted before or after the last period of the ITPP to fill out the questionnaire.

**Interviews:** According to Pan (2014), interview data presented a clearer picture of learners' washback. To better understand the participants' motivation, strategy use and opinions about the ITPP, 16 semi-structured interviews were conducted after the program, 11 interviews with participants from the experimental group and five with those from the control group. The interviews were all conducted in Chinese. Based on the study of Gass and Mackey (2005), interviews can be conducted in the interviewees' first language to remove concerns about language proficiency impacting quality and quantity of the data provided. To gain a better understanding of the data analysis, a few questionnaire items were incorporated into the interview questions, for example, Interview question 9, "In the following aspects, how are you affected by the ITPP? Self-image, motivation to learn, peer relationship, future employment opportunities, teacher and student relationship, anxiety and emotional tension". The interview questions for the experimental group and the control group were different, since some questions were designed specifically for those students who participated in the ITPP. For instance, "What were your methods and strategies of learning English? Did you change methods and strategies after the ITPP? What methods and strategies did you learn from the ITPP? Do you use those strategies after the ITPP? Do they work?" (for English version of interview questions, see Appendix B). Each interview lasted about 30 to 40 minutes, and the process was recorded and transcribed.

Scores: Participants from the experimental group were required to take a formal TOEIC after the ITPP. Their scores on TOEIC were thus collected to be compared with their responses on washback effects in order to uncover the relationship between test performance and washback, as stated in research question 3. Only the scores from students in the experimental group were collected due to the following reason. The study aimed to analyze the relationship between the washback effects of the ITPP and learners' test performance rather than to compare the improvement in test performance of the two groups. Since the control group did not attend the ITPP, collecting the scores from this group would fail to answer the research question.

### 3.5 Data analysis

A questionnaire, scores from the TOEIC, and interviews were utilized to answer the three research questions. Responses on all items in the questionnaire were entered into SPSS version 22 to analyze the data. An independent *t* test was utilized to test the significance of the differences between students' motivation and strategies. In addition, the participants' scores from the TOEIC were collected to compare with responses on the washback effects. The Pearson correlation coefficient was computed to find out the relationship between students' test performance and washback effects. Finally, interview data were used to triangulate the data. In brief, both quantitative and qualitative data were analyzed to bring more objective and holistic perspectives to students' motivation and strategies.

## 4 Results & discussion

The results of three research questions are presented in the following sections.

### 4.1 Research question 1: In what ways does an ITPP affect non-English majors' English learning motivation?

Table 4 shows the comparison of students' motivation toward learning English between the experimental and control group. The overall difference between the two groups in their motivation toward English learning is noteworthy,  $t(110) = -2.26, p < .05$ . Though the two groups both have high motivation ( $M = 4.79$  for the experimental group;  $M = 4.53$  for the control group), the experimental group experienced statistically higher motivation than the control group.

**Table 4. The *t* test for independent samples comparing the mean difference in motivation between the experimental group and control group**

Group	N	M	SD	
Experimental group	52	4.79	1.01	
Control group	60	4.53	1.21	
Independent comparisons				
Comparison	M	SE	t(110)	Sig.
Experiment vs. Control	-.25	.11	-2.26	.03*

\*  $p < .05$

A closer analysis revealed that of the 16 items in the questionnaire, three items (*Item 3: To me, to get a better job is a reason for learning English; Item 4: To me, to be able to communicate with people is a reason for learning English; and Item 7: To me, to have more and better opportunities in the future is a reason for learning English*) showed relatively higher mean values than the other items. The findings were also supported by the interview data, which showed that most of the interviewees in the ITPP had higher motivation for studying English when they were aware that their future career goals would be strongly associated with better English proficiency. Similar to the majority of the interviewees in the experimental group, Interviewee #C2, from the control group,



also reported that “I study English mainly for the purpose of working or studying abroad in the future. As a result, I spend one and a half hours a day learning English.”

Furthermore, the interviewees revealed that the ITPP motivated them to learn English on a fixed schedule. Most of the interviewees in the experimental group mentioned that the environment was a key factor for English learning in their academic career. Interviewee #E5 brought up his idea that “The reason why I participated in ITPP was that I would not otherwise actively study English under certain conditions. In other words, the environment of ITPP did impel me to a higher motivation of learning English.” Moreover, the teaching assistants in an ITPP can motivate students. Interviewee #E8 even commented that “The teaching assistants’ teaching styles affected my motivation. The teaching assistants were interesting and I was looking forward to the classes. They were so hard-working that I didn’t want to let them down.” As mentioned by Interviewee #E6, the teaching assistants in the reading class were enthusiastic and vigorous, who captivated his attention through their teaching. In addition, the time arrangement of the ITPP played an important role in students’ motivation. Interviewee #E3 also made a remark that she could make progress on her TOEIC score through her self-preparation; however, participating in the ITPP contributed to further improvement of her English, because the ITPP was a regular and intensive class.

The results show that an ITPP can motivate students’ English learning, which is in line with the study by Jacques-Bilodea (2010), who noted that ITPPs can escalate students’ motivation, commitment, and engagement. It has also been pointed out in other studies (Burton & Nesbit, 2002; Daniel, 2000) that when courses are intensively organized, second language learners’ motivation, commitment, and engagement increase in turn. The present study indicates that the environment, teaching assistants’ teaching styles, and the time arrangement of an ITPP can sustain students’ motivation and concentration in learning English.

#### 4.2 Research question 2: How does an ITPP affect non-English majors’ English learning strategies?

The research data in Table 5 compare how the two groups of non-English majors employed English learning strategies. Among the 14 strategies listed in the questionnaire, three strategies differ significantly between the experimental and control group: (1) take mock tests,  $t(110) = -3.06$ ,  $p < .01$ ; (2) read English newspapers, articles, and magazines,  $t(110) = -2.09$ ,  $p < .05$ ; and (3) learn through English learning websites,  $t(110) = -3.38$ ,  $p < .001$ . The interview data below depict how this occurred.

**Table 5. A comparison of students’ learning strategies**

Students’ English learning strategies	The experimental group			The control group			t-value	sig
	N	M	SD	N	M	SD		
(1) Take mock tests	52	4.28	0.98	60	3.89	1.25	-3.06	.00**
(5) Read newspapers, articles, and magazines	52	3.33	1.12	60	2.87	1.18	-2.09	.04*
(12) Learn through English learning websites	52	3.42	1.15	60	2.65	1.24	-3.38	.00***

\*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$

The interview results also reported some differences in terms of learning strategies between the two groups. When asked if the ITPP changed their English learning strategies, the majority of the interviewees from the experimental group indicated the skills they learned in the ITPP were then utilized to facilitate their general English learning. Listening skills such as identifying different accents and replaying the online listening materials repeatedly stimulated them to expose themselves to English learning websites. According to Interviewee #E11, he believed that practicing the skill to distinguish between various accents in the ITPP was a crucial component of increasing his listening

comprehension, which proved to be useful when he took the formal TOEIC. Reading skills that included scanning and skimming, understanding basic structure of English articles, and being aware of English grammar helped the interviewees speed up their reading and understand the gist of the materials. Interviewee #E9 reported that “I learned several reading skills in the ITPP, including reading topic and concluding sentences of an article first, and using skimming and scanning. With these, I can read faster and get the main idea more easily and quickly than before.” The result, however, contradicts Pan’s study (2014), which stated that test preparation did not greatly change students’ reading skills.

Another noteworthy aspect was that there was no significant difference in memorizing vocabulary between both groups; nevertheless, from the interview data, it was another strategy frequently used by the students. Several interviewees (#E6, #E7, #E9, #E10, #C1, and #C2) mentioned that categorized and systematic vocabulary improved their memory while preparing for the test. According to Mizumoto (2008), “Learners with higher TOEIC scores had clear goals and attended to vocabulary learning strategies in conscious, coordinated, and structured manners” (p. 27). Furthermore, they disclosed that vocabulary learning strategies generally had the greatest impact on TOEIC scores.

In light of this, an ITPP does influence students’ learning skills, and furthermore, it affects their learning strategies. The findings correspond to Xie’s study (2013), which found that the most frequently used test preparation practice was rehearsing test-taking skills, followed by test preparation management, drilling, memorizing, and affective strategies in turn. Learning strategies focused on the development of language skills via extensive and functional uses of English language were used scarcely.

Hence, the test-taking strategies provided in ITPPs may contribute to positive effects on students’ future English learning. Zimmerman and Pons (1986) also concluded their study by indicating that “Other analyses revealed that students’ use of self-regulated learning strategies yielded a substantial increase in prediction of standardized achievement test scores after the effects of gender and socioeconomic status were removed” (p. 625). Accordingly, the results suggest that teachers and students can focus more on test-taking strategies in ITPPs.

### **4.3 Research question 3: What is the relationship between students’ test performance and washback effects?**

To answer the research question on the relationship between students’ test performance and washback effects, both students’ scores on TOEIC and their responses to the items on washback effects were analyzed. Table 6 shows the mean score, standard deviation and total student number for the TOEIC score and washback effects. The mean TOEIC score was 716.73 (SD=95.80), while the mean for the washback effects was 4.47 (SD=.52). Next, the correlation was calculated. Table 7 displays the correlation coefficients ( $r$ ) between formal TOEIC scores and washback effects of the ITPP on students from the experimental group. The results indicated that there was a negative correlation between test performance and washback effects ( $r = -.09$ ); however, it did not reach the significant level. In other words, there was a negative correlation between these two variables, but the magnitude was negligible.

**Table 6. Descriptive statistics**

	M	SD	N
Washback effects	4.47	.52	52
Formal TOEIC score	716.73	95.80	52

**Table 7. Correlations**

		Formal TOEIC score
Washback effects	Pearson correlations	-.09
	Sig. (2-tailed)	.52
	N	52

In addition to the quantitative results on the positive washback effects ( $M=4.47$ ), qualitative analyses also unveiled that students believed that the ITPP had positive washback effects on them. When asked if they thought it was possible to make an improvement or achieve their goals on formal TOEIC scores before the scores came out, most of the interviewees claimed that they possessed great confidence in their scores. Interviewee #E5 explained that “Of course, if I do not reach my expected score, I will feel frustrated since I have already put a lot of effort into it. On the other hand, I will have a sense of achievement and have the incentive to move forward if I attain my ideal goal.” Yet, only six out of 11 interviewees fulfilled the goal they set for themselves. Interviewee #E6 responded that he believed the ITPP brought about positive washback effects on him no matter how much he improved his score. He enjoyed the test-preparation program a lot. In fact, English learning was not all about the scores in the proficiency tests. Additionally, the majority of the students gave a positive answer without a doubt.

Interviewee #E2 even commented that “I think that ITPP does really have its efficiency. I can integrate what I learned in ITPP into my own learning methods, which gives me more confidence in learning English. I feel one step closer to English. The improvement that I made on formal TOEIC helped build up my confidence. I think the school should hold more courses like ITPP as its subsequent effects were evident.”

In brief, students from the experimental group experienced positive washback effects ( $M=4.47$ ) from the ITPP, while their test performance did not seem to have significantly improved compared to the washback effects. As the students were high language achievers with high motivation, they were likely to make slight but not considerable progress on their TOEIC scores. In response to Cheng’s study in 2004, she stated that “language test scores cannot be interpreted simplistically as an indicator of particular language ability we want to measure. The scores are also affected by the characteristics and contents of the test tasks, the characteristics of the test takers, the strategies test takers employ in attempting to complete the test tasks, as well as the inferences we draw from the test results. These factors undoubtedly interact with each other” (p. 5). The results were similar to the findings from the interview data, suggesting that test scores cannot represent students’ true language proficiency or their growth in learning English.

## 5 Conclusion

This study investigated the washback effects of an Intensive Test-Preparation Program (ITPP) on 112 non-English majors in a university in Taiwan. The results gathered from a questionnaire, interviews, and test scores indicated that the ITPP generated some effects on non-English majors in their English learning. The results can be summed up as follows:

1. ITPPs can motivate students’ English learning.
2. The test-taking strategies in ITPPs contributed positive effects on students’ future English learning.
3. Students experienced positive washback effects from the ITPP, while their test performance did not seem to have achieved significant improvement compared to the washback effects.

As a whole, the ITPP significantly influenced students’ motivation and strategies for learning English. These results indicated that elements of ITPPs such as the environment, teaching styles, and time arrangement can effectively motivate students, and that test-taking strategies should be taught more in ITPPs.

### **5.1 Implications**

Non-English majors should make better use of the test-taking strategies taught in ITPPs to familiarize themselves with the test patterns in the formal TOEIC. Test-taking strategies such as skimming and scanning, identifying different accents, and understanding the basic structure of an English article are reported as useful for test takers. Furthermore, students with less self-discipline and self-regulation are advised to take part in an ITPP, as it provides a fixed schedule and supportive learning environment. Although this study showed that some students did not reach their expected goals after the ITPP, they all made progress in learning English. Since language learning is a long-term process, a lot of improvement may not be apparent in a short period of time; however, ITPPs may successfully sustain students' motivation.

As for teachers, the findings of the study show that energetic and lively teaching styles are effective ways to motivate students' learning, and that most of the students in ITPPs are greatly influenced by the teachers. Aside from that, teachers should put emphasis on test-taking strategies in an intensive program to alleviate students' anxiety about taking high-stakes tests.

In terms of curriculum design, the current study shows that an appropriate environment and proper time arrangement of an intensive course can effectively facilitate students' motivation and help students form the habit of studying English.

### **5.2 Limitations & future directions**

This study has three limitations that future research could improve upon. First of all, the study could have recruited more participants, because a larger sample size could reduce the effect of extreme quantitative data or ineffective outliers. Sufficient sample size can allow the researchers to explore the unknown potentials and assemble a holistic picture for data analysis. In this present study, the ITPP only recruited 52 students in two semesters. Future research that recruits sufficient participants would yield more convincing results. Next, the interviews could have been conducted again after a longer period of time, which could enhance the reliability and reconfirm the consistency of the results. Based on Cheng (2008), it is clear that the future direction of washback and impact studies that investigate the consequences of language testing needs to be multiphase, multimethod and longitudinal in nature. Washback and impact of testing take time to evolve, and therefore longitudinal studies are essential with repeated observations (and measures) of the classroom teaching, including teachers and students as well as policy, curriculum, and assessment documents. Also, researchers need to be immersed in the educational system interacting with a wide range of stakeholders. In the current study, the researchers undertook the interviews right after the ITPP, which might not enable researchers to explore possible changes in participants' English learning. Hence, in future investigations, delayed interviews could help elicit more in-depth data on the changes in students' motivation and strategy use. Another issue that future studies can explore pertains to the duration of the ITPP. In the present study, the ITPP lasted only 5 weeks. Whether longer ITPPs can bring about more positive washback effects can be further examined. All of these directions for future studies can contribute to a greater understanding of washback effects of ITPPs on EFL students' learning experiences.

### **Acknowledgement**

We would like to express gratitude to the Department of Applied Foreign Languages and Language Center at the National Taiwan University of Science and Technology for their administrative support for the research project. We also thank Che-Jui Jerry Chang and Kai-Ting Amy Chang for their help in the data collection and analysis, the participants for taking part in this project, and the editors and anonymous reviewers for their valuable feedback. The research was funded by the Ministry of Science and Technology, Taiwan. (Grant no.: 105-2815-C-011-029-H).

## References

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied linguistics*, 14(2), 115–129.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bailey, K. M. (1999). *Washback in language testing*. Princeton, NJ: Educational Testing Service.
- Brown, J. (1998). An investigation into approaches to IELTS preparation, with particular focus on the academic writing component of the test. *IELTS Research Reports*, 1, 20–37.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10(1), 15–42.
- Burton, S., & Nesbit, P. (2002, November). *An analysis of student and faculty attitudes to intensive teaching*. Paper presented at the Conference of Celebrating Teaching at Macquarie, Sydney, Australia.
- Carroll, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals*, 1, 131–51.
- Chen, L. (2002). *Taiwanese junior high school English teachers' perceptions of the washback effect of the basic competence test in English* (Unpublished doctoral dissertation). Ohio State University, Columbus, OH, USA.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279–301.
- 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, NJ: Lawrence Erlbaum.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge, England: Cambridge University Press.
- Cheng, L. (2008). Washback, impact and consequences. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (pp. 349–364). New York, NY: Springer.
- Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, 6(3), 399–403.
- Chu, H. Y. (2009). *Stakes, needs and washback: An investigation of the English benchmark policy for graduation and EFL education at two technological universities in Taiwan* (Unpublished doctoral dissertation). National Taiwan Normal University, Taipei City, Taiwan.
- Collins, L., Halter, R. H., Lightbown, P. M., & Spada, N. (1999). Time and the distribution of time in L2 instruction. *TESOL Quarterly*, 33, 655–680.
- Daniel, E. L. (2000). A review of time-shortened courses across disciplines. *College Student Journal*, 34, 298–308.
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 191–200). Mahwah, NJ: Lawrence Erlbaum.
- Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, 55(4), 575–611.
- Green, A. (2007). *IELTS washback in context: Preparation for academic writing in higher education*. Cambridge, England: Cambridge University Press.
- Hawkey, R. (2006). *Impact theory and practice: Studies of the IELTS test and ProgettoLingue 2000*. Cambridge: Cambridge University Press.
- Jacques-Bilodeau, M. (2010). *Research project: A study of the long term effects of Intensive English programs on secondary school ESL students*. Retrieved from [www.csbe.qc.ca/projetrd/doc\\_projet/marie\\_jac\\_bil\\_2010.pdf](http://www.csbe.qc.ca/projetrd/doc_projet/marie_jac_bil_2010.pdf)
- Jin, Y. (2006). Improving test validity and washback: A proposed washback study on CET4/6. *Foreign Language World*, 2006(6), 65–73.
- Messick, S. (1982). Issues of effectiveness and equity in the coaching controversy: Implications for educational and testing practice. *Educational Psychologist*, 17, 67–91.
- Miller, K. (2003). The pitfalls of implementing TOEIC preparation courses. *Yoshinogawa Review*. Retrieved from <http://www2.chikoku-u.ac.jp/english-dept/pitfalls.html>.
- Mizumoto, A & S. Takeuchi. (2008). Exploring the driving forces behind TOEIC scores: Focusing on learning strategies, motivation, and study time. *JACET Journal* 46, 17–32.
- Mukundan, J., Mahvelati, E. H., & Nimehchisalem, V. (2012). The effect of an intensive English program on Malaysian secondary school students' language proficiency. *English Language Teaching*, 5(11), 1–7.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.

- Pan, Y. C. (2014). Learner washback variability in standardized exit tests. *TESL-EJ*, 18(2), n2.
- Pan, Y., & Newfields, T. (2012). Tertiary EFL proficiency graduation requirements in Taiwan: A study of washback on learning. *Electronic Journal of Foreign Language Teaching*, 9(1), 108–122.
- Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459–470.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing*, 22(2), 142–173.
- Robb, T. N., & Ercanbrack, J. (1999). A study of the effect of direct test preparation on the TOEIC scores of Japanese university students. *Teaching English as a Second or Foreign Language*, 3, 25–55.
- Serrano, R. (2010). Development of English language skills in oral production by adult students in intensive and regular EFL courses. *Vigo International Journal of Applied Linguistics*, 7, 99–115.
- Serrano, R., & Muñoz, C. (2007). Same hours, different time distribution: Any difference in EFL? *System*, 35, 305–321.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29.
- Stern, H. H. (1985). The time factor and compact course development. *TESL Canada Journal*, 3, 12–27.
- Stoneman, B. W. H. (2006). *The impact of an exit English test on Hong Kong undergraduates: A study investigating the effects of test status on students' test preparation behaviours* (Unpublished doctoral dissertation). The Hong Kong Polytechnic University, Hong Kong, China.
- Watanabe, Y. (2004). Methodology in washback studies. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 19–36). Mahwah, NJ: Lawrence Erlbaum.
- Wu, J. (2008). Views of some Taiwanese students and teachers on English language testing and assessment. *Research Notes*, 34, 6–9.
- Xie, Q. (2013). Does test preparation work? Implications for score validity. *Language Assessment Quarterly*, 10(2), 196–218.
- Yu, C. (2010). *Reliability of self-report data*. Retrieved from <http://www.creative-wisdom.com/teaching/WBI/memory.shtml>
- Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614–628.

## Appendices

### Appendix A

#### Students' questionnaire for the main study – the experimental group

Dear students:

We would like to know your own experience toward learning English in universities, and ask for your opinions about the TOEIC you are going to sit soon. *All information you provide here will be treated in the strictest confidence.* Thank you for your valuable time.

#### **PART ONE Please tick the appropriate answer.**

1. Your gender (1)  Male (2)  Female
2. Your grade level (1)  Freshman (2)  Sophomore (3)  Junior  
(4)  Senior (5)  Master-degree (6)  PhD
3. Your department (1)  Engineering (2)  Electrical Engineering (3)  Applied Sciences  
(4)  Management (5)  Design-related (6)  Others
4. At what level did you begin to receive lessons in English?  
(1)  Kindergarten (2)  Primary school (3)  Junior high school  
(4)  Senior high school (5)  Universities/ Colleges
5. What is the medium of instruction your teacher usually uses to teach you English?  
(1)  English only  
(2)  English supplemented with occasional Chinese explanation  
(3)  Half English and half Chinese  
(4)  Mainly Chinese

6. Have you ever participated in any program mainly about the TOEIC?  
 (1)  Yes (2)  No
7. To above the question, how long did the program last for?  
 (1)  within one month (2)  one to three months  
 (3)  six months (4)  more than six months
8. What's instruction way of your teacher usually uses to teach you English?
- 

**PART TWO Please circle the following on 6-point scale format where 1= strongly disagree, 2= disagree, 3= slightly disagree, 4= slightly agree, 5= agree, 6= strongly agree.**

9. In the following aspects, how are you affected by the ITPP? Self-image motivation to learn peer relationship future employment opportunities teacher and student relationship anxiety and emotional tension.

	strongly negative	negative	slightly negative	slightly positive	positive	strongly positive
(1) self-image	1	2	3	4	5	6
(2) motivation to learn	1	2	3	4	5	6
(3) peer relationship	1	2	3	4	5	6
(4) future employment opportunities	1	2	3	4	5	6
(5) teacher and student relationship	1	2	3	4	5	6
(6) anxiety and emotional tension	1	2	3	4	5	6

10. How do you agree with the following reasons for learning English?

	strongly disagree	disagree	slightly disagree	slightly agree	agree	strongly agree
(1) To acquire basic knowledge and forms of English	1	2	3	4	5	6
(2) To be able to graduate	1	2	3	4	5	6
(3) To get a better job	1	2	3	4	5	6
(4) To be able to communicate with people	1	2	3	4	5	6
(5) To meet the requirements of the society	1	2	3	4	5	6

(6) To be able to watch English movies and listen to English programs	1	2	3	4	5	6
(7) To have more and better opportunities in the future	1	2	3	4	5	6
(8) To enter a high school	1	2	3	4	5	6
(9) To fulfill parents' and teachers' expectation	1	2	3	4	5	6
(10) To understand the cultural background of English-speaking countries	1	2	3	4	5	6
(11) To speak English fluently	1	2	3	4	5	6
(12) To be able to handle the job related to English better than colleagues in the future.	1	2	3	4	5	6
(13) To be able to make an English presentation for future over-sea clients, foreign colleagues and boss.	1	2	3	4	5	6
(14) To be able to write emails in English and communicate with over-sea clients, foreign colleagues and boss.	1	2	3	4	5	6

11. How often do you use the following strategies for learning English?

	never (0%)	seldom (20%)	sometimes (40%)	often (60%)	usually (80%)	always (100%)
(1) Play language games	1	2	3	4	5	6
(2) Take mock tests	1	2	3	4	5	6
(3) Study textbooks	1	2	3	4	5	6
(4) Communicate in English	1	2	3	4	5	6
(5) Read newspapers, articles, magazines	1	2	3	4	5	6
(6) Listen to radios	1	2	3	4	5	6
(7) Watch TV programs, movies	1	2	3	4	5	6
(8) Memorize vocabulary	1	2	3	4	5	6
(9) Learn grammar	1	2	3	4	5	6
(10) Join extracurricular activities	1	2	3	4	5	6
(11) Watch video clips	1	2	3	4	5	6



(12) Learn through English learning websites	1	2	3	4	5	6
(13) Memorize idioms and phrases	1	2	3	4	5	6
(14) Imitate native speakers' accents	1	2	3	4	5	6

12. How do you agree with the following opinions?

	strongly disagree	disagree	slightly disagree	slightly agree	agree	strongly agree
(1) Students like examinations	1	2	3	4	5	6
(2) My learning is improved by practicing mock exam.	1	2	3	4	5	6
(3) Examinations force me to study harder.	1	2	3	4	5	6
(4) I work harder when I prepare for the examinations	1	2	3	4	5	6
(5) Taking examinations is a valuable learning experience.	1	2	3	4	5	6
(6) I work hard to achieve my best in examinations.	1	2	3	4	5	6
(7) Examination is one of the motivations for my learning.	1	2	3	4	5	6
(8) My score on an examination is a good indication of how well I have learned the material.	1	2	3	4	5	6
(9) Mock examinations are important ways to learn.	1	2	3	4	5	6
(10) My score on an examination is a good indication of how well I will be able to apply what have been learned.	1	2	3	4	5	6
(11) I like to learn English.	1	2	3	4	5	6
(12) I learn English actively.	1	2	3	4	5	6
(13) The way I use to learn English now is effective.	1	2	3	4	5	6
(14) My English ability is better than my peers.	1	2	3	4	5	6
(15) I believe learning English is useful	1	2	3	4	5	6
(16) Examinations should NOT be used as a sole determiner of student grades.	1	2	3	4	5	6
(17) My score on the listening part of an examination is a good indication of how well I can understand real-life English.	1	2	3	4	5	6

(18) My score on the reading part of an examination is a good indication of how well I can understand real-life English.	1	2	3	4	5	6
--------------------------------------------------------------------------------------------------------------------------	---	---	---	---	---	---

13. How do you agree with the following opinions after you participated in intense TOEIC-oriented program?

	strongly disagree	disagree	slightly disagree	slightly agree	agree	strongly agree
(1) I like the teaching methods the teachers use.	1	2	3	4	5	6
(2) I like the teaching methods the teaching assistants use.	1	2	3	4	5	6
(3) This intensive program is useful for improving my score on an examination.	1	2	3	4	5	6
(4) This intensive program is useful for improving my listening ability.	1	2	3	4	5	6
(5) This intensive program is useful for improving my reading ability.	1	2	3	4	5	6
(6) I feel happy when I am in class.	1	2	3	4	5	6
(7) I like this way of learning English.	1	2	3	4	5	6
(8) I like the time arrangement of the curriculum in this intensive program.	1	2	3	4	5	6
(9) I like the schedule of this program.	1	2	3	4	5	6
(10) I think the school should initiate this kind of program more.	1	2	3	4	5	6
(11) I am satisfied with this program.	1	2	3	4	5	6
(12) My classmates who participated in the program inspired me to keep learning.	1	2	3	4	5	6

*End of the Questionnaire  
Thank you very much for your help.*

## **Appendix B**

### **Semi-structured interview questions – the experimental group**

1. What were your methods and strategies of learning English? Did you change methods and strategies after the ITPP? What methods and strategies did you learn from the ITPP? Do you use those strategies after the ITPP? Do they work?

2. Which elements had the greatest/smallest influence on you during the ITPP (e.g. teachers, classmates, the interaction in class, learning methods, the environment, course design)? Why? In what aspects?
3. Do you think ITPP is effective? Does ITPP change your perspectives on learning English (in positive or negative ways)? Would you like to participate in ITPP again? Why?
4. Before you joined the ITPP, how much time did you spend on learning English? When participating in the ITPP, how much time did you spend on learning English?
5. Before you participated in the ITPP, what were your expectations of this program? Did the program meet your expectations? Were there any positive and negative influences on your learning motivation, relationship among peers, and your feelings?
6. What media do you usually use to learn English (e.g. newspaper, radio program, video, magazine)? Why? Is there any change after you participated in the ITPP?
7. Do you think you actively learn English now? Why/why not?
8. Why did you choose to join the ITPP?
9. In the following aspects, how are you affected by the ITPP? Self-image, motivation to learn, peer relationship, future employment opportunities, teacher and student relationship, anxiety and emotional tension,

#### **Semi-structured interview questions – the control group**

1. Which of the following factors has the largest/smallest influence on you when you learn English (i.e. teachers, classmates, the interaction in class, learning methods, the environment, course design)? Why? In what aspects?
2. What are your methods and strategies of learning English? Do they work?
3. One item in the questionnaire states that “My score on an examination is a good indication of how well I have learned the material.” Do you agree with that? If not, what should be the indicator of how well you have learned the material?
4. One item in the questionnaire states that “My score on the listening/reading part of an examination is a good indication of how well I can understand real-life English.” Do you agree with the statement? Why/ why not?
5. How do you learn English? How much time do you spend on learning English?
6. What media do you usually use to learn English (e.g. newspaper, radio program, video)? Why? How do you use the media?
7. Do you think you actively learn English now? Why/ why not?
8. How do you prepare for the test? Can you explain that in details?