

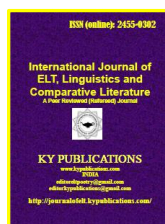


## The Impact of Washback-based Approach of the Achievement Test on Developing the EFL Learning Process

(A Case Study of First Year Secondary School Students in Saudi Arabia)

IBRAHIM AHMED EISA ABDALLAH\*, Prof.Dr. ISHRAGA BASHIR ELHASSAN,  
Dr.EINAS AHMED ABDEL RAHMAN

\* E-mail: [Abdallahibrahim659@yahoo.com](mailto:Abdallahibrahim659@yahoo.com)



### ABSTRACT

This study investigates the impact of washback- based approach of the achievement test on developing the EFL learning process. The study adopts experimental and statistical methods to analyze and verify the results in which frequencies, percentages; means and their standard deviation were used. The analysis was carried out through the Statistical Package for Social Sciences (SPSS) technique. The sample of the study comprises of 30 students of higher secondary school in the Eastern Region of Saudi Arabia to answer the questions of the test. One tool has been used for collecting data that relevant to the study, namely a test (pre- and posttests) to the first-year secondary students to provide answers to the study question and hypothesis. The researcher designed a specific teaching plan that aiming at improving the first-year secondary students' language knowledge and enable them to use the language communicatively in the different context of life. For carrying out the plan, all available techniques, materials, and aids were exploited. However, before students being exposed to this plan, the researcher exposed the target sample to a diagnostic test (pre-test) in order to check the actual problem and need of students' language knowledge. Soon after the pre-test administration and registration of the results, the researcher started carrying out the plan systematically. After four months, the same test was administered to the same group for the second time as an achievement test (post-test) and the results were registered. The results of the two tests (pre/posttests) were compared to see whether the teaching plan had a significant effect on the students' performance. The results have shown great improvements in the performance of the learners and they were influenced by the treatment plan positively. The study offered some recommendations and suggestions for further studies that encourage the EFL learners to be aware of the impact of wash back-based approach the achievement test on developing the learning process.

**Key words:** Admission, knowledge, improvement, influence

### Introduction

In recent years, the impact of major tests and examinations on language learning has become an area of significant interest for testers and teachers. One aspect of test impact is the washback, which is traditionally described as the positive and negative effects that result from a test. It is said to create a narrowing of the curriculum in the classroom so that teachers and learners focus solely on the areas to be tested. It is widely accepted in language teaching that major tests and examinations inevitably have an influence both within the



education system and beyond it. There are two terms which are used to refer to this influence: impact and washback. The broader concept is impact, which Wall (1997: 291) defined as "...any of the effects that a test may have on individuals, policies or practices, within the classroom, the school, the educational system or society as a whole.". On the other hand, there have been attempts to generate positive washback by means of examination reform to encourage teachers and learners to adopt more modern communicative approaches to language learning. The test that is the subject of the present study is the achievement test or the attainment test, which has become one of the preferred method of assessing the EFL learners who are seeking admission into further education or employment opportunities. Courses which claim to prepare students for the test have become an increasingly common feature of the programs offered by both private and public-sector language schools.

### Literature Review

#### The Origin of Washback

Although the subject of the effects of examinations has long been discussed in the literature of general education Vernon (1956); it has been common in the literature on testing that washback, as it is known for us now, has come to attract the attention of test researchers only at the beginning of the 1990's. Before that date, testing specialists and applied linguists used different terms to refer to the idea of examination influence. Language testing researchers have realized that the emergence of the concept washback is the result of considerable reforms and advances that have taken part in the domain of language testing mainly during the last two decades at the end of the twentieth century.

In a comprehensive study on how the concept washback has come to emerge, Tsagari (2007) designs an artificial time framework divided into three distinct but successive phases that clearly display how the concept has evolved overtime in the scene of language testing. These phases are the "pre-1990's", the "1990's", and the "post 1990's". Tsagari identifies that the initial phase was mainly characterized as the period of time when writers recognized the existence of the examination influence phenomenon but no one accounted for it. Second, the 1990's phase was thought to be different from the previous one and was basically dominated by the publication of a seminal paper by two prominent language testing researchers, Alderson and Wall, who are greatly indebted the fact they were the first who questioned the nature of examination influence; and more importantly, they managed to re-conceptualize this phenomenon by proposing a set of hypotheses. The third phase, the "post 1990's", or as Tsagari names it the 'reality phase', was significant since substantial models of washback have been developed in order to accurately explain and analyze the nature of this phenomenon. Such a way of looking at the subject is seen by testing experts to be a serious attempt and a step forward in the study of washback in language testing.

#### Definitions of Washback

Washback, sometimes referred to as backwash (Biggs, 1995, 1996 in Cheng, 2000), can be generally understood as the effect of an examination on teaching and learning (Chen, 2002; Hughes, 2003; Cheng, 2003). Not all scholars, however, have agreed to its definition. Alderson and Wall (1993) restricted the use of the term 'washback' to "classroom behaviors of teachers and learners rather than the nature of printed and other pedagogic material" (p. 118). They would also consider washback to be what teachers and learners do that "they would not necessarily otherwise do" (p. 117). Messick (1996) states that in order to be considered washback, good or bad teaching has to be "evidentially linked to the introduction and use of the test" (p. 16). Moreover, Wall (1997) makes a clear distinction between washback and test impact. The latter would refer to the effect of a test on "individuals, policies or practices, within the classroom, the school, the educational



system or society as a whole" (cited in Cheng and Curtis, 2004, p.4). Other researchers Andrews, (2002) do not make that distinction and they consider that narrow and wider effects can be included under the term washback. For the purposes of this study, washback will be understood in the wider sense, that is, including what some scholars call 'impact'.

The concept of washback has been associated with validity. Morrow (1986) refers to "washback validity" to describe the quality of the relationship between testing and teaching and learning" cited in Cheng (2000, p.4). For Messick (1996) washback is part of construct validity, and it is an inherent quality of any kind of assessment, especially when the results are used for important decisions. For him, washback contributes to the consequential aspect of construct validity, but information about the operative level of test validity should help one distinguish test washback per se from the effects of good or bad educational practices regardless of the quality of the test. So, washback in itself is not a reliable criterion to establish test validity. It is other test properties, like authenticity and directness that are likely to produce washback.

Messick (1996) claims that:

"If a test is deficient because it has construct under representation or construct-irrelevant characteristics, then good teaching cannot be considered an effect of the test, and conversely, if a test is construct-validated, but there is poor teaching, then negative washback cannot be associated with the test. Only valid test can increase the likelihood of positive washback"

### Types of Washback

Alderson and Wall (1993) distance themselves from a simplistic assumption about the way a test can influence behaviors. Therefore, they developed some washback hypotheses according to what is influenced: teaching, learning, content, rate, sequence, degree, depth, attitudes and also the number of teachers or learners affected by a test. Which hypotheses will be put forward depends on the nature of the test, the educational context, and the nature of the decisions that are taken on the basis of the test results. Actually, there seems to be a number of variables in society, education, and schools that determine how washback will appear.

When studying washback, it is also possible to focus on participants (teachers, students, material developers, publishers), process (actions by participants towards learning), and products (what is learned and the quality of learning), as suggested in Hughes's trichotomy model (Hughes, 1993 as cited in Bailey, 1996). Watanabe (2004) proposes disentangling the complexity of washback by conceptualizing it in terms of: Dimension (specificity, intensity, length, intentionality and value of the washback), aspects of learning and teaching that may be influenced by the examination, and the factors mediating the process of washback being generated (test factors, prestige factors, personal factors, macro-context-factors).

Usually researchers focus on one aspect or type of washback. In Alderson and Wall's study in Sri Lanka (Alderson & Wall, 1993; Wall, 1996), the introduction of a test of English as a foreign language proved to produce faster changes in the content of teaching than changes in teaching methodology. Cheng (1997), in the preliminary results of a study of the washback effect of the Hong Kong Certificate of Education Examination in English in Hong Kong secondary schools, reports that washback effect "works quickly and efficiently in bringing about changes in teaching materials [...] and slowly and reluctantly and with difficulties in the methodology teachers employ" (p.1). Cheng introduces the term 'washback intensity' to refer to the "degree of washback effect in an area or a number of areas that an examination affects most" (p.7).



Andrews et al. (2002) found out in their study that the impact of a test can be immediate or delayed. According to these researchers, washback seems to be associated primarily with 'high-stakes' tests, that is, tests used for making important decisions that affect different sectors., for example, determining who receives admission into further education or employment opportunities, he identifies as 'high' such situations when admission, promotion, placement or graduation are dependent on the test.

Another aspect that has been studied is whether the test has been used as a lever for change (Pearson, 1988 in Cheng, 1997), so everything, from textbooks to staff, works to achieve better scores. Cheng (2000) reports:

"Tests are often introduced into the education system to improve teaching and learning, especially in centralized countries where tests are considered an efficient tool for introducing changes into an educational system without having to change other educational components. Therefore, in some countries these tests can be considered "the engine for implementing educational policy" Cheng (2000, p. 6)".

#### Factors Affecting Washback

It has been demonstrated that it is simplistic to believe that a test can result in all desired changes in teaching and learning. Education is a complex phenomenon and there are many factors involved in bringing about changes, like the school environment, messages from administration, expectations of teachers and students, for example. Saif (2000) argues that an analysis of the needs and objectives of learners and educational systems should be carried out as a starting point for the research in washback. Wesche (1983 in Bailey, 1996), points out that when tests reflect the situations, content and purpose where learners will use the language, they are likely to improve motivation. Shohamy et al. (1996) consider factors like the status of the subject-matter tested, the nature of the test, and the use to which the test scores are put (p. 300). Wall (1996) provides a list of factors which might have prevented the examination in Sri Lanka from providing an effective 'lever for change'. These ranged from teachers' factors (lack of understanding of exam, the nature of the change desired, resistance to change, unfamiliarity with the test format and content) to more macro factors like gap between designers of test and teachers, lack of well trained teachers, overload of teachers, etc. Besides, according to Andrews et al. (2002), the innovating effect of a testing innovation is affected by the teachers and how they interpret the innovation, which may differ from what the conceivers of the test had in mind. Another variable can be the published materials in use Andrews et al. (2002).

As can be seen, washback is a very complex notion. It can refer to the effect of an examination in the classroom, but also in the school, in the educational system and also in the society. Besides, this effect does not always take place directly but it is mediated by a number of factors, like the teachers' perception of the test, the status of the test as well as that of the subject – matter tested, the macro – context where the examination is used, the purpose of learning the language in the context, among others. Additionally, in order to study the washback effect, it is necessary to look at the people that participate in the educational process, to the actual classroom events and activities, and to the outcomes of these processes.

#### Language Testing and Washback Effect

Test is a tool of any exam and possesses central position in the framework of an examination system. Focus, form, content and manipulation of a test can affect validity of exam and ultimately affects the quality of English language teaching and learning that aims at imparting language skills. Washback has become a well versed common academic phenomenon in all the institutions. Comprehensively washback can be defined as the influence of testing on teaching and learning Gates (1995).



Washback is not something new it has always been present in the academic career of every student and teacher in the form of the question: Will this be included in the test? Naturally, teachers and students are inclined to adapt teaching and learning according to the requirement of test especially when the test results would have a bearing on their future Buck (1988). Washback has become an important aspect of present day education because test scores have played a decisive role in the achievement of educational and social pursuits Cheng (1997).

Despite their beneficial or harmful washback, public exam and test have maintained their place because they hold a mirror to all the other activities going on under the umbrella term of teaching-learning process and other processes in education system. It is important to understand that washback is a neutral term that can be considered as positive or negative (Bachman & Palmer, 1996). On the basis of its scope Bachman and Palmer (1996) differentiated test impact from test washback. They considered test impact a 'macro level' a phenomenon having social and institutional implications while observed test washback at 'micro level' limited to teaching, learning, teachers and students. In the same way, Wall (1997) claimed that test impact is any kind of effect that an exam or test might have on individuals, policies and other educational practices, while test washback is the effect of test specifically on teaching and learning.

Considering distinction between test impact and test washback it was inferred that it is neither easy to measure negative washback effect of exams nor to modify exams to exert a positive effect because washback not only influences teaching and learning but it is also influenced by many other factors.

### Statement of the Study Problem

The researcher as a teacher of English language has been met by a countless number of complaints and discontent among teachers of English language as well as students about the results achieved by students in their final exams. Many students who performed well in Exams and obtained full marks, unfortunately were unable to manage their learning. This phenomenon would hypothetically be associated with a negative washback-based approach of the achievement test, which is the effect of test on the learning process. The researcher would also like to discover, whether the EFL achievement tests at their current situation have positive or negative washback.

### Objective of the Study

This study attempts to achieve the following objectives:

1. 1.To investigate the impact of washback-based approach of the achievement test on developing the EFL learning process.
2. To explain how washback-based approach of the achievement test can promote EFL learners' proficiency.

### Question of the Study

The current study attempts to answers the following question:

What is the impact of the washback-based approach of the achievement test on developing the EFL learning process?



### Hypothesis of the Study

This study tries to find out respondents' reaction to this hypothesis:

Washback- based approach of the achievement test has positive impact on developing the EFL learning process.

### Significance of the Study

The importance of investigating more about washback-based approach of the achievement test is crucial.

\*Firstly, this study provides some of the needed knowledge and understanding to help future teachers and language test developers to design instruction and tests that properly account for washback.

\*Secondly, language classroom teacher has a vested interest in knowing how washback affects the learning process, and how to best recognize the effects of washback-based approach, either positive or negative, on her/his teaching.

\*Thirdly, learning more about washback in this environment has important implications for teacher education programs that are designed to train new instructors, to introduce valuable classroom techniques and strategies, and to review important issues in the development of course syllabi and teaching materials. It is hoped that this study will be of great value to teachers of English language and students of Secondary School in Saudi Arabia in particular in dealing with investigation into the washback-based approached effects on EFL achievement tests. It is also hoped that the findings of this study might be useful to learners of English as a foreign language in general.

### Limits of the study

The study is limited to:

1. The limits of Title: The Impact of Washback-based Approach of the Achievement Test on Developing the EFL Learning Process
1. 2.The limits of Place: The study was applied and carried out at Higher Secondary School in the Eastern Region of Saudi Arabia as a case study.
2. 3.The limits of Time: The study was carried out and applied during the academic year: 2016/2017).

### Methodology of the Study

The study adopted the experimental method throughout the research. The researcher used one tool that aiming at gathering the needed data and to investigate the impact of washback-based approach of the achievement test on developing the learning process. The First test was administered during the beginning of the academic year 2016/ 2017 for first grade secondary as a diagnostic test (pretest) to discover the actual problem of the students in language learning and the second test was given to them after 15 weeks as an achievement test (posttest) to evaluate the whole process. Their ages are between 15 to 16 years and they have been studying English language for six years.

### Population and sampling

The population of the present study consists of secondary school students in the Eastern Region of Saudi Arabia to check the impact of washback-based approach of the achievement test on assessing the learning process. The sample of the study consists of (30) students of higher secondary school for the administration of the pre- and posttest. The students were homogenous in terms of learning experience,



language abilities and ages. That is, most of the students are around 18years of old and they have been studying English language for six years.

#### Tool of data collection and statistical methods

The researchers used pre and posttests as an effective tool to investigate the impact of washback-based approach of the achievement test on developing the EFL learning process.

The data were analyzed through a descriptive statistical method in which means, significance difference in students' scores and their standard deviations are used to support the question of the research. The analysis was carried out through the Statistical Package for Social Sciences (SPSS). The researcher establishes the test reliability by employing the retest procedure which involves giving the test twice to the same group of students which concluded that there is higher correlation between the scores of each student in the two administrations of the test.

#### Analysis and Discussion of the Study Tool

This section has covered the data analysis of the study. The researcher presented and discussed the results of the study according to following question.

\* What is the impact of washback- based approach of the achievement test on developing the EFL learning process?

After tabulating and analyzing the data of the study through the test, the results are shown in the following tables:

Table 1: The subjects' results in the pre-and post-tests

Student No.	Questions Items										Total	
	1		2		3		4		5			
	Reading 10		Grammar 10		Vocabulary 10		Conversation 10		Writing 10			
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1-	4	8	6	6	5	7	5	10	0	3	20	33
2-	6	10	5	5	5	10	6	10	0	4	22	39
3-	5	9	5	8	4	5	5	10	4	5	23	37
4-	8	10	7	7	5	7	5	10	5	8	30	42
5-	7	9	3	5	4	7	3	9	0	3	17	33
6-	9	10	3	5	1	6	0	4	0	3	13	28
7-	5	10	7	7	8	10	8	10	5	7	33	44
8-	3	6	1	3	4	5	1	4	0	5	9	23
9-	4	6	1	3	0	2	0	3	0	2	5	16
10-	3	7	3	8	4	5	0	6	0	3	10	29
11-	5	10	5	10	10	10	5	10	5	7	30	47
12-	9	10	2	4	0	3	0	4	0	6	11	27
13-	8	10	8	10	8	8	10	10	5	5	39	43
14-	8	10	10	10	10	10	10	10	8	9	46	49
15-	4	7	2	5	0	3	0	5	0	3	6	23
16-	9	9	4	6	3	5	8	10	5	6	29	36





Student No.	Questions Items										Total  50 Marks	
	1		2		3		4		5			
	Reading 10		Grammar 10		Vocabulary 10		Conversation 10		Writing 10			
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
17-	8	9	0	3	5	5	10	10	0	2	23	29
18-	7	8	4	4	6	7	3	4	5	5	25	28
19-	8	10	2	5	9	9	10	10	5	7	34	41
20-	8	10	8	9	7	8	8	10	5	7	36	44
21-	6	7	6	6	5	9	5	10	3	4	25	36
22-	5	8	2	4	7	8	0	2	0	3	14	25
23-	4	4	1	3	1	5	0	3	0	2	6	17
24-	5	10	5	7	8	8	5	10	0	2	23	37
25-	4	6	0	4	0	3	0	0	0	2	4	15
26-	6	6	0	2	0	2	0	3	0	3	6	16
27	7	10	3	5	3	7	5	5	2	4	20	31
28-	4	9	6	6	5	8	5	6	3	4	23	33
29-	8	10	8	9	10	10	9	10	8	9	43	48
30-	7	7	4	5	7	8	10	10	5	7	33	37

Table 2: T-test for the differences between the scores in pre and posttests in reading part

Tests	Means	SD	DF	T- value	p-value
Pre –test	6	0.21	29	14	0.040
Post –test	9	0.2			

For the scores gained from the pretest (*PET test*) the mean value calculated. was (6), and the mean value for the post test was (9) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.040) being smaller than (0.05), our null hypothesis is rejected, this indicate that there is significance difference in students' scores after we retesting them on reading part

Table (3): T-test for the differences between the scores in pre and posttests in grammar part

Test	Means	SD	DF	T- value	p-value
Pre-test	4	0.4	29	16	0.0401
Post test	6	0.3			

For the scores gained from the pretest (*PET test*) the mean value calculated. was (4), and the mean value for the post test was ( 6) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.0401) being smaller than (0.05), our null hypothesis is rejected, this indicate that There is significance difference in students' scores after we retesting them on grammar part

Table 4: T-test for the differences between the scores in pre and posttests in vocabulary part

Group	Means	SD	DF	T- value	p-value
Pre-test	5	0.33	29	14	0.02
Post test	8	0.43			





For the scores gained from the pretest (*PET test*) the mean value calculated. was (5), and the mean value for the post test was (8) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.02) being smaller than (0.05), our null hypothesis is rejected, this indicate that There is significance difference in students' scores after we retesting them on vocabulary part

Table 5: T-test for the differences between the scores in pre and posttests in conversation part

Test	Means	SD	DF	T- value	p-value
Pre-test	5	0.29	29	15	0.00
Post test	9	0.4			

For the scores gained from the pretest (*PET test*) the mean value calculated. was (5), and the mean value for the post test was (9) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.00) being smaller than (0.05), our null hypothesis is rejected, this indicate that There is significance difference in students' scores after we retesting them in conversation section

Table 6: T-test for the differences between the scores in pre and posttests in writing part

Test	Means	SD	DF	T- value	p-value
Pre-test	2	0.2	29	14	0.0212
Post test	4	0.3			

For the scores gained from the pretest (*PET test*) the mean value calculated. was (2), and the mean value for the post test was (4) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.0212) being smaller than (0.05), our null hypothesis is rejected, this indicate that there is significance difference in students' scores after we retesting them on writing part

Table 7: T-test for the differences between the scores in pre and post tests for over all

Test	Means	SD	DF	T- value	p-value
Pre-test	23	0.05	29	14	0.0000
Post test	35	0.33			

For the scores gained from the pretest (*PET test*) the mean value calculated. was (23), and the mean value for the post test was (35) Moreover T-test was employed on these scores for hypothesis testing purposes. As the result of T-test suggests (P-value 0.0000) being smaller than (0.05), our null hypothesis is rejected, this indicate that There is significance difference in students' scores on the overall part.

As shown from the previous tables the results of written pretest and posttest explains that there is a significant difference between the scores in pre and posttests in all parts (reading, grammar, vocabulary, conversation and writing). In addition to, the total difference between the scores is (23) for the pretest and (35) for the post test (overall figure). The above analysis proves that there is a development in the performance of the learners and the respondents have influenced positively during the treatment program.

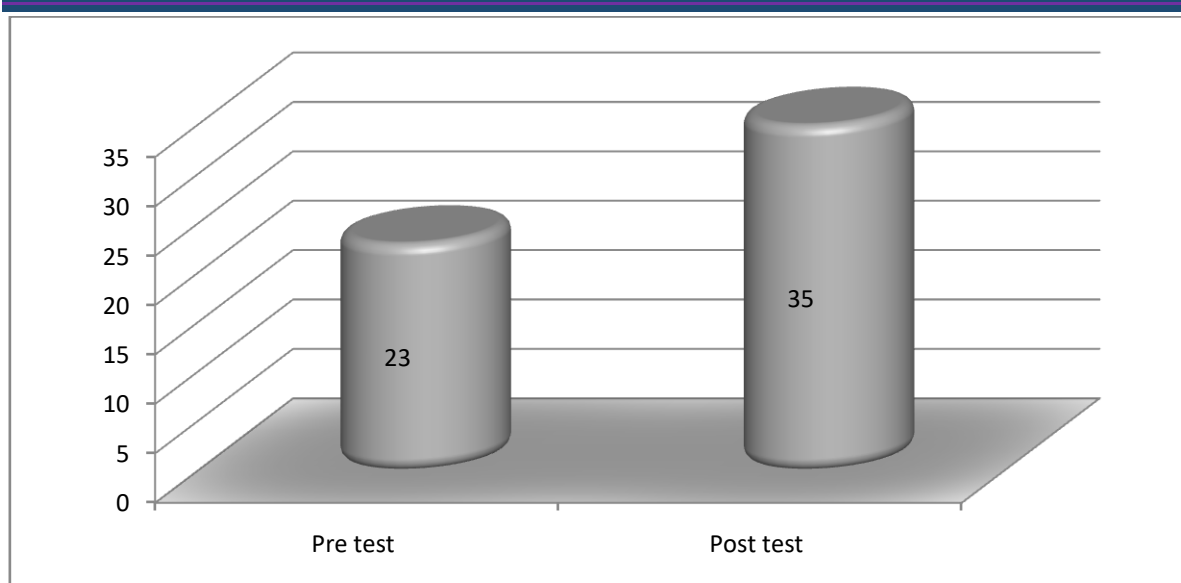


Figure 1: shows the difference between the scores on the overall in pre and posttests.

### Summary of the results

The researcher has discussed the results of the study and reached to the following out comes:

1. There was a remarkable progress shown in the performance of the subjects and they were influenced by the treatment plan positively.
2. Learners focus more on the test taking than on language learning strategies, this actually will lead to negative washback.
3. The process of exposing learners to the actual learning experience not just teaching them to pass the test was an influential factor.
4. The result of writing test shows that the performance of most study subjects in writing skill is weak. There for, writing skill as on one of the most important language competence is more challenging to them. Tables (1 and 6).

### Recommendations

The researcher recommends the following items that could be adopted by the other researchers:

1. More attempts should be done to generate positive washback by means of examination reforms
2. EFL students in Saudi Arabia should be recommended to focus on language learning strategies and how to use it in public or in real life contexts.
3. Teachers and learners should adopt more modern communicative approaches to language learning.
4. Students should be encouraged by teachers to study the language for communicative purposes not only for passing the test.

### Suggestions for further studies

The researcher suggests the following topics that could be conducted by the other researchers:

1. Investigating more about washback-based approach of the achievement test is crucial.
2. Using modern aids and equipment to teach and practice English language skills is of great value.



3. Exploring up to date effective teaching and learning methods to promote the positive washback and avoid the negative one.
4. Teaching plan should cover every section in the textbook although some sections are unlikely to be tested in the examination.
5. Testing the overall competence of the students in English such as listening, speaking, reading, and writing is required.

#### BIBLIOGRAPHY

- Alderson, J.C., & Wall, D. (1993). *Does washback exist? Applied Linguistics. International Journal of Educational Development*, 14, 115-129.
- Andrews, S., Fullilove, J. & Wong, Y. (2002). *Targeting washback – a case study. System*, 30, 207-223. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Bailey, D.M. (1996). *Working for washback: A review of the washback concept in language testing. Language Testing*, 13, 257-279. Brumfit, C., & Oxford: Oxford University Press.
- Biggs, J. (1995). *Assumptions underlying new approaches to educational assessment. Curriculum Forum*, 4 (2), 1-22.
- Biggs, J. (1996). *Testing: To educate or to select? Education in Hong Kong at the cross-roads*. Hong Kong: Hong Kong Educational Publishing Co.
- London: Longman.
- Buck, G. (1988). *Testing listening comprehension in Japanese university entrance examinations. JALT Journal*, 10, 12-42.
- Chen, L. (2002). *Washback of a public exam on English teaching. (ERIC Document Reproduction Service No. ED472167)*.
- Cheng, L. (1997). *Does Washback influence teaching? Implications for Hong Kong. Language and Education*, 11 (1), 38-54.
- Cheng, L. (2000). *Washback or backwash: A review of the impact of testing on teaching and learning. (ERIC Opinion papers ED 442280)*.
- : A Hong Kong case study. *Journal of Classroom Interaction* 38 (1), 1-10.
- Cheng, L., Watanabe, Y. & Curtis, A. (Eds.) (2004). *Washback in language testing: Research contexts and methods*. Mahwah, N.J.: London: Lawrence Erlbaum.
- Gates, S. (1995). *Exploiting washback from standardized tests*. Tokyo: Japan Association for language Teaching. Language in India [www.languageinindia.com](http://www.languageinindia.com)
- Hughes, A. (1993). *Backwash and TOEFL 2000. Unpublished manuscript*, University of Reading in England.
- Hughes, A. (2003). *Testing for language teachers. (2nd Ed.)* Cambridge: Cambridge University Press.
- Messick, S. (1996). *Validity and washback in language testing. Language Testing*, Cambridge, England: Cambridge University Press.
- Morrow, K. (1986). *The evaluation of tests of communicative performance. In M. Portal (Ed.), Innovations in language testing. (pp. 1-13)* London: NFER/Nelson.
- Pearson, G. (1988). *Tests as levers for change. In D. Chamberlain & R. J. Baumgardner (Eds.) ESP in the classroom practice and evaluation (pp. 98-107). Modern English Publications*.
- Saif, S. (2000). *Theoretical and empirical considerations in investigating washback: A study of ESL/EFL learners [Abstract]. DAI*, 61, 03, 966
- Shohamy, E. (1996). *Test impact revisited: Washback effect over time. Language Testing*, 13(3), 298-317.
- Tsagari, D. (2007). *Investigating the Washback effect of the FCE examination (Cambridge/ESOL) in Greece. Greece, UCLES*.



---

Vernon, P.E. (1956). *The measurement of abilities (2nd ed.)*. London: University of London Press.

Wall, D. (1996). *Introducing new tests into traditional systems: Insights from general education and from innovation theory*. *Language Testing*, 13 (3), 334-354. *Washback of the FL Text...* 19 Arizona Working Papers in SLAT – Vol.12

Wall, D. (1997). *Impact and washback in language testing*. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (pp. 291-302). Dordrecht: Kluwer Academic.

Watanabe, Y. (2004). *Methodology in washback studies*. In L. Cheng, Y. UK: Penguin.

Wesche, M.B. (1983). *Communicative testing in a second language*. *Modern Language Journal*, 67, 41-55.

---