科技部補助

大專學生研究計畫研究成果報告

＊ ********* ** ＊
＊ 計 畫 ＊
＊ ： 英語檢定畢業門檻對於科技大學外語系學生以及畢業 ＊
＊ 名 稱 生的英語學習回沖效應研究 ＊
＊ ********* ** ＊

執行計畫學生： 林佳蓁
學生計畫編號： MOST 103-2815-C-041-009-H
研 究 期 間 ： 103 年 07 月 01 日至 104 年 02 月 28 日止，計 8 個月
指 導 教 授 ： 董幸正

處理方式： 本計畫涉及專利或其他智慧財產權,2 年後可公開查詢

執 行 單 位： 嘉藥學校財團法人嘉南藥理大學應用外語系

中華民國 104 年 03 月 31 日

The Impacts of the English Graduation Benchmark upon Technical University English Majors

Lin Jia-zhen (林佳蓁)
Chia-Nan University of Pharmacy & Science
j829955118@gmail.com

I. Introduction

This paper is intended to explore the washback effects of English graduation benchmark (hereafter EFL benchmark) upon technical university English majors and graduates. Here, the English majors are specifically referred to the daytime English juniors and seniors. Here, two major fields become the major concern for the washback effects of EFL benchmark: (1) education, and (2) society (Alderson and Wall, 1993; Biggs, 1995, 1996; Hamp-Lyons, 1997; Mousavi, 2002: 46; Popham, 1983; 1987). The former encompasses the curricular design, strategic pathways in language teaching, the selection of teaching materials, ways of assessment, and the overall contents of pedagogy. Distinctively, the latter spans wider ranges of washback effects, including the socio-cultural values, policy of education, types of ideology. Additionally, the washback effects operate in bilateral directions (i.e. forward washback effects, and backward washback effects) (Pearson, 1988: 98). The forward washback effects speculates the possible influences of tests to be implemented. The backward washback effects is primarily concerned with the educational or the social reforms in concert with test formats.

The issue of this current research paper, virtually, has been widely detected in the following fields: (1) affection (e.g. English learning motivation, English learning anxiety), (2) educational policy (e.g. the policy of formal English test administration), (3) recognition (e.g. English learning strategies), (4) education (e.g. the test-oriented classroom instruction), and (5) society (e.g. the impacts of formal English language tests) (Cheng Y. L., 2009; Huang L. H., 2010; Chu H. Y., 2008; Chou. C. H., 2010; and Chen T. H., 2008). In summary of the previous research contributions, the EFL benchmark indeed manifests its powers in bolstering students' test achievements (Cheng Y. L., 2009；Chou. C. H., 2010，Chu H. Y., 2008). Besides, the test anxiety empirically exists except for the English majors（Chen T. H., 2008；Chu H. Y., 2008）. Yet, the EFL benchmark narrows its scope to the test achievements, possibly biasing the normal track of EFL study students' (Chu H. Y., 2008). In order to clarify the advantages and the possible negative impacts of EFL benchmark, the writer of this study is intrigued to explore the washback effects of EFL benchmark upon technical English majors and the extended impacts of such a test policy upon English graduates who are currently dedicated themselves in various workforces.

In order to address the aforementioned statements, the following research questions are formulated for further discussion:

Q1 Is the EFL benchmark advantageous to English majors' development in their general English language proficiency?

Q2 Is the EFL benchmark affective to English majors' motivation in their EFL study?

Q3 Is the EFL benchmark negative to English majors' anxiety in their EFL learning and test participation?

II. Literature Review

(I) Introduction

The notion *washback* (Alderson and Wall, 1993) or *backwash* (Biggs, 1995, 1996) refers to the nature of a test on teaching and learning (Mousavi, 2002: 46). In other words, tests also act as the measurement-driven instruction since it ought to bring about changes in teaching and hence learning (Popham, 1983; 1987).

The impact of language tests, on the other hand, is a relatively new topic in language testing (Mousavi, 2002: 305). Briefly defined, the *test impact* refers to the influence of tests on the behavioral changes in the educational context and society. The educational context encompasses curriculum, teaching methods, teaching and learning strategies, material and courseware, assessment practices, and the contents of instruction (Mousavi, 2002: 305). These aspects of impact have also been of particular interest to researchers in the field of washback. The *washback* is briefly defined as the influence of testing upon teaching and learning (Alderson and Wall, 1993; Hamp-Lyons, 1997). Therefore, the terms *impact* and *washback* are interchangeably used when the influence of tests is confined to the educational context.

Nonetheless, *test impact* is distinctively different from *test washback* to some extent. The *washback* has been used in general education such as the curriculum, teaching materials and language teaching (Alderson and Wall, 1993; Hamp-Lyons, 1997). Virtually, the scope of washback is also extended to the roles of teachers, teaching and learning strategies, and the attitudes toward teaching and learning methods (Alderson and Wall, 1993: 121). Mialnovic and Saville (1996) contended that the washback concept was the complex interactions between factors which made up the teaching and learning context.

Impact, compared with washback, is a broader term encompassing the effects of testing not only on individual classrooms but on the educational system as a whole and on society more generally (Hamp-Lyons, 2000: 586). The social aspects

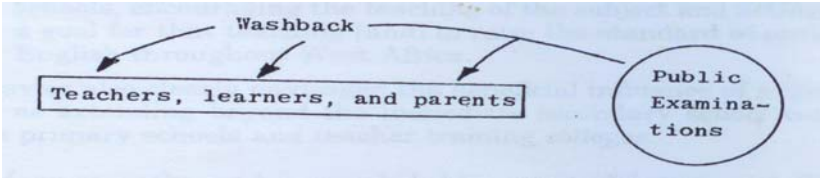
included the social values, the educational policy, cultural differences, ideology, ethics, and morality. For example, Parry (1994) attributed test-takers' poor performance on the reading test to their insensitivity to the foreign cultures in the texts.

In a nutshell, washback is of particular interest in the teaching and learning context. Test impact, however, does not draw exclusively on the educational context. The social, political, and ethnical dimensions are also included for the test impact research issues.

(II) The Ways That Washback Works

In the preceding section, a test is said to exert its influences on teaching and learning after its formal administration. This influence is seen in a backward direction (Pearson, 1988: 98) which is illustrated in the following figure.

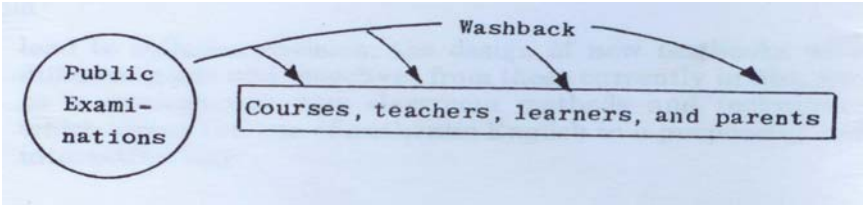
Figure 1 The washback effects of a test in a backward direction



(Pearson, 1988: 98)

According to Figure 1, a test influences the attitudes, behavior, and motivation of participants (teachers, learners, and parents). However, the influence of a test possibly operates in a forward direction as courses are designed in direct response to the features of a test (Pearson, 1988: 98). This is demonstrated in the following:

Figure 2 The washback effects of a test in a forward direction



(Pearson, 1988: 98)

As indicated in Figure 2, a test determines the course design and directions of teaching and learning in class. For example, in the study of Wall and Alderson (1993), a new English test affected the contents of teaching in Sri Lanka such as the increased emphasis on the instruction of listening skills in class. In the past, much attention had been devoted to the reading and writing abilities in class before the introduction of a new test in Sri Lanka.

Pearson (1988: 98-99) made a contribution to the directions of the test washback, but he did not explicitly state the processes of washback. The processes of test

washback were later put forward by Fullan (1991) in the division of three phases: initiation, implementation, and continuation. The initiation process refers to the time that a test first appears and the comments it receives before it is ready for adoption. In order to achieve the intended influences, a test is closely related to three R's: relevance, readiness, and resources. Relevance includes the interaction of need, practitioners' understandings of the innovation, and utility. Readiness involves schools' capacity to initiate, to develop, or to adopt a given innovation. Resources indicate the accumulation and provision of support as a part of the change process.

When teachers perceive the influences of a test, they begin to cope with the changes of test formats (i.e. implementation process). Sometimes, when implementing new test formats, teachers might feel that the effects of new test formats are not as efficient as what they would have thought. Consequently, new test formats seem to bring about overload on teachers and this result corresponds to what Goodlad *et al*. (1970) termed *false clarity* or *painful unclarity*. The former occurs when teachers change their teaching superficially. To put it in another way, teachers seem to change their instructional behaviors, but their attitudes toward the innovative skills remain unchanged. The latter term refers to teachers' frustration, anxiety, and abandonment of effort that they have made in teaching since they are called upon to implement the changes that they really do not understand.

Compared with the study by Pearson (1988), Fullan (1991) pointed out several variables which were determinable to the desired influences that a test promoted: the understanding of new test formats, teachers' attitudes toward the innovative ideas, the affective factors that a new test skill imposed, and so forth. Based on the innovation theory by Fullan (1991), Alderson and Wall (1993) proposed the Washback Hypothesis as follows:

(1) A test will influence teaching.
(2) A test will influence learning.
(3) A test will influence what teachers teach.
(4) A test will influence how teachers teach.
(5) A test will influence what learners learn.
(6) A test will influence how learners learn.
(7) A test will influence the rate and sequence of teaching.
(8) A test will influence the rate and sequence of learning.
(9) A test will influence the degree and depth of teaching.
(10) A test will influence the degree and depth of learning.
(11) A test will influence the attitudes to the contents, method, etc. of teaching and learning.
(12) Tests that have important consequences will have washback.

(13) Tests that do not have important consequences will have no washback.
(14) Tests will have washback on all learners and teachers.
(15) Tests will have washback effects for some learners and some teachers, but not for others.

(Mousavi, 2002: 830)

According to these fifteen hypotheses listed above, a test exerts its influences on all aspects of teaching and learning (See hypotheses 1 and 2), especially the contents of teaching and methodology (See hypotheses from 3 to 6), rates and sequences of teaching and learning (See hypotheses 7 and 8), the quality and quantity of teaching and learning (See hypotheses 9 and 10), teachers' and learners' attitudes toward tests (See hypotheses 11), and the consequences that a test had in teaching and learning (See hypotheses from 12 to 15). The hypotheses (1) and (2) are the most general ones in which anything being related to the influences of tests on teaching and learning (e.g. teachers' and students' attitudes toward the contents of textbook or ways of textbook compilation) is included for discussion. The hypotheses from 3 to 6 refer to the influence of tests on the contents of teaching and methodology. Liying Cheng (1997: 36), for example, investigated the washback effects of public examinations on the curriculum change in Hong Kong. In Cheng's words, the positive washback occurred when the exam acted as a vehicle for an intended curriculum change. However, the negative washback appeared when teachers exclusively focused on the test-taking skills that were related to the public examinations instead of the language learning activities in the textbooks. Alderson and Wall (1993) observed the impact of a new English examination in Sri Lanka on the contents of language teaching, teaching methodologies, and ways of assessment. They discovered that the new English exam was likely to bring about the positive washback effects on the teaching materials when these teaching materials resembled text types in the exams. In addition, teachers would be writing tests that mirrorred the contents of textbooks or *Teachers's Guide*. However, the negative washback occurred when teachers limited their instructional strategies as well as teaching materials to the test item types in the new English exam which was introduced by Alderson and Wall (1993). What's worse, in order to make students familiarize with the new English exam, teachers tended to adapt practice tests from publications designed to prepare students for the exam.

The former six hypotheses are concerned with content and ways of teaching which are profoundly influenced by tests. The hypotheses 7 and 8 refer specifically to rates and sequences of teaching and learning (Alderson and Wall, 1993). Take the Sri Lanka study for example. The new O-level English exam led to the changes of teaching procedures such as the primacy of the listening and speaking abilities in class.

In addition to the teaching rates and orders, the quality and quantity of teaching and learning was assumed to be influenced by the language tests. Leinhardt and her colleagues (Cooley & Leinhardt, 1981; Leinhardt, 1983; Leinhardt & Seewald, 1981) maintained that overlap between what was taught and what was tested served as an important indicator of test performance. For instance, Cooley and Leinhardt (1981) discovered that students performed well on tests if they had been taught the specific materials covered by the test and if they had been frequently exposed to the test format. By the same token, Saville (2004) investigated the washback effects of IELTS on teachers' perceptions towards their changes of the language teaching methods and materials. Consequently, 90 percent of the participant teachers agreed that the IELTS exerted great influences on their selection of teaching materials (Saville, 2004). In addition, 63 percent of teachers responded that IELTS made their teaching methods more test-oriented. In addition, influenced by the reading skills measured in test items in IELTS, teachers developed test-relevant communicative micro-skills such as identifying main points, identifying overall meaning, predicting information, retrieving and stating factual information, planning and organizing information, and distinguishing fact from opinion (Saville, 2004). Moreover, the hypothesis 11 deals with teachers' attitudes toward the influence of tests, especially the nationwide tests. In Taiwan, Chang Wu-chang (1996) surveyed senior high school teachers' attitudes toward the English composition and Chinese-English translation tests which were severely criticized to call for abolishment due to a dearth of objective scoring scales. Notwithstanding this, 61 percent of English teachers in Chang's written survey (1996) strongly suggested to remain these two test formats under the consideration of the ETJCEE washback effect. Given that the writing tasks had been abolished, senior high school students would have been reluctant to compose their writings.

Although these 15 hypotheses listed above seem to exist individually, they are indeed mutually interrelated. For example, when tests are used as levers for change, the design of new teaching materials is amenable to the innovative test formats. In addition, new tests are believed to result in the curricular reforms and the innovation of teaching and learning approaches (Biggs, 1995). Furthermore, the fifteen washback hypotheses by Alderson and Wall (1993) had strong bearings on the validity of a test (i.e. consequential validity). This kind of relationship, however, was contingent upon the effects of the test itself on the aspects of teaching and learning (Alderson and Wall, 1993; Buck, 1988; Hughes, 1989; Shohamy, 1992). In other words, to achieve the positive washback effects, a test has to be valid and reliable. Hughes (1989) proposed the following ways to promote the positive washback effects which were closely relevant to the test validity and reliability:

1. Test the abilities whose development you want to encourage.
2. Sample widely and unpredictably.
3. Use direct testing.
4. Make testing criterion-referenced.
5. Base achievement tests on objectives.
6. Ensure that the test is known and understood by students and teachers.
7. Where necessary provide assistance to teachers.

According to these seven principles listed above, only the efficient tests led to the beneficial and powerful impacts upon teaching and learning objectives. However, the influence of testing was not limited to teaching and learning. Other aspects such as the entire educational system, curricular design, and teaching materials were also possibly influenced by tests. McClung (1979) pointed out three kinds of validity in relation to the washback effects of tests: content validity, curricular validity, and instructional validity. The content validity referred to the link between test contents and educational objectives. The curricular validity was concerned with the relationships between the test contents and materials used in schools. The instructional validity was defined as the extent to which test material had been taught in the classroom (Airasian & Madaus, 1983). Frederiksen and Collins (1989) added the systemic validity in which a test acted as a promoter of the whole teaching and learning system. Further, Alderson and Wall (1993), according to fifteen washback hypotheses listed above, re-classified test validity as four types: (1) washback validity, (2) systemic validity, (3) content validity, and (4) consequential validity (Alderson and Wall, 1993). The *washback validity* was a broader term which was believed that the validity of a test had to be gauged by the degree to which it had a positive influence on teaching (Morrow, 1986). In other words, this kind of validity was centered on the relationship between the quality of a test and associated teaching (Cheng, 1998; Wall & Alderson, 1993; Wall, 1999). The *systemic validity* referred to the effects of language tests on the curricular reform or on educational system changes (Frederikson & Collins, 1989; Kellaghan, Madaus, & Airasian, 1982; Smith, 1991; Vernon, 1956). Messick (1996), for example, noted that a test's validity had to be appraised on the grounds of the washback that it manifested to the curriculum reform.

The third type of validity, the *content validity*, is concerned with the influence of a test on the textbook exercises (Wall, 1993). A test is said to have the content validity if it greatly influences the contents of textbooks (i.e. text types or tasks appearing on the exam), methods of teaching (i.e. the close relationship between the in-class instruction and the skills being assessed on the exam), and ways of assessment (e.g. marking students' work by the criteria laid down in the textbook).

The *consequential validity* which was proposed by Messick (1989) extended the

scope of test validity to the political and social contexts (cited in McNamara, 1998: 304-305). Messick's discussion arouse from the upheaval in social science research in which the social order no longer followed the disciplinary procedures (Lynch, 1994; Lynch and Hamp-Lyons, 1996). Instead, the minds of individuals (Jacoby and Oches, 1995) and the rapidly changing environments became the major concerns in the language testing research. For example, when labor mobility led to the immigrant flows, the assessment of language skills had assumed the importance in the context of immigration policy (Davies, 1996; Hawthorne, 1996; 1997).

Thus, according to types of validity mentioned above, a test served as the *measurement-driven instruction* when it created impacts upon the educational processes, the goals of teaching and learning, the educational innovation, and the narrowing contents of the curriculum prior to the test. In addition, types of validity (i.e. washback validity, systemic validity, content validity, and consequential validity) were actually in the complex interplay. For this reason, a unified validity concept was called for. Messick (1975, 1989) came up with an assessment model, which was designed to make inferences about the test scores, variables operating within the social context, and the educational effectiveness. Hughes (1993) proposed the notion of trichotomy, in which washback took place in the following three aspects: (1) participants, (2) processes, and (3) products. Each of them was mutually affected. Table 1 presented these three major aspects as follows.

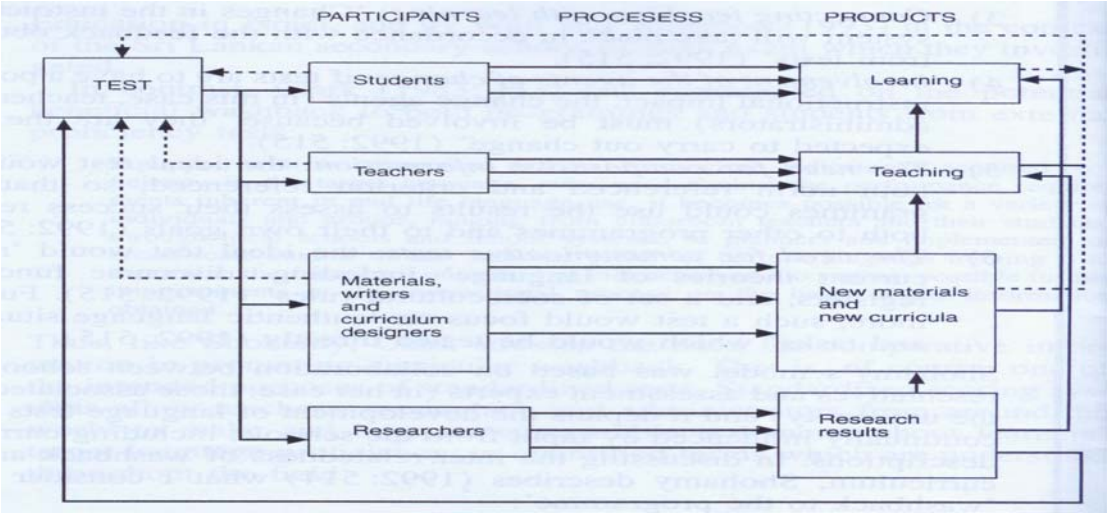
Table 1 The Trichotomy Backwash Model

(a) Participants-students, classroom teachers, administrators, material developers and publishers, whose perceptions and attitudes toward their work may be affected by a test.
(b) Processes- any actions taken by the participants which may contribute to the process of learning.
(c) Products-what is learned (facts, skills, etc.) and the quality of the learning.

(Hughes, 1993: 2)

On the basis of the Trichotomy Backwash Model (Hughes, 1993) and fifteen Washback Hypotheses (Alderson & Wall, 1993), Bailey (1996) developed a basic model of washback, as illustrated in the following figure:

Figure 3 Bailey's (1996) basic model of washback



As Figure 3 indicated, the model of washback in Bailey (1996), the same as Hughes (1993), consisted of three major aspects: participants, processes, and products. The participants included teachers, students, material writers, curriculum designers, and researchers. The processes of washback, as seen in Figure 1, were operated in both the forward and backward directions. The forward direction in the black lines referred to the influences of tests on participants. The backward direction in the dotted lines indicated the possible influences from the participants and products on the test. Virtually, the processes of washback in Bailey (1996) had been put forward by Pearson (1988), but Pearson did not elaborate the variables which affects the directions of washback. Finally, the products in Bailey's model (1996) were invented through the processes of action taken by participants such as the actions of learning, teaching, new materials, and research results.

In conclusion, the ranges of test washback effects on teaching and learning are broader and more complicated than what we would have thought since a test exerts its influences on participants, teaching and learning process, and the products (e.g. teaching material, curricular design) which are virtually in the complex interplay. However, the washback models and hypotheses mentioned above barely describe the positive and detrimental effects of tests and these effects are elaborated in the following section.

(III) The Positive And Negative Washback of Tests

A test is generally believed to promote and reward better teaching and learning. For instance, a 'Spoken English' exam exercised a strong influence on the goal of foreign language instruction when being introduced in West Africa (Davis, 1968). In a similar vein, Johnson and Wong (Read, 1981), discovered that the 'Scaling Test of the Junior Secondary Education Assessment System', being introduced in Hong Kong in the year of 1982, led to the syllabus revision, the use of authentic English, the

change of textbook design and classroom teaching methods, etc. However, tests may bring about the detrimental effects upon educational practices when they impose the restrictions on curricula, teachers, and students (Oxenham, 1984: 113). In addition, the harmful influences of tests would make teachers and students under great pressure such as anxiety, shame, anger, and inferiority. Thus, every test brought about beneficial and detrimental washback effects (Alderson and Wall, 1993) and we attempted to point out factors that led to the positive and negative washback effects, respectively in this section.

To begin with, as noted earlier, Hughes (1989) suggested seven ways to promote the positive washback effects, which were closely relevant to the test validity and reliability. These seven ways in Hughes (1989) had been listed in the earlier section, and it was replicated as follows:

1. Test the abilities whose development you want to encourage.
2. Sample widely and unpredictably.
3. Use direct testing.
4. Make testing criterion-referenced.
5. Base achievement tests on objectives.
6. Ensure that the test is known and understood by students and teachers.
7. Where necessary provide assistance to teachers.

Hughes (1989) provided the groundwork for Bailey (1996: 268-272) to add two variables which led to the positive test washback:

8. Exams should reflect the goals of classroom instruction.
9. Exams should contain authentic examples of communicative behavior.

According to the principles summarized from Hughes (1989) and Bailey (1996), the positive washback effects on teaching and learning occurred when tests reflected the goals of classroom instruction as well as validly and authentically measured the defined language proficiency. In addition, both instructors and pupils had to be aware of the test formats, but they could not teach or learn to the test since the tests were unpredictable. To avoid the test-oriented instruction, test makers had to create a variety of tasks in a test so that test takers could not predict and practice particular test formats.

The variables of positive washback effects in the studies by Hughes (1989) and Bailey (1996) also provided the groundwork for Lee (2000), who described the washback effects in the division of two categories: dynamic and static effects. The former referred to the use of valid tests to improve instructional and learning effectiveness, while the latter the use of tests to evaluate the merits and demerits of the current instructional program. The following table which was made by the writer of this study helped to clarify the dynamic and static effects of tests.

Table 2 The model of washback effects in the study of Lee (2000)

Washback	Dynamic	Static
Positive	The use of valid tests to promote the effectiveness of teaching and learning	The use of valid tests to identify the merits and demerits of the current instructional practice
Negative	The use of invalid tests to determine the goal of teaching and learning	The use of invalid tests to evaluate the instructional program or to identity the proficiency levels of students

As indicated in Table 2, only a valid test achieved the intended goals of education and induced more positive results of teaching and learning. However, Lee (2002) was concentrated solely on the validity of tests when giving accounts of the positive and deleterious washback effects. Other factors such as teachers' perception toward the tests were also determinable to the intended washback effects of tests. According to the research studies relevant to this issue (Alderson and Wall, 1993; Aski, 1998; Cheng, Watanabe & Curtis, 2004; Davies, 1968; Messick, 1996; Noble and Smith, 1994; Vernon, 1956; Wall, 1996), the following four aspects seemed to be helpful in attempting to sketch out the desired washback effects of tests: (1) goals of classroom instruction, (2) unpredictability in testing, (3) authenticity in testing, and (4) teachers' awareness of test demends.

The goals of classroom instruction, as Valette (1992) described, were classified into the *proficiency-first* programs and the *grammar-oriented* or *accuracy-first* programs. The former emphasized creative language use and tolerated grammatical inaccuracy if the message was comprehensible. The latter program (i.e. grammar-oriented program) was primarily concerned with the correct manipulation of the grammatical rules, which was regarded as the prerequisite for the effective communication. Generally speaking, scant attention was paid to the sub-skills in language instruction because these skills were not greatly effective to students' communication ability. Virtually, as Hughes (1989) suggested, these two programs had to be interrelated in the language teaching so that they would achieve the positive classroom washback effect. In other words, in the language instruction, we had to make a balance between the communication-oriented curriculum and the form-focused instructional methods. By doing so, teachers and students would not concentrate on particular skills.

Next, the notion *unpredictability* referred to a variety of tasks included in a test

so that test takers would not predict the test formats in the incoming tests. Otherwise, the negative washback effects of assessment produced if teachers composed tests which mirrorred the contents of past examinations. In order to administer an unpredictable test, test developers had to sample widely. Since a test contained the full ranges of abilities, test takers could not concentrate on particular abilities (Aski, 1998: 480; Noble and Smith, 1994).

The ranges of abilities in a test, as Madsen (1983) indicated, were broadly categorized into *the sub-skills* (e.g. pronunciation, vocabulary, and grammar) and *the communication-skills* (e.g. listening, speaking, reading, and writing). The traditional discrete-item exercises had been criticized not to reflect the real communication abilities. Nevertheless, Aski (1998: 481) argued that discrete point tests still positively influenced students' learning when this type of test evaluated students' ability to produce grammatically correct responses. It was true that a communication-skill test relatively measured how test takers manifested linguistic forms in tasks, but this kind of test did not draw much attention to the accuracy of linguistic elements. Consequently, students' oral or written expression was comprehensible but structurally inaccurate. In order to reduce the negative washback effects of these two types of tests (i.e. discrete-point tests and communication-skill tests), Omaggio and Omaggio-Hadley (1980) advocated a *hybrid exam*, which was a blend of open-ended discourse-based activities with the discrete-item exercises. As Omaggio *el al.* (1980) noted, language learners in a hybrid exam demonstrated their knowledge of specific linguistic points and how they functioned these linguistic elements in natural discourse.

Further, in discussing the positive washback effects of a test, Wiggins (1994) and Bailey (1996: 268-272) noted the importance of authenticity, in which exams had to contain authentic examples of communicative behavior. Wiggins (1994) noted that a test was not an add-on to instruction, nor a set of decontextualized drills or exercises. Instead, a test aimed for the truly authentic representations of L2 communicative interactions. However, according to Shohamy and Reves (1996: 54-57), Spolsky (1985: 37), and Stevenson (1983: 163), contexts created in a test were artificial by its nature and language used in a test was the "authentic test language" rather than the real life language. Nevertheless, it was still possible to create a pseudo-authentic test when this kind of test mirrored test takers' real life situation or nearly the same as the communicative, interactive, and contextualized class activities.

Moreover, teachers ought to be aware of the test formats and their requirement if a test primarily served as a guidance to change instruction. Teachers' perception toward a test did not mean they purported to teach for the exam. Instead, teachers' attention to the new test formats was to adjust their teaching methods. For instance,

in the study by Alderson and Wall (1993), the revised Sri Lankan "O" level English exam had imposed considerable impacts on the content of English lessons and on the way teachers designed their classroom tests. Likewise, Liying Cheng (2004: 147) investigated teachers' perceptions toward the new exam formats of the 1996 Hong Kong Certificate Examinations in English (HKCEE). Because much weight was attached to the oral and listening components in the 1996 HKCEE, most of the teachers in Cheng's study agreed to make more efforts in the teaching of listening and speaking in class.

Teachers' perceptions toward new test formats were beneficial to their language instruction in class. However, the detrimental effects of tests produced if teaching and learning were both directed effectively to passing examination papers (especially new test formats) rather than to mastering language learning activities (Davies, 1968; Wiseman, 1961). Also, the negative washback effects of new test formats existed when they narrowed curricular offerings and instruction modes (Smith, 1991).

Finally, the *age and experience* of teachers had bearings on the washback effect of tests. Usually, the experienced teachers were believed to skillfully adopt the new test item types in their test design or in the classroom instruction because inexperienced teachers felt greater anxiety and accountability pressure in language teaching (Fish, 1988).

In a nutshell, according to four major factors described above, the washback effects of testing depended on *who* it is that actually conducted the investigation within a particular education context, as well as *where*, the school or university contexts, *when*, the time and duration of using such assessment practices, *why*, the rationale, and *how*, the different approaches used by different participants within the context.

III. Research Design

This sections sketches the test participants, research tools, procedures of investigation, and ways of analysis.

The test participants in this research paper are invited from the daytime English juniors (N = 81), English seniors (N= 79), and English graduates (N = 89) at one university of technology in the southern Taiwan. They are requested to supply their written responses to the provided questionnaire by the writer of this research paper (See the Appendix). The self-designed written questionnaire is broadly categorized into two major sections: (1) the washback effects of EFL benchmark for English majors, and (2) the extended impacts of EFL benchmark on English graduates. The former section encompasses arrays of test items perceiving the wielding powers of EFL benchmark in the fields of test achievements, test preparation, preferred test tools

for assessment, and the career management in the future. The latter section lays the primary emphasis on the effects of EFL benchmark in English majors' actual performances in their workforce, such as their foreign language behaviors and possible advantages in their job promotion. Afterwards, the written responses from selected English majors are further analyzed by means of one-Way ANOVA and the Independent T-test.

IV. Research Results

(I) The Impacts of Graduation Benchmark of English Proficiency Tests

As Table 1 presents, most of the participants presented the positive attitudes toward the EFL benchmark in relation to the English proficiency improvement. This can be further warranted by the average scores, which falls beyond Score 3. In particular, due to the pressure on the forthcoming general English proficiency tests in the mid-May (i.e. CSEPT Level 2), the surveyed English juniors treated the EFL benchmark as a main driving force to their test preparation. That is the major reason why English juniors are comparatively more supportive of the EFL benchmark.

Table 1 Graduation Benchmarks and English Proficiency Improvement

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.16	81	.37
Seniors	3.01	71	.49
Graduates	3.15	89	.55

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.15 .01	.08 .07	.18 .99	.05 .17	.34 .20
Seniors	.15 .13	.08 .08	.18 .23	.34 .33	.05 .06
Graduates	.01 .13	.07 .08	.98 .23	.20 .06	.17 .32

2. The Impacts of EFL Benchmark on English Majors' Early Test Preparation

Now that the EFL benchmark is officially regulated as the main threshold to examine English majors' general English language proficiency levels, the surveyed

research participants are supposedly diligent in their test preparation. Yet, based on the results in Table 2, both English juniors and graduates voiced their counter opinions upon the expected assumption, keeping their normal track of English learning (Juniors: 2.95；Graduates: 2.88). Comparatively, several of the invited English seniors are loaded with the EFL benchmark, and they tend to be moderately positive toward this issue.

Table 2 The Impacts of EFL Benchmark on English Majors' Early Test Preparation

Students	Average Scores	Number	Standard Deviation
Juniors	2.95	81	.52
Seniors	3.01	71	.49
Graduates	2.88	89	.62

3. The Dedication to Passing the EFL Benchmark at the Expense of Other Academic Studies

This issue is concerned with the increasing weight of EFL benchmark upon English majors' academic achievements. Virtually, there reaches agreement between Table 2 and Table 3. Now that a handful of sampled English majors map out of their test preparation early, they are surely inclined to dedicate themselves to the general English language proficiency tests at the expense of other academic studies. This echoes English graduates' written responses (capturing 1.80 scores on this issue), devaluing the EFL benchmark as the major juggernaut to dominate their overall English learning program on campus.

Table 3 The Dedication to Passing the EFL Benchmark

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.40	81	.63
Seniors	2.80	71	.50
Graduates	1.80	89	.66

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.40 .61	.10 .09	.00 .00	.64 .38	.15 .84
Seniors	.36 1.00	.10 .10	.00 .00	.15 .77	.64 1.24
Graduates	.61 1.00	.09 .10	.00 .00	.84 1.24	.38 .77

4. The EFL Benchmark and the Test-Oriented English Language Study

This issue remains the similar direction as the previous two issues (i.e. early test preparation and the sacrifice for other academic studies), navigating the effects of EFL benchmark to the test-oriented English language study. As Table 4 indicates, there presents the downgrading consensus upon test-oriented English language study in concert with the EFL benchmark (Juniors: 2.81; Seniors: 2.40; Graduates: 1.71). The lower response scores from English graduates (i.e. 1.71 average scores) give birth to the statistically significant differences via the ANOVA calculation. To put it alternatively, most of the written respondents in this research paper are quite positive to the whipping effects of the EFL benchmark (i.e. the test score improvement). Yet, these invited research participants decline to distort their normal track of EFL learning despite the EFL benchmark greatly affects their academic certificate procurement.

Table 4 The EFL Benchmark and the Test-Oriented English Language Study

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.81	81	.63
Seniors	2.40	71	.64
Graduates	1.71	89	.64

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.42 .10	.10 .10	.00 .62	.16 .15	.68 .34
Seniors	.42 .32	.10 .10	.00 .00	.68 .58	.16 .07
Graduates	.10 .32	.10 .10	.62 .00	.34 .07	.15 .58

5. The Washback of the EFL Benchmark upon the English Language Teaching and Assessment

The issue title resembles the former one; devaluing the test-oriented English language classroom instruction. Truly, owing to the forthcoming test pressure (i.e. CSEPT Level 2) on campus, English juniors beg for the test-centered classroom instruction, which results in the comparatively higher scores of response (i.e. Average Score = 3.14). In wake of the CSEPT pressure, both English juniors and graduates declined to be immersed in the ‘test-centered’ classroom instruction (Seniors: 2.82; Graduates: 2.90). To put it simply, only the EFL benchmark approaches do the surveyed English majors call for the test-based drills

Table 5 The Washback upon the English Language Teaching and Assessment

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.14	81	.54
Seniors	2.82	71	.57
Graduates	2.90	89	.60

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.32 .24	.09 .09	.00 .03	.09 .02	.55 .45
Seniors	.32 .08	.09 .09	.00 .67	.55 .31	.09 .14
Graduates	.24 .08	.09 .09	.03 .67	.45 .14	.02 .31

6. The Desire for More Test-Based Courses in Response to the EFL Benchmark

Despite the invited English majors (especially the English juniors) were less positive to the dominance of EFL benchmark to the overall test-centered language teaching and testing practices, they were modestly supportive of the test-based courses for the learning guidance. This is further envisioned in the average scores from the surveyed respondents, which falls around the score 3. Still, similar to the research results in the previous issue, the English juniors were comparatively in favor of the test-centered course management since they are loaded with the pressure of EFL benchmark.

Table 6 The Desire for Test-Based Courses

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.20	81	.52
Seniors	3.00	71	.45
Graduates	3.08	89	.61

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.21 .13	.09 .08	.06 .28	.00 .07	.42 .33

Seniors	.21 .08	.09 .08	.06 .65	.42 .29	.00 .13
Graduates	.13 .08	.08 .08	.28 .65	.33 .13	.07 .29

7. The Urgent Need for the Test-based Courses at the Expense of Other Courses

Based on the results in Table 6, the surveyed students' call for the test-based courses does not mean their willingness to sacrifice other academic courses. This presents the lower average scores from the respondents, who mostly disagree to place the test-centered courses as the foci. 'Keeping the normal track of EFL study' is the prerequisite for the advanced education of English study. Being skillful in the test performances is not tantamount to the expert manipulation of EFL language skills in the fluent language performances.

Table 7 The Urgent Need for the Test-based Courses

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.78	81	.74
Seniors	3.14	71	.49
Graduates	2.60	89	.90

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.36 .18	.12 .11	.01 .28	.66 .10	.06 .46
Seniors	.36 .55	.12 .12	.01 .00	.06 .25	.66 .84
Graduates	.18 .55	.11 .12	.28 .00	.46 .84	.10 .25

8. The Motivation to Apply for the Remedial Course ahead of the Formal CSEPT Test Administration due to the EFL Benchmark

As the Table 8 exhibits, most of the respondents were inclined to empower themselves with the test-taking skills through the remedial courses provided on campus. This statement is further evidenced in the lower response scores, being located around 2.90 (Table 8). This result yielded to the statistically insignificant differences ($t>.05$). Nevertheless, among the surveyed research groups, the English graduates were much more desirous of the remedial course program.

Table 8 The Motivation to Apply for the Remedial Course

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.78	81	.65
Seniors	2.66	71	.75
Graduates	2.93	89	.58

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.16 .15	.11 .11	.56 .31	.15 .40	.38 .09
Seniors	.12 .27	.11 .10	.56 .04	.38 .53	.15 .01
Graduates	.15 .27	.10 .10	.31 .04	.10 .01	.40 .53

9. The Management of Passing the EFL Benchmark by Taking the Less Challenging Formal English Proficiency Test

Among batteries of well-recognized formal English proficiency tests, the invited English majors supposedly met the requirement of the officially regulated EFL benchmark by taking the less challenging English proficiency tests, such as CSEPT Level 2. Virtually, the results in Table 9 ran counter to our assumption, in which the selected English majors were not wishy-washy to their selected English proficiency tests. That is, the current university English majors are desired to take challenges for the formally and globally recognized English proficiency tests (e.g. TOEIC, TOEFL) in wake of their qualification for the EFL benchmark. For the sampled English majors, invalid certificates are nothing to their career development (Juniors: 2.95; Graduates: 3.17). Nevertheless, the English majors expressed the slightly positive attitudes toward this issue since satisfying the regulated EFL benchmark was what counted most.

Table 9 The Management by Taking the Less Challenging Test

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.95	81	.69
Seniors	2.66	71	.61
Graduates	3.17	89	.66

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.29 .22	.11 .10	.03 .10	.03 .47	.55 .03
Seniors	.29 .51	.11 .10	.03 .00	.55 .76	.03 .25
Graduates	.22 .51	.10 .10	.10 .00	.03 .25	.47 .76

10. The Test Load of EFL Benchmark

The issue of test load varies with English majors' grade levels. Despite being pressured by the EFL benchmark, the English juniors were seemingly keeping their normal tracks of EFL study. This statement is warranted by the lower response scores, averaging below 3.0 (i.e. the standard score of agreement). Likewise, the English graduates expressed the similar attitudes, laying the test load as the secondary consideration in their academic study. By comparison, the invited English seniors slightly consent to the test load since the score values dominate their certificate procurement.

Table 10 The Test Load of EFL Benchmark

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.58	81	.77
Seniors	3.17	71	.61
Graduates	2.31	89	.73

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.59 .27	.12 .11	.00 .05	.87 .00	.30 .54
Seniors	.59 .85	.12 .11	.00 .00	.30 .58	.87 1.13
Graduates	.27 .85	.11 .11	.05 .00	.54 1.13	.00 .58

11. Students' Intention to Shy Away from the Oral Assessment

Currently, the EFL benchmark places the heavy weights upon the skills of listening and reading for assessment. The productive language proficiency (i.e. the speaking and writing abilities) failed to capture the tantamount attention in the selected English majors. This statement is supported through the numerical evidence in Table 7, in which most of the respondents argued against our earlier assumption, striving to maintain the balance among four English language skills (i.e. listening, speaking, reading, and writing). Accordingly, even though the selected English majors captured the satisfying certificate of TOEIC Listening and Reading Tests, they were believed to challenge for the productive tests to manifest their speaking and writing achievements.

Table 11 Students' Intention to Shy Away from the Oral Assessment

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.70	81	.75
Seniors	2.70	71	.68
Graduates	2.51	89	.68

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.00 .20	.11 .11	1.00 .19	.28 .07	.28 .46
Seniors	.00 .20	.11 .11	1.00 .21	.28 .08	.28 .47
Graduates	.20 .20	.11 .11	.19 .21	.46 .48	.07 .08

12. Students' Intention to Shy Away from the Writing Assessment

Currently, several of the widely-recognized tests lay the emphases on the receptive skills (i.e. listening and reading abilities) (e.g. CSEPT tests), and they are eligible for the measurement of regulated EFL benchmark. This possibly leads to English majors' avoidance in the formal English proficiency tests which measure the productive skills (e.g. the graded-level of GEPT). Virtually, through the statistical results in Table 12, most of the selected respondents argued against such a statement in view of the lower average scores (juniors: 2.62; seniors: 2.68; graduates: 2.42). The lower average scores among these three surveyed groups failed to yield the statistically significant values through the one-way ANOVA computation ($t>.05$). Yet, English seniors captured the relatively high average scores (2.68), approaching

the threshold of *Agreement*. This indicates a few of the sampled English juniors are less confident in their productive skills, especially in their written performances.

Table 12 Students' responses to avoid assessing their writing abilities

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.62	81	.72
Seniors	2.68	71	.69
Graduates	2.42	89	.60

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Junior Seniors	.06	.11	.86	.33	.21
Graduates	.20	.10	.15	.05	.45

13. The Correlation between the Statistical Values in Question 11 and Question 12

Table 13 targets the correlation of statistical values between question 11 (i.e. the oral assessment) and question 12 (i.e. the writing assessment). The purpose of this investigation is to detect whether or not the invited English majors screwed their preference in a particular productive language skill. Table 13 presents the results in the following:

Table 13 The statistical correlation between Q11 and Q12

		Q11	Q12
Q11	Pearson Correlation	1	.45
	Sig. (2-tailed)		.00
	Numbers	241	241
Q12	Pearson Correlation	.45	1
	Sig. (2-tailed)	.00	
	Numbers	241	241

Evidently, the values of correlation reach .45, locating on the rank of low-correlation (See Table 13). This indicates the sampled respondents indeed prefer one productive language skill. As we overhaul the average values in Table 11 and Table 12, the sampled English majors seemingly avoid the oral assessment in view of the higher average scores (approaching the lowest requirement of *Agreement*).

14. The Priority of American Pronunciation in Selected English Proficiency Tests

The higher upperclassmen incline to U.S.A diction test and bring about conspicuous divergence. Yet, the lower degree respondents point out that they are willing to understand British accent when they taking listening examination.

Table 14 English major's preference of American pronunciation in EFL tests

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.04	81	.58
Seniors	2.61	71	.64
Graduates	2.75	89	.82

(B)

Student	Mean	Standard Deviation	Conspicuousness	95% Confidence Interval	
				Minimum	Maximum
Juniors	.43 .28	.11 .11	.00 .03	.15 .02	.71 .55

As Table 14 manifests, English juniors consent to select formal EFL tests with the American pronunciation. Their positive responses yields the statistical significance in comparison with the other two research groups (i.e. seniors and graduates) ($t<.05$). Remarkably, with the shifting testing trends, the American pronunciation no longer dominates the test sections of listening and speaking. Such a belief of global English leads to the disagreement of English seniors and graduates regarding the superiority of American English in the formal EFL tests. Yet, it might be worrisome for English juniors since they relatively depend on the American English pronunciation, limiting their scopes to get immersed in other types of English pronunciation.

15. English Majors' Responses to Formal English Tests with More Text-Base Reading Assessment but Less Vocabulary and Grammar Tests

This issue is intended to investigate English majors' preferences of test sections measuring their general English proficiency levels. The results are summarized in Table 15 that follows.

Table 15 English majors' responses to formal English tests assessing more knowledge of language components rather than text-based reading proficiency

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.33	81	.71
Seniors	2.85	71	.47
Graduates	2.57	89	.67

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.51 .24	.10 .10	.00 .05	.76 .48	.26 .00

As a matter of fact, the lower grade respondents comparatively prefer more text-based reading comprehension tests, leading to the statistical significance with the English seniors (t: = .00<.05). Evidently, the English seniors prefer seeds of test items measuring the knowledge of components, implying their bottom-up reading processes (i.e. commencing from the study of vocabulary and grammar). Distinctively, the surveyed English juniors and English graduates prefer more percentage scores allotted to the text-based reading test sections, downgrading the values of language components for assessment. Pedagogically implied, both English juniors and English graduates employ the top-down textual reading processes, locating the study of language components at the secondary consideration.

16. English Majors' Responses to the Employment of Test-Taking Skills in Their Test Participation

This issue, through the results in Table 15, presents the increasing attention to lower grade test participants. Despite most of the respondents would not heavily manipulate the test-taking tips on their test performances, the English juniors relatively begged for the test-taking skills to succeed in their test performances on the grounds of higher average score values (juniors: 2.62; seniors : 2.61; graduates: 2.37). Even though the average scores among these three research groups fail to reach the statistical significance through the computation of one-way ANOVA, the respondents currently on campus express their positive attitudes toward the test-taking skills. Yet, the English graduates expressed the contrary statement, bolstering the general English language proficiency levels as the primary target.

Table 16 The Employment of Test-Taking Skills in The Test Participation

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.62	81	.72
Seniors	2.61	71	.75

Graduates	2.37	89	.86

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.11 .25	.13 .12	1.00 .12	.30 .05	.32 .54

17. English Majors' Attitudes toward the Transfer of Test Achievements to Their General English Language Behaviors

In speaking of the transferability from the test achievements to the general foreign language behaviors, both English juniors and graduates express the moderate attitudes, approaching the average values to the lowest threshold of Agreement (i.e. score = 3) (English juniors: 2.93; English graduates : 2.91). No statistically significant results produced between these two research groups, confirming the statement in the beginning of this section. Yet, the sampled English seniors are negative to the interlocking relationships between EFL test achievements and the foreign language behaviors. The response scores from English seniors (i.e. 2.66 average scores) indeed reach the statistical significance in comparison with their two research counterparts (t: junior-senior =.02<.05; senior-graduates =.03<.05). Evidently, the surveyed English seniors are inferred not to outperform their counterparts in the formal English test participation, leading to the statistical significance.

Table 17 Transfer of Test Achievements to English Language Behaviors

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.93	81	.59
Seniors	2.66	71	.63
Graduates	2.91	89	.54

(B)

Student	Mean	Standard Deviation	Conspicuousness	95% Confidence Interval	
				Minimum	Maximum
Juniors	.26 .02	.09 .09	.02 .98	.03 .20	.50 .24
Seniors	.26 .25	.09 .09	.02 .13	.50 .48	.03 .02
Graduates	.02 .25	.09 .09	.98 .03	.24 .02	.20 .48

18. The Advantages of English Teachers' Constant Support of EFL Test Participation

Depending on Table 18, English juniors captures scores of 3.06, reaching the rank of *Agreement*. Likewise, the sampled English graduates attained similar results, confirming the constant support of EFL teachers in their test participation (average scores = 2.97). These two surveyed respondents present the statistically insignificant results, implying their positive attitudes toward the advantages of EFL teachers' support. Contrarily, the invited English seniors only scored 2.73, leading to statistical significance in comparison with the other two research groups (t <.05). In other words, the EFL teachers' constant support seems to be invalid in English seniors' test participation.

Table 18 Teachers' Constant Support of EFL Test Participation

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.06	81	.48
Seniors	2.73	71	.68
Graduates	2.97	89	.57

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.33 .10	.09 .09	.00 .56	.10 .12	.56 .31
Seniors	.33 .23	.09 .09	.00 .04	.56 .46	.10 .00
Graduates	.10 .23	.09 .09	.56 .04	.31 .00	.12 .46

19. English Majors' Test Anxiety of the EFL Benchmark

Among these three surveyed research groups, the English seniors were anxious of the EFL benchmark, which was greatly affective to their graduation (Seniors: 2.93). That was the major reason why the English seniors outscored the remaining two research groups, and the disparities of the average scores led to the statistical significance through the one-way ANOVA computation (t= .00<.05). Evidently, English seniors were apprehensive of their formal English test performances.

Table 19 English Majors' Test Anxiety of the EFL Benchmark

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.49	81	.84

Seniors	2.93	71	.57
Graduates	2.40	89	.73

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.44 .09	.12 .11	.00 .72	.73 .19	.14 .36
Seniors	.44 .53	.12 .12	.00 .00	.14 .24	.73 .81
Graduates	.09 .53	.11 .12	.72 .00	.36 .81	.19 .24

20. The Whipping Effect of Rising EFL Benchmark to English Majors' EFL Study

Currently, the CEFR B1 was regulated as the lowest requirement for the selected English majors' general English language proficiency levels. Here, we turn to different lens in detecting this issue: the rising EFL benchmark levels to drive English majors' motivation to prepare the formal EFL tests. Such a whipping effect was surmised to efficiently bolster EFL learners' test achievements. Such a statement presented a slightly moderate response from the selected English majors. Despite being categorized in the Section of *Disagreement*, the average scores approached the lowest requirement of *Agreement* (i.e. the lowest score=3). Particularly, such a test policy change won the wide support from the sampled English graduates, stoking higher demands upon English majors' efficiency in their test performances.

Table 20 The Whipping Effect to English Majors' EFL Study

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	2.79	81	.70
Seniors	2.75	71	.82
Graduates	2.93	89	.62

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.04 .14	.12 .11	.93 .43	.24 .41	.33 .13
Seniors	.04 .19	.12 .11	.93 .26	.33 .47	.24 .09

Graduates	.14 .19	.11 .11	.43 .26	.13 .10	.41 .47

21. The Advantages of EFL Benchmark upon the Pursuit of Advanced Education the Application of Jobs

The EFL benchmark, through the results in Table 21, won the positive responses from both English juniors (average scores = 3.31) and English graduates (average scores = 3.04). Interestingly, the English seniors were not highly positive to the effects of EFL benchmark in their career management in view of lower average scores (2.76). Such a disparity in seniors' responses gave birth to the statistical significance in comparison with English juniors (t= .00<.05) and English graduates (t= .04, <.05). It was speculated that English seniors were not much clear of their future directions in their career development. Otherwise, the EFL benchmark was highly instrumental in the sampled English majors' (especially juniors' and graduates') management in their career or advanced education.

Table 21 The Advantages upon the Pursuit of Advanced Education the Application of Jobs

(A)

Students	Average Scores	Number	Standard Deviation
Juniors	3.31	81	.58
Seniors	2.76	71	.64
Graduates	3.04	89	.75

(B)

Student	Mean	Standard Deviation	Significance	95% Confidence Interval	
				Minimum	Maximum
Juniors	.55 .26	.11 .10	.00 .04	.28 .01	.82 .52
Seniors	.55 .28	.11 .11	.00 .03	.82 .55	.28 .02
Graduates	.26 .28	.10 .11	.04 .03	.52 .02	.01 .55

22. The Advantages of EFL Benchmark upon the Pursuit of Advanced Education the Application of Jobs

The EFL benchmark, through the results in Table 21, won the positive responses from both English juniors (average scores = 3.31) and English graduates (average scores = 3.04). Interestingly, the English seniors were not highly positive to the

effects of EFL benchmark in their career management in view of lower average scores (2.76). Such a disparity in seniors' responses gave birth to the statistical significance in comparison with English juniors (t= .00<.05) and English graduates (t= .04, <.05). It was speculated that English seniors were not much clear of their future directions in their career development. Otherwise, the EFL benchmark was highly instrumental in the sampled English majors' (especially juniors' and graduates') management in their career or advanced education.

Table 21 EFL Benchmark upon the Pursuit of Advanced Education the Application of Jobs

Students	Average Scores	Number	Standard Deviation
Juniors	3.31	81	.58
Seniors	2.76	71	.64
Graduates	3.04	89	.75

Multicomparison 21

Student	Mean	Standard Deviation	Significance	95% 信賴區間	
				Minimum	Maximum
Juniors	.55 .26	.11 .10	.00 .04	.28 .01	.82 .52
Seniors	.55 .28	.11 .11	.00 .03	.82 .55	.28 .02
Graduates	.26 .28	.10 .11	.04 .03	.52 .02	.01 .55

(II) The EFL Benchmark to the Competitive Edges of English Graduates in the Workforces

22. The Advantages of EFL Benchmark in English Graduates' Career Management

Among the invited English graduates (N= 89), only two major types of occupation (i.e. service and teaching) attract quite a number of devotees (Teaching = 12; Service = 36). Accordingly, this research section zooms in on these two major groups of occupation for the in-depth investigation. To begin with, regarding the EFL benchmark advantages in relation to the competitive edges, those working as English teachers or tutors present their positive attitudes, allocating the average scores (M=3.33) to the rank of *Agreement*. Comparatively, those who dedicate themselves in the service industries express the counter opinion in the effects of the EFL benchmark (M = 2.92). Such the disparities in the average scores yielded the

statistical significance through the computation of Independent T-test (See Table22 (B), t<.05). Evidently, the EFL benchmark seemingly narrows its wielding powers to the fields of education.

Table 22 The Advantages in English Graduates' Career Management

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	3.33	.49	.14
Service	36	2.92	.60	.10

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	.08	.78	-2.16	46	.04	-.42	.19	-.81	-.03
Equal Variances Not Assumed			-2.39	22.97	.03	-.42	.17	-.78	-.06

23. The Extent of English Majors' Support of the EFL Benchmark Regulation to Under-classmates

Despite the EFL benchmark wane its impacts upon the industries of service (See Table 21), these two surveyed groups are highly positive to the whipping effects of EFL benchmark to stimulate the EFL study of under-classmates in view of the high average scores in Table 23(teaching = 3.67； service = 3.39). Based on the Table 2, the graduates demonstrated that no matter what careers they in, they agree — almost highly agree — with the EFL Benchmark because of the influence of career workplace upon it (Teacher: 3.67; The others: 3.39). Such a minor gap in the score values fail to capture the statistical significance (t>.05) in Table 23 (B). Thus, the EFL benchmark was lauded as the catalyst to stimulate English majors' achievements in their foreign language study. Yet, such a whipping effect is not tantamount to its advantages of the career management.

Table 23 The Extent of the EFL Benchmark Regulation to Under-classmates

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching Service	12 36	3.67 3.39	.49 .49	.14 .08

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	.55	.46	-1.69	46	.10	-.28	.16	-.61	-.05
Equal Variances Not Assumed			-1.69	18.96	.11	-.28	.16	-.62	-.07

24. English Graduates' Attitudes toward Acquiring Professional Certifications before Graduation

The results in Table 24, contrary to our previous statement in Table 22, reflect the surveyed English graduates' expectation of obtaining the higher ranks of formal English proficiency tests (Teaching = 3.00; Service = 3.28). That is to say, the EFL benchmark does not fully guarantee English graduates to hunt for their ideal jobs. Yet, the EFL benchmark is treated as the prerequisite of EFL graduates' competitive edges in their career management despite the effects of EFL benchmark vary with types of occupation.

Table 24 Attitudes toward Acquiring Certifications before Graduation

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching Service	12 36	3.00 3.28	1.48 .66	.43 .11

(B)

	Levene	t-evaluation	
			95% Confidence

								Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	37.75	.00	.90	46	.37	.28	.31	-.34	.90
Equal Variances Not Assumed			.63	12.49	.54	.28	.44	-.68	1.23

25. The Extent of Agreement on Capturing Higher Ranks of EFL Benchmark at the Expense of Professional Course Participation

Although the EFL benchmark predominates the English majors' procurement of academic certificates, most of the sampled English graduates express their strong disagreement on the dedication to the EFL benchmark at the sacrifice of the professional course participation. This is further warranted by lower scores from two major types of respondents, ranging around 2 scores on average (Teaching = 1.00; Service = 2.06). Particularly, those working in the field of teaching voice their strong disagreement on this issue, arguing that the EFL benchmark is not the whole picture of the advanced EFL education.

Table 25 The Agreement at the Expense of Professional Course Participation

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	1.00	.00	.00
Service	36	2.06	.71	.12

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	14.29	.00	5.08	46	.00	1.06	.21	-.64	1.47
Equal Variances			8.86	35.00	.00	1.06	.12	-.81	1.30

Not Assumed									

26. The Extension of EFL Benchmark to Related Fields of Professional English Education

Here, we turn different lens in detecting the ranges of EFL benchmark, extending its scope to other related fields of expertise. Currently, the general English language proficiency tests are touted as the sole indicator of English majors' growth in their professionalism. Yet, such an evaluation is severely lamented for biasing the normal direction of advanced English education. Other certificates of professionalism merit the serious consideration, including the TKT and the nationally-administered test of interpretation by Center of Language Training and Testing (LTTC) in Taiwan. When being requested such an issue, those working in the field of teaching express their disagreement on posing arrays of formal tests upon English majors' regular study on campus (average = 2.33). By contrast, the ones in the industries of service are supportive of such a policy, enriching the expert knowledge of English majors through the catalyst of various formal professional assessment and evaluation. Such a wide range of average score disparities in these two surveyed groups gives birth to the statistical significance (t = .00 < .05). Interestingly, it is the teaching staff that demonstrates their negative attitudes toward multiple assessments and evaluation in keeping track of English majors' professional development.

Table 26 The Extension to Related Fields of Professional English Education

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching Service	12 36	2.33 3.28	.49 .66	.14 .11

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum

Equal Variances Assumed	1.37	.25	4.54	46	.00	.94	.21	.53	1.36
Equal Variances Not Assumed			5.26	25.25	.00	.94	.18	.57	1.31

27. The Response of English Graduates to Launch the EFL Benchmark Early to Proctor English Majors' General English Language Proficiency Development

Being consistent to the early statement in Table 23, the EFL benchmark lends itself to regulating English majors to map out their test-oriented preparation scheme, which wins the evidence from the higher average scores of the sampled research groups (Teachihg: 3.67; Service = 3.39). In other words, the teaching staff is mesmerized by the whipping effects of EFL benchmark to bolster English majors' general English language proficiency levels despite the surveyed English majors on campus were still wishy-washy about their formal English test preparation.

Table 27 The Response to Proctor English Majors' General English Language Proficiency Development

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	3.67	.49	.14
Service	36	3.39	.69	.11

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	3.38	.07	-1.29	46	.20	-.28	.22	-.71	.16
Equal Variances Not Assumed			-1.52	26.45	.14	-.28	.18	-.65	.10

28. English Graduates' Attitudes toward the Intermediate Level of EFL Benchmark to English Majors

For long, the intermediate-level of EFL benchmark has been severely attacked not to stimulate English majors' general English language proficiency. This echoes with the written reaction from the surveyed teaching staff, placing the lower average scores (2.67) to upgrade the EFL benchmark. It is speculated that these teaching staff is primarily responsible for the young-kid English education. Thus, the upper-intermediate level of EFL benchmark seems redundant. Distinctively, those working in the field of service voice their counter opinion on this issue, confirming the advantages of EFL benchmark in the career promotion. Even though the disparities of average scores fail to produce the statistical significance, the ones accountable for the business affairs relatively place the increasing emphasis on the higher ranks of EFL benchmark.

Table 28 Attitudes toward the Intermediate Level to English Majors

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	2.67	.49	.14
Service	36	3.03	.74	.12

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	.43	.51	1.58	46	.12	.36	.23	-.10	.82
Equal Variances Not Assumed			1.92	28.53	.06	.36	.19	-.02	.75

29. The Agreement of Alternative Assessment for the EFL Benchmark

Being in tune with Table 26, the multiple assessment of the EFL benchmark is strongly requested by these two surveyed groups in lieu of higher average scores (Teacher: 3.00; The others: 3.28). This recasts our further doubts about the dominating powers of paper-based English tests in English majors' academic achievements. Accordingly, the EFL benchmark is strongly called for to fly from the limited scope of paper-based evaluation.

Table 29 The Agreement of Alternative Assessment for the EFL Benchmark

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	3.00	.85	.26
Service	36	3.28	.66	.11

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	.70	.41	1.17	46	.25	.28	.24	-.20	.75
Equal Variances Not Assumed			1.03	15.63	.32	.28	.27	-.29	.85

30. The Attitudes toward the Widely-Recognized English Tests as the Main Indicator of English Majors' Test Achievements

For long, the widely-recognized formal English tests are highly instrumental to English graduates' career development. Yet, contrary to our expectation, the teaching staff in this research survey express the negative attitudes (M = 2.00), accepting the locally-administered English tests to validly keep track of the general English language development. Such a belief was further challenged by the ones in the industries of service based on the average scores (3.25), and the statistical significance yields. To put it alternatively, the local teaching staff devalues the test validity and reliability, requesting the cost-effective formal English tests as the priority (i.e. the test practicality).

Table 30 The Attitudes toward the Widely-Recognized English Tests

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	2.00	.00	.00
Service	36	3.25	.84	.14

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum
Equal Variances Assumed	51.75	.00	5.11	46	.00	1.25	.24	.76	1.74
Equal Variances Not Assumed			8.92	35.00	.00	1.25	.14	-.97	1.53

31. The Advantage of EFL Benchmark to Promote English Graduates' Performances in Their Workforces

Remarkably, through the driving force of the EFL benchmark, both selected research groups present the positive attitudes, locating the average scores to the rank of *Agreement*. Evidently, the EFL benchmark indeed promotes English majors' competitive edges in their foreign language behaviors as well as the professional development.

Table 31 The Advantage to Promote Performances in Their Workforces

(A)

Occupation	Numbers	Average Scores	Standard Deviation	Standard Error of the Average Scores
Teaching	12	3.00	.85	.25
Service	36	3.06	.71	.12

(B)

	Levene		t-evaluation						
								95% Confidence Interval	
	F	Sig.	t	Degree of Freedom	Sig. (2-tailed)	Mean Difference	Standard Deviation	Minimum	Maximum

Equal Variances Assumed	.78	.38	.22	46	.83	.06	.25	-.45	.56
Equal Variances Not Assumed			.20	16.47	.84	.06	.27	-.52	.63

V. Conclusion

This section primarily summarized the findings in this research paper. The EFL Benchmark is credited as the driving force to English majors' test preparation (Table 1, Table 20). Yet, except for the English seniors, quite a number of English juniors and graduates remained their normal track of EFL study, let alone their call for early test preparation, the test-oriented classroom, the test anxiety, and the remedial instruction to satisfy the regulated EFL benchmark or the (Table 2 to Table 8, Table 18). Empirically, most of the English majors would not highly transfer from their test achievements to their general English language behaviors (Table 17), which devalued the test-first classroom instruction or the benchmark. Despite being moderately loaded with the EFL benchmark (Table 10), the sampled English majors still preferred the widely-recognized EFL tests measuring their four language skills (i.e. TOEIC test) (Table 11 to Table 13; Table 15) on the consideration of the test practicality (e.g. the development of career or the pursuit of advanced education) (Table 9, Table 21). Particularly, being impacted by the TOEIC listening test design (i.e. the multiple accents in the conversation tests), most of the invited English majors (except for the English juniors) accepted the various accents ventured in the English listening tests (Table 14). Additionally, the test-taking skills were not the primary concern for the invited English majors (Table 16). What counted most is to get immersed in the normal English course participation.

With regards to the English graduates, particularly the ones in the fields of education and the industries of service, the EFL benchmark indeed moderately bolster their general English language proficiency (Table 22) and the competitive edges in the workforce (Table 23 & Table 24). Nevertheless, the English graduates declined to extend the impacts of EFL benchmark at the expense of other professional course management (e.g. the business English) (Table 25 & Table 26). The EFL benchmark is credited to act as the indicator not the dominator to keep track of English majors' learning progress (Table 27). If possible, the English graduates preferred alternative assessments for the EFL benchmark, not adopting the formal English proficiency test as the juggernaut to English majors' learning achievements (Table 29 to Table 31).

References

一、中文書目

周啟葶(2007)。**高中生英語自我效能、英語學習焦慮、英語學習策略與英語學習成就關係之研究**。國立台灣師範大學教育學系博士論文，未出版。

陳超明（主編）(2010 a)。S. Morris 著。**TOEIC 900 (I)**。台北市：聯經出版事業股份有限公司。

_____(2010,b) **TOEIC 900 (II)**。台北市：聯經出版事業股份有限公司。

陳智華(2009 年 11 月 29 日)。技職英檢班 陳超明：根本浪費錢。**聯合報**, AA4 版。

陳致遠、王文萱、鄭家文（譯）(2008)。金智淵、李寶英、李聖龍著。**New TOEIC 閱讀完全攻破 BASIC**。台北市：寂天出版社。

彭尊聖、謝雅婷（譯）(2008)。金秀賢、李聖熙、金妙熹著。**New TOEIC 聽力完全攻破 BASIC**。台北市：寂天出版社。

彭尊聖、賴瑞玲、謝雅婷（譯）(2010)。金智淵、李聖龍、崔伊玲、李恩奎著。**NewTOEIC 閱讀完全攻破 INTENSIVE**。台北市：寂天出版社。

彭尊聖、羅竹君、謝雅婷（譯）(2009)。李寶英、金真璟、金學仁、李美英著。**New TOEIC 聽力完全攻破 INTENSIVE**。台北市：寂天出版社。

孫蓉華(2010, 3 月 6 日)。填鴨式 落伍了學學有用英文吧。**聯合報**, AA4 版。

張一慈、劉中、賴翠玲 (2010)。**多益一路通**。台北市：文鶴出版有限公司。

藍孝蘋、徐毓秀（譯）(2007 a)。M. Broukel, B. Rogers 著。**多益致勝大全:聽力篇**。台北市：寂天出版社。

_____(2007 b)**多益致勝大全:閱讀篇**。台北市：寂天出版社。

二、英文書目

Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129.

Biggs, J. B. (1995). Assumptions underlying new approaches to educational assessment. *Curriculum Forum*, 4(2), 1-22.

_____ (1996). *Testing: to educate or to select? Education in Hong Kong at the cross-roads*. Hong Kong: Hong Kong Educational Publishing Co.

Chen, S. H. (陳姝樺) (2009). *The reality behind college students' motivation to learn English considering Taiwan's English education policy and recent developments in the new media environment* [從台灣英語教育政策以及新媒體環境發展探討大學生英語學習動機的現實]. Unpublished master's dissertation, Da-Yeh University, Changhua.

Chen, T. H. (陳姿惠) (2008). *Impacts of compulsory standardized examinations on college students' L2 learning motivation* [強制性標準化測驗對大學生外語學習動機之影響]. Unpublished master's dissertation, Chao Tung University, Hsinchu.

Cheng, Y. L. (鄭雅玲) (2009). *An investigation of the influence of a graduation requirement on English learning from both English student' and teachers' perspectives* [英語畢業門檻相關規定對大學生英語學習與英語教師教學影響之研究]. Unpublished master's dissertation, Leader University, Tainan.

Chou, C. H. (周承萱) (2010). *Perceptions of the test of English for international communication and its washback effects: A case study at a private university in Taiwan* [多益英語測驗意見調查以及回沖效應之研究]. Unpublished master's dissertation, Hsuan Chuang University, Hsinchu.

Chu, H. Y. (朱秀瑜) (2008). *Stakes, needs and washback: An investigation of the English benchmark policy for graduation and EFL education at two technological universities in Taiwan* [利害關係、需求與回沖效應:以台灣之大學英文畢業門檻為例]. Unpublished doctoral dissertation, National Taiwan Normal University, Taipei.

Evans, V. & Dooley, J. (2004). *Prepare and practice for the TOEIC test*. EU: Express Publishing.

Gardner, R. C. & Lambert, W. E. (1972). *Attitudes and motivation in second language learning*. Newbury House: Rowley, Mass.

Gwon, O. K. (2010). *Score high on listening and reading tests for TOEIC*. Taipei: Cosmos Culture Ltd.

Hamp-Lyons, L. (1997). Washback, impact, and validity: Ethical concerns. *Language Testing*, 14, 195-303.

Horwitz, E. K. (1991). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. In E. K. Horwitz, & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications*. (pp.37-40). Englewood Cliffs. N J: Prentice Hall.

Huang, L. H. (黃齡萱) (2010). *English graduation threshold among the college students' EFL anxiety, motivation and learning strategies* [英檢畢業門檻對大學生的英語焦慮、英語學習動機、英語學習策略之影響研究]. Unpublished master's dissertation, Tzu Chi University, Hualien.

Lee, C. H. (李俊賢) (2011). *The study of technological university students' self-attributions for success and failure in English learning* [科技大學學生英語學習成敗歸因之研究]. Unpublished master's dissertation, National Kaohsiung First University of Science and Technology, Kaohsiung.

Lin, C. I. (林靜憶) (2009). *A study of the implementation of English education policy*

at universities of technology in Taiwan [科技大學英語教育政策執行之研究]. Unpublished doctoral dissertation, National Pingtung University of Education, Pingtung.

Lin, L. (2006). *Prepare series for the New TOEIC test: introductory course*. NY: Pearson Education Inc.

_____(2007a). *Prepare series for the New TOEIC test: intermediate course*. NY: Pearson Education Inc.

_____(2007b). *Prepare series for the New TOEIC test: advance course*. NY: Pearson Education Inc.

Mousavi, S. A. (2002). *An encyclopedic dictionary of language testing*. Taipei: Tung Hua Book Co. Ltd.

Pearson, I. (1988). Tests as levers for change (or 'putting first things first). In D. Chamberlain & R. J. Baumgardner (Eds.), *ESP in the classroom: practice and evaluation* (pp. 98-107). London and Oxford: The British Council in Association with Modern English Publications (Macmillan), 1988.

Popham, W. J. (1983). Measurement as an instructional catalyst. In R. B. Ekstrom (Ed.), *New directions for testing and measurement: Measurement, technology, and individuality in education* (pp. 19-30). San Francisco: Jossee-Bass.

_____ (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.

Thompson, C. (2007). *The best TOEIC*. Seoul: Moonjinmedia Co., Ltd.

Tung, H. C. (2006, April 22). The influences of English language proficiency tests. 2006 International Conference on English Instruction and Assessment. Symposium conducted at National Chung ChengUniversity, Chiayi, Taiwan.

Tung, H. C. (2009, May 5). The investigation on CHNA English majors' TOEIC test performances:2007-2008. The Proceedings of 2009 CHNA International Conference on English Teaching in the R.O.C.(p.141-161). Symposium conducted at Chia Nan University of Pharmacy & Science, Tainan, Taiwan.

Tung, H. C. (2010, October, 15). The Earlier, The Better? The effects of test-intensive instructional program on English Majors' test achievement. Paper presented at 2010 International Conference on English/Foreign Language Education at Min-Hwei College of Health Care Management. Symposium conducted at Min-Hwei College of Health Care Management, Tainan, Taiwan

Appendix The Written Questionnaire for This Research Paper

英語檢定門檻相關調查問卷

親愛的學長姐以及同學，您們好：
誠摯地邀請您填寫這分問卷，本研究主要想了解目前各位對於英語檢定作為畢業門檻的相關反應，您提供的資料將對本研究有很大的參考價值，請您依照您的真實情況作答，非常謝謝您的參與！

應用外語系三年級甲班　林佳蓁 敬上　民國103年9月11日

第一部分: 基本資料

1. 性別：　□男性　　□女性

2. 在校生目前（103 年度本學期）就讀年級：

　□大學三年級　　□大學四年級

3. 畢業學長姐的學歷（**在校**生請**勿**勾**選**）：

　□大學/大專　　□ 研究所

4. 畢業學長姐目前的職業：

　□研究所學生　□服務業　□製造業　□農林漁牧業　□軍公教

　□科技資訊業　□金融業　□傳播媒體業　□家管

　□其它行業(請簡述):_____

第二部分:「英語檢定門檻」相關問題（所有參與者均必須勾選）

- 請依據您目前的情況確實回答問題並打ˇ。

問卷題目	非常同意	同意	不同意	非常不同意
1. 系上「英語檢定畢業門檻」幫助我提高英語檢定考試成績。				
2. 為了要通過「英語檢定畢業門檻」，我提早（如大二）開始努力準備英語檢定考試。				
3. 為了要通過「英語檢定畢業門檻」，我投注許多時間準備。甚至影響到我平日其它修課的課業也沒關係。				
4. 為了要通過「英語檢定畢業門檻」，我平日的英語聽說讀寫訓練朝英檢考試題型方向努力（如多益模擬試題)。比較不會拿一般的課堂聽說讀寫書籍（如 Listen in）來加強準備。				
5. 為了要通過「英語檢定畢業門檻」，我希望一般聽說讀寫課程的老師在教學或甚至小考，都能朝英檢（如多益）考 試題型來進行。				

6. 既然「英語檢定畢業門檻」如此重要，我希望系上增加與英檢相關的必選修課程（如「多益或全民英檢實務課程」等）。				
7. 如果系上有安排英檢相關選修課程，而該課程與我喜愛的其它必修或選修課程（如「基礎日文」）修課時段有牴觸時，我以選英檢相關課程為優先。				
8. 因為系上「英語檢定門檻」，我才會參加短期英檢輔導班級。				
9. 若「英語檢定畢業門檻」沒有特別規定要參加哪一類別的英檢考試（如多益或全民英檢等），我會先選擇較容易通過的英檢考試來應考。待有信心後，再投入較有挑戰的英檢考試準備。				
10. 「英語檢定畢業門檻」規定，讓我應考時很有壓力，甚至臨場表現不佳。				
11. 我選英檢考試時，會避開有口說評量考試（如劍橋英檢）				
12. 我選英檢考試時，會避開有寫作評量考試（如劍橋英檢）				
13. 我選英檢聽力考試時，希望以美國英語的發音為主。				
14. 我選英檢閱讀考試時，希望能少些單字或文法題，而多些文章閱讀題型（如文章克漏字或是文章閱讀考試等）				
15. 我在英語檢定考試當中，很倚賴課堂教師所提到的解題技巧（如音相近的選項不要選）來作答。				
16. 我相信英檢考試有通過者（如英文寫作考試），在一般英文課程（如英文寫作作業）也會有傑出表現。				
17. 導師或英檢課程老師督促，有助於提升我英檢考試成績。				
18. 系上「英語檢定畢業門檻」規定，讓我對英語學習感到十分焦慮不安。				
19. 目前系上「英語檢定畢業門檻」規定以通過中級英檢程度為主。但我認為應該提升到中高級英檢程度才有鞭策學習的效果。				
20. 我相信「英語檢定畢業門檻」規定，對我未來的工作應徵或升學(研究所)有幫助。				

第三部分:「英語檢定門檻」對工作效益度問題（限畢業學長姐勾選）

- 請依據您目前的情況確實回答問題並打ˇ。

問卷題目	非常同意	同意	不同意	非常不同意
1. 因「英語檢定畢業門檻」而考取的英檢證照，對我當時應徵工作或是升學(研究所)有幫助。				
2. 經過職場工作歷練後，我肯定系上的確要對學弟妹設定「英語檢定畢業門檻」，激勵他們考取英檢考試證照。				
3. 我認為就讀外語系，最重要的就是在畢業前一定要拿取好的				

英語檢定證照成績。				
4. 我認為如果為了拿取好的英語檢定證照成績，而使自己一般專業課程成績差一點是沒關係的。				