

WASHBACK EFFECT OF ALCPT ON CADETS IN TAIWAN: A CASE STUDY

Chiung-Wen Chang

Department of Foreign Languages, R.O.C. Military Academy, Taiwan
E-Mail: chiungwenchang@gmail.com

Abstract. *This paper presents a study investigating the washback effect of the American Language Course Placement Test (ALCPT) on cadets in a Taiwanese military academy. Over the last decades, washback studies have become even more important in the field of applied linguistics due to the critical role of the high stakes tests for students, teachers, and societies. This form of specific purposes assessment ALCPT affects both military personnel and civilians working within the military structure in terms of posting, promotion and remuneration, and it could be argued that it has serious consequences if not carried out professionally. Based on Hughes (1993) and Bailey's (1996) washback models and empirical washback studies, this study focused on cadet participants' attitudes toward aspects of the ALCPT and their corresponding behavior influenced by the test in their learning process. The data in this study were obtained through questionnaires and follow-up interviews conducted with cadets. A total of 140 cadets participated in this study on a volunteer basis, of whom ten agreed to take part in the follow-up interviews. The results revealed that ALCPT exerts a negative impact on cadets. The results also show that the exam not only has a micro level effect, which can be defined as the effect on an individual or a small group of individuals, but also a macro effect on a relatively large population of individuals studying or working in military academy contexts. Practical implications as well as suggestions for the instruction of cadets' language learning are offered to conclude the paper.*

Key words: *Washback effect, ALCPT, cadet's English learning*

1. INTRODUCTION

Taiwan has relied heavily on the United States (US) for military arms and supplies. Most of those arms need ongoing operational support and training from the US military. The ALCPT was designed by the US Defense Language Institute in the 1950s and was employed by the US forces as the benchmark EFL test. The most current statistical reference indicated that very few cadets reached the score of 70, the minimum required score for training in the US.

Since Republic of China Military Academy (ROCMA) is the main source of future Taiwan military officers, it needs an innovative instruction method to help cadets to improve their ALCPT scores, to achieve a much higher success rate in language learning. "Definitions of washback are nearly as numerous as the people who wrote about it." (Bailey, 1999:3) Researchers have adopted different perspectives as to its definition and scope according to their specific research context and target population. In view of this, the term *washback* is used in this study to indicate the influence of the ALCPT on cadets' learning in Taiwan's educational context. Although numerous studies have been conducted

regarding the washback effect of the English language tests in different contexts (e.g., Mizutani, 2009; Watanabe, 2004; Ferman, 2004; Chen, 2002; Cheng, 1998; Wall and Alderson, 1993), to date no empirical research on the washback effect of ESP tests within the Taiwan EFL context has been reported on. Moreover, students' voices have frequently been neglected in past research (Jiang, 2007; Tsagari, 2007), as is the case in the general literature on washback. As Hamp-Lyons (1997, p. 299) comments, many more studies are needed of students' views and their accounts of the effects on their lives of test preparation, test-taking and the scores they have received on tests. For these reasons, the present study attempts to investigate the washback effect of the ALCPT on cadets, for the purpose of examining how a high-stakes test influences cadets' learning in a Taiwanese educational context.

2. LITERATURE REVIEW

2.1 Definition of washback in this study

Numerous explanations of the term 'washback' can be found throughout the published research and literature on language testing. The term can be defined according to two major perspectives: one from a narrower view within the classroom at a micro level, and the other from a wider and more holistic view beyond the classroom at a macro level. As suggested by Bachman and Palmer (1996), washback, at a macro level, refers to the extent to which a test influences society, ranging from government policymaking, school administration, publishing, and general opportunities, to parents' expectations of their children. At a micro level, washback refers to the extent to which a test influences the classroom, mainly in the change or innovation of curricula and teachers' methodologies and the influence of students' learning. Bailey (1996) used the phrase "washback to the learners" to indicate the effects of tests on students, and "washback to the program" to indicate effects of tests on teachers, administrators, curriculum developers, counselors, etc.

To summarize, the narrower definition of washback focuses on the effects that a test has on teaching and learning. The wider or more holistic view of washback (also defined as test impact) looks beyond the classroom to the educational systems and society at large. All in all, tests can have "significant impact not only on individuals but also on practices and policies—in the classroom, the school, the educational system and in society as a whole" (Wall 2005).

In this study, a narrow interpretation of washback will be adopted: washback at a micro level within the classroom to investigate test washback on cadets, that is, the washback effect on learning.

2.2 Washback models

A detailed overview of the washback effect is provided by Alderson and Wall (1993), whose fifteen hypotheses incorporate a variety of possible aspects of washback, including the impact on what to teach/learn, how to teach/learn, the rate and sequence of teaching/learning, the degree and depth of teaching/learning and the attitudes towards teaching/learning. Alderson and Wall's model (1993) includes a number of washback hypotheses, and Hughes's (1993) model distinguishes between participant, process and product. Drawing on Hughes's (1993) model, Bailey (1996) developed new model (Figure

1) to show how washback might operate. This model offers a sound explanation of how washback might take effect: a test first influences the perceptions and attitudes of the participants towards their tasks, and these perceptions and attitudes in turn influence what processes the participants will take to achieve their intended products. Third, it indicates not only the effects of testing upon *participants*, *processes* and *products* but also the interaction between all three and their reaction toward the test. It emphasizes “participants’ perceptions and attitudes and how these factors affect what they do” (Bailey, 1999:10). These aspects, therefore, are the research foci of this study. Given my major interest in test-takers, my study design will be based on Hughes (1993) and Bailey’s (1996) model, and particular attention is paid to washback on cadets.

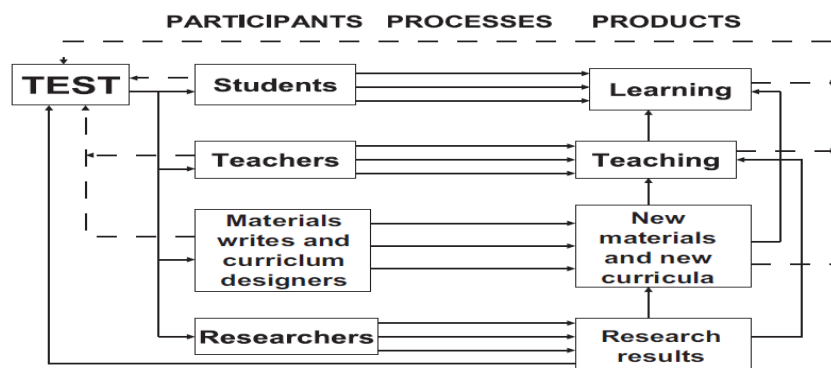


Fig. 1 A basic model of washback (Bailey, 1996, p. 264)

2.3 Empirical studies exploring washback on student learning

Safi (2006) aimed to examine how a high stakes performance test would influence the teaching activities and learning outcomes of an ITA (International Teaching Assistant) program. The results implied that the ITA test had some influence on teaching content, teaching methodology, and student learning. Students in the experimental groups outperformed those in the control group. The findings suggested a positive relationship between preparation for the (ITA) test and learning outcomes.

In Shohamy et al. (1996) teachers reported that the low stakes Arabic exam may have promoted learning at lower levels but not at upper levels as the students were committed to learning the subject anyway by that stage. In relation to an oral EFL exam, they believed that it had undoubtedly brought a focus on oral proficiency but the Reading component had not affected reading in class, as this part of the exam was considered to be poorly designed.

Andrews et al. (2002) aimed to see how introducing an oral component to a high stakes test (Hong Kong Advanced Supplementary ‘Use of English’) would impact the performance of those who take the test. They conducted simulated oral tests with three groups of candidates, matched for their ability over a three year period. The results showed that the nature of the washback varied across the student groups: only a small improvement in performance between the first and the third group was indicated, leading researchers to conclude that the washback effect of the test was delayed since it was more noticeable in the second than in the first year of the test. The researchers linked the impact of the test to the mediation of the exam-related published materials.

Another study that attempted to measure learning outcomes is that of Read and Hayes (2003). The researchers had two small groups of students (total 17) take retired versions of the International English Language Testing System (IELTS) exam as pre- and post-tests to two IELTS courses (intensive and general). The results did not show any significant improvement overall (with the exception of the listening tests at one of the two schools investigated), nor between the groups of students. As in previous studies, the researchers concluded that time is needed for washback to occur:

It is generally recognized that students need an intensive and usually extended period of study to achieve any substantial increase in their score on a proficiency test like IELTS (ibid: 110)

Chu (2009) asked 667 students from two Taiwanese universities about their perceived impact of the General English Proficiency Test (GEPT) graduation exam on their out-of-school practice. The top three learning activities according to a 5-point scale used to prepare for the GEPT test were “listening to English songs” ($M=2.33$), “memorizing vocabulary items” ($M=1.57$) and “watching English TV programs” ($M=1.53$). It is interesting that the students preferred to use entertaining audio-visual materials for test-preparation. However, as indicated by the author, “the low mean scores in the parentheses indicate that students did not do these tasks often and regularly” (p. 122). In other words, the GEPT graduation requirements had a fairly low degree of washback on learning behavior.

The findings of washback studies focusing on student learning are disparate and too mixed to provide a definite conclusion. In fact, of the washback studies reviewed only one documents any demonstrable gains in student learning that can be tied to the use of a test (Saif, 2006). The remaining studies have either used dubious approaches, did not find any considerable gains, or found negative results. It is not always clear what was meant by ‘*student learning*’. Some researchers looked at actual student performances, others at students’ perceptions (of their motivation or learning strategies), and others at the comprehensiveness of skills tested in the exam. To conclude, investigating washback on ‘the product of learning’ is still an area in need of research (stressed also in Wall, 2000; Alderson & Banerjee, 2001).

2.4 Empirical studies exploring washback on students’ feelings and attitudes

The purpose of this section is to consider whether and how the learners’ perspective is described in research studies in terms of attitudes, beliefs, learning strategies, self-esteem, anxiety and motivation towards high-stakes exams, commonly known in the literature as ‘*washback to the learner*’ (Bailey, 1996).

Li (1990) reported on students’ (and teachers’) positive attitudes towards an exam and motivation to study. The researcher explained that there seemed to be a new enthusiasm for learning English outside the classroom on the part of students, which included more after-class learning and led to high sales of simplified English readers. However, it is not clear whether what Li reported is based on the teachers’ opinions about their students’ attitudes only or whether she had direct access to the students’ views. Read and Hayes (2003) also reported, via a students’ questionnaire, positive feelings about the IELTS exam and motivation among learners (and teachers).

Cheng (1998) investigated the effect of the Hong Kong Certificate of Education Examination (HKCEE) on the students’ perceptions and attitudes toward their English

learning. The results showed that they had mixed feelings towards the HKCEE exam, recognizing on the one hand that the exam made them work hard to achieve good scores, while at the same time considering that they were not an accurate reflection of all aspects of their study. There might be individual differences among students in the way they perceive and react to exams.

Ferman (2004) examined the washback effects of the EFL oral test. The researcher concluded that the test resulted in differential washback among learners. In particular, she found that average ability level students were significantly different from other students: their anxiety level was the highest and they were most adversely affected by potential failure in the test. Therefore, the researcher suggested that to ensure the desired washback, individual differences among students need to be taken into account.

Gosa (2004) sought to identify possible washback effects that took place inside and outside classrooms as experienced by her Romanian students using student diaries. The analysis revealed that the personal environment of the students was affected by test washback *'to a greater extent than their classroom one, not only regarding the tasks practised but in almost every respect'* (ibid: 226). She concluded that students' expectations, feelings, attitudes, perceptions, beliefs, learning styles, and anxiety should be taken into account when trying to promote positive washback, as they are likely to interact with the test and, therefore, intervene in the washback process.

Watanabe (2001) interviewed university students about their test preparation to examine if the Japanese university entrance examination would motivate students. The results showed that washback to the learners was far from uniform, although those exams which students perceived to be most important for their future university career had greater impact than those which were perceived to be less important. On the other hand, those tests which were perceived to be less important induced less impact on the learner.

Also Yıldırım (2010) conducted a research study in Turkey on a similar high-stakes test, English Component of the Foreign Language University Entrance Exam (ECFLUEE), offered to secondary school graduates wanting to enroll in an English language program (Linguistics, English/American literature or English language teaching). He found that the exam exerts not only a negative impact but also results in an orientation problem in pre-service ELT programs due to a lack of English communicative skills. ECFLUEE is just another version of ILE in that both offer multiple choice questions on syntax, lexis and some sub-skills of reading.

Shih (2007) has explored the effects of a graduation exam on 29 Taiwanese university students. Although 58% ($n=17$) of the participants considered the test to be important, few actually spent time preparing for it. This echoes Bright and Von Randow's (2004) point: merely knowing that a test is important is not enough to guarantee that most people will actually study harder for it. Shih speculated that extrinsic factors (e.g., personal factors such as students' part-time jobs), intrinsic factors (e.g., students' learning attitudes), and test factors (the immediate importance of the test; the way the test-driven policy was implemented) were responsible for the phenomena perceived.

Even though there are a number of studies acknowledging the importance of the learners' views in the study of washback, their number is still fairly limited and the findings are contradictory, indicating a complex relationship between examinations and students' perspectives. The effects of factors such as motivation, language proficiency, previous education, current context of education, and time are also worth taking into consideration when investigating washback to the learners. The stakes of the test, the status of the language

taught, and the difficulty and importance of the test as perceived by the students are important variables, too. Furthermore, studies have considered various affective and cognitive factors such as students' attitudes, study habits, views, anxiety, personal environment, perceptions, beliefs, learning styles, motivation, etc. with conflicting results. This could be one of the reasons why investigating washback on learners is so complex. More evidence-based support for claims of positive and negative washback on students is needed that would help resolve the conflicting results.

3. RESEARCH METHODOLOGY

This study employs survey research, questionnaires and interviews, to collect data. The following two research questions were explored in this study to examine the nature of the washback effect of the ALCPT on cadets:

- 1) What is the nature of the washback effect on cadets' attitudes toward the ALCPT?
- 2) What is the nature of the washback effect on cadets' learning behaviors as a result of the ALCPT?

3.1 Participants

There are two types of participants in this study: the cadets who completed the questionnaires, and the cadets who agreed to be interviewed. The questionnaire respondents involved in this study are 140 cadets. They had learned English as a subject in school settings for at least eight years, and their English proficiency ranges from low intermediate to intermediate. They had neither studied abroad before, nor currently attended additional English classes outside of the classroom. In all, 85 students (61%) had the experience of taking language proficiency test, while 55 (39%) had never taken any language proficiency test before. A total of 10 cadets who indicated a willingness to be interviewed on their surveys were selected for the interview phase of this study. Their demographic information is summarized in Table 1.

Table 1 Demographic information about the participants

Category	Questionnaires Respondents (n=140)	Interviewees (n=10)
Major		
Information management (IM)	28	2
Management Science (MS)	31	2
Civil Engineering (CE)	28	3
Chemistry (C)	23	3
Mechanical Engineering (ME)	30	0
Experience of Language Proficiency Test		
GEPT	82	6
TOEIC	2	0
TOEFL	1	0
None	55	4

3.2 Instruments

Questionnaire: The aim of the study was to examine the nature of the washback effect of the ALCPT on cadets. The questionnaire consisted of three parts using a slightly modified questionnaire described in the reference (Wang, 2010). The first part related to the cadets' demographic information. All the items in this part were designed on a multiple-choice basis. The second part dealt with the cadets' attitudes toward the ALCPT related to ten aspects, including its difficulty, time allotment, test format, reliability and so on. The last part was about their behaviors in the learning process from two perspectives: learning contents such as the quality and quantity of learning materials and learning methods which include both the coaching methods and the recommended methods. Items in the second and third parts were designed on a five-point Likert scale of agreement, where 5 = strongly agree, 4 = agree, 3 = undecided, 2 = disagree, and 1 = strongly disagree. The five-point Likert scale was employed as it is one of the most commonly accepted scales in the education field (Green, 2007; Cheng, 2004; Gu et al, 1995; Cohen & Manion 1989). The researcher translated the questionnaire into the cadets' L1 (Chinese), checking the translation with a Chinese teacher. The author then piloted the questionnaire with three cadets and amended it, revising several items for clarity. The cadets were assured of anonymity and informed that their participation was entirely voluntary. Cronbach's alpha for the reliability of the questionnaire used in this study was 0.78.

Interview: The interviews were conducted after the questionnaires were collected. The aim was to gather qualitative data to back up and assist the interpretation of the quantitative data. After the participants were informed about the content and objectives of the study, the interviews were conducted with one participant at a time, each lasting between 20 and 30 minutes. All interviews were audio taped and transcribed verbatim. The interview sessions were completed in Chinese over a period of three weeks in November 2013. Five open-ended questions aimed to identify the possible washback effects that ALCPT may exert, and the way it shapes the language learning experiences of the cadets.

3.3 Data analysis

The quantitative data were analyzed using the Statistical Package for Social Sciences (SPSS version 12.0). Frequencies and means were used to analyze single items. The data collected through semi-structured interviews were coded through constant comparative analysis (Glaser & Strauss, 1967, p. 101-116), which refers to a process of repeated sifting through the data to distinguish similarities and patterns of reference in the interview transcripts. Analyses of these similarities and patterns gradually led to an evolving coding system for the categories. This repetitive process continued until coherence and a saturation of understanding the data was accomplished. The researcher used peer review as a technique to verify her coding of the data. The results of both codings indicated 85.6% reliability. The total analysis period took six weeks.

4. RESULTS AND DISCUSSION

Quantitative data

This section aims to present the statistical results and the data analyses in order to answer the research questions addressed in the present study.

4.1 Students' attitudes toward ALCPT

Students' attitudes were operationally defined as students' apprehensions and understandings of aspects of ALCPT. This part consisted of one category of 9 items. The cadets were asked to state to what extent they agreed with the following 9 statements. The results of this part are reported in Table 2

Table 2 Cadets' Attitudes toward the ALCPT according to the Mean Scores

Statement	Mean
ALCPT should NOT be used as the sole measure of cadets' language ability.	4.43
Cadets' learning is improved by practicing mock exam papers of ALCPT.	3.96
Most of the cadets work hard to achieve their best in the ALCPT	3.75
ALCPT forces cadets to study harder	3.50
ALCPT is one of the motivations for cadets' learning	3.43
Most of the cadets dislike the ALCPT	3.30
Taking ALCPT is a valuable learning experience	3.20
Mock examinations are important ways to learn	3.04
A cadet's score on the ALCPT is a good indication of how well s/he has learned the material.	2.13

Table 2 lists the statements in descending order according to the mean scores given by the cadets. Seven out of the nine items got the mean scores above 3 indicating that the cadets, at least to some extent agreed with the items. However, when they were asked to state their viewpoints on Item 6 (A cadet's score on the ACLPT is a good indication of how well s/he has learned the material), the majority of them disagreed with this belief. As can be seen, the item 'ACLPT should NOT be used as the sole measure of cadets' language ability' produced the strongest agreement among the surveyed cadets. This was followed by the item 'Cadets' learning is improved by practicing mock exam papers of the ALCPT'. This reflected the cadets' mixed feelings about the examinations. On the one hand, they did not look at the examinations as an accurate reflection of all aspects of their study, while on the other, they considered mock exam papers to be an important lever for learning. These mixed feelings towards the examinations are also demonstrated by the following two items: 'Most of the cadets dislike the ALCPT' but 'ALCPT is one of the motivations for cadets' learning.'

What is more, the results show that the cadets generally claim that the ALCPT fails to measure their true language ability. Nevertheless, this can not be regarded as a sign of negative washback effect. In the researcher's personal communication with cadets during the pilot study stage, they argued that they could perform better if the test format was different. It seems that the test looks invalid to them. Thus, it is hard to summarize this as either a positive or negative washback effect of the test. But this result does pinpoint one problem. It is argued by some researchers that the concept of face validity should be abandoned (Zou, 2005). The results from this study seem to oppose this opinion. A good test is one that appears to measure the ability it claims to measure, as judged by an untrained observer. The results of the current study indicate that the ALCPT lacks face validity. Accordingly, face validity should be taken into consideration in future ALCPT development and validation processes. Although the term 'face validity' has been abandoned,

its meaning can be preserved and be merged into the concept of content validity, since the “test takers will perceive tests as more interactive and authentic, and will therefore be more motivated, which could lead to enhanced preparation and hence to better performance” (Bailey, 1999:14).

4.2 Students' learning behaviors

Students' behaviors were operationally defined as students' actions influenced by the ALCPT in their learning process from two perspectives: learning contents and learning methods, based on the suggestions of Bailey (1996), Gu (2005), Qi (2004), the ALCPT specification and the fifteen washback hypotheses (Alderson & Wall, 1993). Thirteen items were used in this part. Three subcategories—learning contents, coaching methods and recommended learning methods—were grouped together for the convenience of analysis, but are reported separately in the following sections.

4.2.1 Learning contents

Students' feedback regarding learning contents influenced by the ALCPT were examined according to five aspects, the results of which are reported in Table 3.

Using authentic materials and increasing the quantity and variety of the listening materials are preferred learning behaviors according to cadets' responses. Therefore, they are signs of positive washback effects of the ALCPT. As for using mock tests, simple matters of positive or negative washback effects cannot be drawn without analyzing the quality of the mock tests and the degrees and intensities of students' uses. This study mainly examined the nature and scope of the ALCPT's washback effect; therefore, the analysis of the above points is beyond the scope of this study, but it is strongly recommended as the topic of further research.

Table 3 Cadets' learning behaviors—learning contents

Behavior (learning contents)	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree	Mean
1.Using textbooks	4.3	47.1	37.1	9.3	1.4	2.44
2.Using authentic material	15.7	51.4	28.6	3.6	0.7	2.78
3.Using mock tests	5.0	57.9	27.9	8.6	0.7	2.58
4.Increasing quantity	5.7	38.6	42.9	10.7	2.1	2.35
5.Increasing variety	4.3	38.6	44.3	11.4	1.4	2.33

4.2.2 Coaching methods

Three aspects of students' coaching methods influenced by the ALCPT were examined in this study. Table 4 shows the results related to these three aspects.

It is shown in the table that the cadets generally do adopt those coaching methods, and these are the negative influence the listening subtest has produced on their learning. This paper defines washback as the influence of testing on learning, at a micro level. Thus, the negative influence can be regarded as a negative washback effect. However, those cadets' coaching methods can never be caused by a test itself. There should be other factors, together with the test, that affect what students actually do.

Table 4 Cadets' learning behaviors—coaching methods

Behaviors (coaching methods)	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree	Mean
Applying test-taking strategies	9.3	42.9	32.1	12.9	2.9	2.43
Demanding test-preparation classes	4.3	29.3	47.9	13.6	5.0	2.14
Learning what occurs in the test first	7.1	45.7	42.9	2.1	1.4	2.55

4.2.3 Recommended learning methods

Students' recommended learning methods influenced by the ALCPT were examined according to five aspects. Table 5 reports the descriptive results of those items. All five recommended learning methods are the preferred learning methods indicated by ALCPT. As a result, they are signs of positive washback effects in this study. Accordingly, teaching and learning of those aspects should be integrated into the teaching curriculum for listening.

Table 5 Students' learning behaviors—recommended learning methods

Behaviors (Recommended learning methods)	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree	Mean
Main idea and important details	5.7	56.4	32.9	5.0	0	2.63
Speaker's attitude, opinions etc.	6.4	47.1	36.4	9.3	0.7	2.49
Implied meaning	6.4	46.4	43.6	2.9	0.7	2.55
Phonological features	4.3	53.6	36.4	4.3	1.4	2.55
Logical relationships	8.6	49.3	38.6	2.1	1.4	2.61

Qualitative Data

This section is organized in terms of the research objectives. The first objective was to identify the possible washback effects that ALCPT may exert, and the second one was the way in which ALCPT shapes the language learning experiences of the cadets.

4.3 Research objective 1: Washback effects of ALCPT

The data that relate to the washback effect of ALCPT are quite clear. All the participants took the same direction and talked about the same issue for the first question "Can you share your opinions about the effectiveness of the ALCPT exam? For instance, does it have any negative/positive/neutral effect on you?", second question "Is ALCPT an effective exam that evaluate academic English skills and knowledge?" and for the fourth question "Do you believe that preparing for ALCPT helps you develop your English?" Some of the responses of the participants are as follows:

"ALCPT is not a good exam because it only tests reading and listening. It also affects me rather negatively. All I have to do is to memorize some vocabulary lists and exam tips. This is annoying."

"ALCPT cannot be an effective exam. I know many people who got high scores in ALCPT but cannot even write a complete sentence in English."

"Well, I am quite experienced about such exams. All I need to do is to memorize more vocabulary items and study more grammar. ALCPT therefore has a negative impact not only on me but also on many people: my teachers, my classmates and even on the system because we spend our precious time here."

"In order to pass ALCPT, I went to an online platform to practice on-line mock tests"

Similar results were also indicated in the studies in which the negative influence of high-stakes tests were discussed (Alderson & Hamp-Lyons, 1996; Cheng et al., 2004). The findings of Yildirim (2010) and the present study also indicate that preparing exams like ALCPT may harm the natural development process of L2 learning and possibly L2 acquisition. The negative washback effect is also associated with the methodology the teacher adopts (Madaus & Kellaghan, 1992; Nolan, Haladyna, & Haas, 1992; Shepard, 1990).

Another important finding also enables discrimination between the participants. The cadets that were interviewed in the study were found to be more negatively influenced by the ALCPT. They reported that ALCPT simply does not contribute to their communicative skills in English.

4.4 Research objective 2: ILE and language learning

The third and fifth interview questions, "How does ALCPT affect your approach to learning a foreign language?" and "Has ALCPT influenced your feelings about learning English?" were utilized to collect data regarding language learning experiences of the participants. The insights provided by the participants were shown in these comments:

"In the spring semester we study the textbook, and there are some pair work and group activities, like conversation... However, in the fall semester, it's all vocabulary lists and mock exam papers... I think 80% is geared toward the ALCPT. We do ALCPT mock tests in class and we analyze the tests... I think these kinds of activities are specific for ALCPT practice. The teacher hopes that we progress fast and get higher scores."

"ALCPT has actually turned language learning into a kind of a memory game. I don't even dream of writing or speaking English. But I can tell you which words are asked in ILE and which ones are not. So my experience of learning a foreign language is mostly challenging, discouraging, being perfectly honest, a little bit boring."

It changes the teaching content. The teachers emphasize more grammar, and we do a lot of mock exercises in class. English class seems to aim at improving our scores in the ALCPT. However, I think my English level has not improved. We are just trained to do the test well.

English class is just for teaching to the test, one section after the other. What we practiced is what's in the test... Students actually aren't in the mood for [other activities] because they are not related to the ALCPT.

"ALCPT affected my view and understanding of learning a foreign language badly. Therefore, I feel bad about learning a language."

"Well, I feel disappointed. I might develop my English more effectively if ALCPT asked me to write a paragraph or utter something in English. In this format, the exam is nothing but a procedure for all of us. I regret it as we spent money and time, and so does the school! Generally, I like English but no way, not this way!"

The findings regarding the language learning experiences of the cadets show that, during the preparation for ALCPT, the participants utilized mostly cognitive strategies which led them to memorization and short term study habits. Various similar studies have previously focused on the impact of a specific test on the learning strategies or study skills of testees (Cheng et al., 2004; Green, 2007; Song, 2009). However, the findings of this study also maintain that a poor examination may also affect the motivation of the cadets, which may also lead to a test anxiety or to avoidance behavior towards the subject matter. These findings may be interpreted in terms of the theoretical discussion of Watanabe (2004) in which washback effect is associated with certain motivational theories.

In the light of the findings of the study, ALCPT was found to have a negative and unintended washback effect on cadets. These types of effects can be addressed as 'Micro level' or the test washback (Wall, 1997) and 'Macro level' or the test impact (Wall, 1997). Micro level effects were categorized in terms of the effects of ALCPT on the individual's learning and life. In this respect, these effects were L2 competences that remain weak and undeveloped, a lower level of motivation for learning the L2 and finally an overemphasis of cognitive learning due to the content of the exam. Learning a foreign language requires not only cognitive learning but also affective or social learning styles and strategies (Brown & Hudson, 1998), but getting prepared for ALCPT, the findings illustrate, requires little or no affective and social learning processes and contexts. The macro level effects, which refer to the social level effect of the exam, were the time spent preparing for the exam as well as the teachers' burden that was mentioned several times by the participants. Some participants believed that ALCPT was a heavy burden on their teachers' shoulder in that the teachers were considered unsuccessful if the students cannot, somehow, succeed in the exam. Many studies in the literature verify this critical point (Anderson, Muir, Bateson, Blackmore & Rogers, 1990; Noble & Smith, 1994; Watanabe, 2004). The final inference from the macro level effects would be that there was a general accord with the idea that ALCPT was a poor test, which should be updated and replaced with a more effective test.

With regard to the learning experiences, the findings can also be discussed under micro and macro level categorizations. As for micro level learning experiences, the participants were found to display short-term study habits and low order thinking skills which were reported as the memorization of the exam tips and the strategies that encourage a behaviorist or mechanical approach to foreign language learning, and finally as memorization and rote learning. This finding is supported by the study conducted by Anderson (Anderson et al., 1990), showing quite similar results. These micro level learning experiences were strongly related to the macro level ones, perhaps in a causal relationship. Participants reported suffering from anxiety and inhibition, and a feeling of failure, both of which led them to avoidance behavior regarding learning English, which supports the previous literature such as Cheng et al. (2004) and Shohamy (2001).

5. CONCLUSION AND IMPLICATIONS

The number of studies on the washback effect in language testing literature has been growing rapidly due to its critical impact on learning, learners and teachers and even the society or the complete educational system of a country. The high-stakes tests across countries have also led scholars to focus on the consequences of such tests on students, teachers and institutions (Alderson & Wall, 1993; Bachman & Palmer, 1996; Biggs, 1995;

Cheng, 2005; Messick, 1996; Shohamy, Donitsa-Schmidt & Ferman, 1996; Taylor, 2005; Yildirim, 2010).

In the present study, the results revealed that ALCPT exerts a negative impact on cadets. The results also show that the exam does not only have a micro level effect, which can be defined as the effect on an individual or a small group of individuals, but also a macro effect on a relatively large population of individuals studying or working in military academy contexts.

In this study, the reason why ALCPT exerts a harmful and negative washback effects on the cadets is attributed to the content and style of the test itself, which is a multiple-choice test claiming to assess grammar, vocabulary and reading-skills knowledge in the L2 – English in the context and scope of this study. The unanimous reactions of all participating cadets of this study were that they need to develop more than what ALCPT assesses. These findings are compatible with the related literature (Burrows, 2004; Song, 2006; Wall, 1999; Watanabe, 2000, 2001; Yildirim, 2010).

The issues raised in this study present some implications for the ALCPT stakeholders. For ALCPT test developers, I believe that the lack of positive washback on cadets is an indication that consideration needs to be given to modification of the structure of traditional, discrete-point ALCPT and employing more communicative language testing. The military academy should be aware that cadets might be engaging in activities that allow them to increase their ALCPT scores without having the English language proficiency necessary to perform at an academic level. Moreover, the academy should be aware that stringent reliance on the use of ALCPT scores as an indication of language proficiency might be negatively effecting the way in which cadets approach language learning. English teachers should be aware that the participants indicated that listening was a skill for which they felt they needed more practice.

Whilst this study is necessarily limited in terms of its scope and depth, it is hoped that it has been able to provide a picture of washback from a specifically cadet-oriented perspective, which has hitherto been somewhat neglected, while at the same time providing some reflections on how the washback effect might be minimized, with cadets' perception in mind. The study also goes some way to supporting the views that a test can influence 'what learners learn,' 'how learners learn' and 'the degree and depth of learning' (Alderson & Wall, 1993). A detailed inquiry into the impact of the test on cadets helps us understand how this influence is actually reflected within the military academy context in Taiwan. The deeper the understanding we have of our cadets, the better we are likely to help them to learn and minimize negative washback.

REFERENCES

- Andrews, S., J. Fullilove, and Y. Targeting Wong. 2002. "Washback—a Case-Study." *System* 30(2): 207–223.
- Alderson, J. C., & J. Banerjee. 2001. "Language Testing and Assessment (Part 1)." *Language Teaching* 34(4): 213–236.
- Alderson, J. C., and L. Hamp-Lyons. 1996. "TOEFL Preparation Courses: A Study of Washback." *Language Testing* 13(3): 280–297.

- Anderson, J. O., W. Muir, D. J. Bateson, D. Blackmore, and W. T. Rogers. 1990. *The impact of provincial examinations on education in British Columbia: General report*. Victoria: British Columbia Ministry of Education.
- Alderson, J. C. and Wall, D. 1993. "Does Washback Exist?" *Applied Linguistics* 14(2): 115–129.
- Bachman, L. F. and A. Palmer. 1996. *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: OUP.
- Bailey, K. M. 1996. "Working for Washback: A Review of the Washback Concept in Language Testing." *Language Testing* 13(3): 257–279.
- Bailey, K. M. 1999. *Washback in Language Testing*. TOEFL Monograph Series. Report Number: RM-99-04, TOEFL-MS-15. Princeton, NJ: Educational Testing Service.
- Bright, C., and J. Von Randow. 2004, September. "Tracking language test consequences: The student perspective". Paper presented at the Ninth National Conference on Community Languages and English for Speakers of Other Languages (CLESOL 2004), Christchurch.
- Brown, J.D., and T. Hudson, (1998). "The alternatives in language assessment". *TESOL Quarterly* 32(4): 653–675.
- Burrows, C. 2004. "Washback in classroom-based assessment: A study of the washback in Australian adult migrant English program". In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*. 113–128. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Cheng, L. 1998. "Impact of 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 of teachers' perceptions toward their classroom teaching". In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*. 147–170. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Cheng, L. 2005. *Changing Language Teaching through Language Testing: A Washback Study*. Cambridge: CUP.
- Chen, L. -M. D. 2002. "Taiwanese junior high school English teachers' perceptions of the washback effect of the basic competence test in English". PhD dissertation, The Ohio State University.
- 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." Doctoral dissertation. National Taiwan Normal University.
- Ferman, I. 2004. "The washback of an EFL national oral matriculation." In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods*. 191–210. Mahwah, NJ: Lawrence Erlbaum Associates.
- Glaser, B. G., and A. L. Strauss. 1967. *The discovery of grounded theory*. New York: Aldine.
- Gosa, C. M. C. 2004. "Investigating Washback: A Case Study Using Student Diaries." PhD thesis, Department of Linguistics and Modern English Language, Lancaster University.
- Green, A. 20007. "Washback to learning outcomes: a comparative study of IELTS preparation and university pre-session language courses." *Assessment in Education*, 14(1): 75–97.
- Gu, Y. Q., Q. F. Wen, and D. X. Wu. 1995. "How often is often? Reference ambiguities of the Likertscale in language learning strategy research." *Occasional Papers in English Language Teaching* 5: 19–35.

- Hamp-Lyons, L. 1997. "Washback, Impact and Validity: Ethical Concerns." *Language Testing* 14(3): 295–303.
- Hughes, A. 2003. *Testing for language teachers (2nd edition)*. Cambridge: Cambridge University Press
- Jiang, X. 2007. "A review of testing researching articles published in the past 10 years: review and future direction." *Foreign Language World* 2: 89–96.
- Li, X. 1990. "How Powerful Can a Language Test Be? The MET in China." *Journal of Multilingual and Multicultural Development* 11(5): 393-404.
- Madaus, G. F., & Kellaghan, T. 1992. "Curriculum evaluation and assessment." In P. W. Jackson (Ed.), *Handbook of research on curriculum*. 119–154. New York: Macmillan.
- Messick, S. 1996. "Validity and washback in language testing." *Language Testing* 13: 241–256.
- Mizutani, S. 2009. "The mechanism of washback on teaching and learning." PhD thesis, The University of Auckland.
- Noble, A. J., and M. L. Smith. 1994. *Measurement-driven reform: Research on policy, practice, repercussion* (Tech. Rep. 381). Tempe, AZ: Arizona State University, Center for the Study of Evaluation.
- Nolen, S. B., T. M. Haladyna, and N. S. Haas. 1992. "Uses and abuses of achievement test scores." *Educational Measurement: Issues and Practice* 11(2): 9–15.
- Qi, L. 2004. *The Intended Washback Effect of the National Matriculation English Test in China: Intentions and Reality*. Beijing: Foreign Language Teaching and Research Press.
- Read, J., and B. Hayes. 2003. "The Impact of IELTS on Preparation for Academic Study in New Zealand." *IELTS International English Language Testing System Research Reports* 4:153-206.
- Saif, S. 2006. "Aiming for Positive Washback: A Case Study of International Teaching Assistants." *Language Testing* 23(1): 1–34.
- Shepard, L. A. 1990. "Inflated test score gains: is the problem old norms or teaching the test?" *Educational Measurement: Issues and Practice* 9: 15–22.
- Shih, C. 2007. "A new washback model of students' learning." *The Canadian Language Review*, 64(1): 135–162.
- Shohamy, E. 2001. *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. London: Longman.
- Shohamy, E., S. Donitsa-Schmidt, and I. Ferman 1996. "Test Impact Revisited: Washback Effect over Time." *Language Testing* 13(3): 298–317.
- Song, D.Y. 2009. "Applicability of an EBB Scale to English Speaking Performance Assessment in the Korean Secondary School." *English Language Teaching* 21(1), 111–135.
- Taylor, L. 2005. "Washback and Impact." *ELT Journal* 59(2): 154–5.
- Tsagari, D. 2007. "Review.Washback in Language Testing: What Has Been Done? What More Needs Doing?" <http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED497709>
- Wall, D. 1997. "Impact and Washback in Language Testing." In C. Clapham and D. Corson (Eds.), *Encyclopaedia of Language and Education. Vol. 7. Language Testing and Assessment*. 291-302. Dordrecht: Kluwer Academic.
- Wall, D. 1999. "The impact of high-stakes examinations on classroom teaching: A case study using insights from testing and innovation theory." Doctoral dissertation, Lancaster University.

- Wall, D. 2000. "The Impact of High-Stakes Testing on Teaching and Learning: Can This Be Predicted or Controlled?" *System* 28: 499–509.
- Wall, D. 2005. *The Impact of High-Stakes Examinations on Classroom Teaching: A Case Study Using Insights from Testing and Innovation Theory*. Cambridge: CUP.
- Wall, D., and J. C. Alderson. 1993. "Examining Washback: The Sri Lankan Impact Study." *Language Testing* 10(1): 41–69.
- Wang, W. 2010. "Investigating the Washback Effect of the New CET 4 Listening Comprehension Subtest on Language Learners." *Chinese Journal of Applied Linguistics* 33(5): 28–39.
- Watanabe, Y. 2000. "Washback Effects of the English Section of Japanese University Entrance Examinations on Instruction in Pre-College Level EFL." *Language Testing Update* 27: 42–47.
- Watanabe, Y. 2001. "Does the University Entrance Examination Motivate Learners? A Case Study of Learner Interviews." In Akita Association of English Studies (Ed.), *Trans-Equator Exchanges: A Collection of Academic Papers in Honour of Professor David Ingram*. 100–110.
- Watanabe, Y. 2004. "Teacher factors mediating washback" in Cheng L, Y Watanabe and A Curtis (Eds.). *Washback in Language Testing: Research Contexts and Methods*. 129-146. NJ: Lawrence Erlbaum.
- Yıldırım, Ö. 2010. "Washback Effects of a High-Stakes University Entrance Exam: Effects of the English Section of the University Entrance Exam on Future English Language Teachers in Turkey." *Asian EFL Journal* 12(2): 92–116. <http://www.asianefl-journal.com>