

People's Democratic Republic of Algeria

Ministry of Higher Education and scientific Research

Mohamed Kheider University of Biskra

Faculty of Arabic Language Arts and Foreign Languages

Branch of English



**INVESTIGATING WASHBACK EFFECTS
ON THE STUDENTS' PERFORMANCE
IN STANDARDIZED TESTS**

A CASE STUDY OF THIRD YEAR SCIENTIFIC CLASSES AT HIGH SCHOOL
OF SMATTI MOHAMED EL ABED IN OULED-DJELLAL

A Dissertation Submitted for the Partial Fulfillment of the Master's
Degree in Sciences of Languages

Presented by:

Miss. Dalila Bezziou

Supervised by:

Dr. Bacher Ahmed

Board of Examiners

Dr. Bacher Ahmed

Mrs. Hassina Nachoua

Mrs. Rabehi Salima

June 2013

References

- Alderson ,J C and Wall, D (1993). *Does WashbachExist?*Applied Linguistics.
- Alderson ,J C. Claphan , C and Wall , D (1995). *Language Test Construction and Evaluation*. Cambridge University Press . Printed in U.K.
- Bachman , L F (1990) . *Fundamental Considerations in Language Testing* .Oxford University Press. Printed in China.
- Bachman , L F and Cohen, A D (1998) . *Interfaces Between Second Language Acquisition and Language Testing Research* . Cambridge University Press.
- Biggan, J (2008).*Succeeding With Your Master's Dissertation*. Open University Press.Printed in U.K.
- Brown, H. D (2004).*Language Assessment: Principles and Classroom Practices* . Pearson Education , Inc. Printed in U.S.A.
- Brown, H.D (2007). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. 3rd ed. Pearson Education, Inc. Printed in U.S.A.
- Burgess, S and Head, K (2005).*How to Teach for Exams*. England. Pearson Education Limited.
- Cheng, L (2005) .*Changing Language Teaching Through Language Testing : A WashbachStudy*. Cambridge University Press. Printed in U.K.
- Cohen , A.D (1980) . *Testing Language Ability in the Classroom* . Newbury House Publishers, Inc.
- Cohen .L ,Manion. L and Morrison . K (2000) . *ResearchMethods Education*. 5th ed. RoutledgeFalmer.
- Davies, P andPearse, E(2000).*Success in English Teaching* .Oxford University Press. Printed in China.
- Davies,S(2006). *The Essential Guide to Teaching*. Pearson Education Limited.1st published in Great Britain.
- Denscombe, M (2007) .*The GoodResearch Guide*. 3rd ed. Open UniversityPress.
- Elatia, S(2003). *History of the Baccalaureate*. University of Illinois at Urbana- Champaign .
- Fulcher, G and Davidson , F (2007). *Language Testing and Assessment*. Routledge.
- Harmer, J(2001).*The Practice of English Language Teaching*. Pearson Education Limited.

- Hawkey, R (2004). *Studies in Language Testing*. Cambridge University Press. Printed in U. K
- Hedge, T (2000). *Teaching and Learning in the Language Classroom*. Oxford University Press. Printed in Hong Kong.
- Hughes, A (2003). *Testing for Language Teachers*. Cambridge University Press. Printed in U.K.
- Kellaghan, T. Madaus, G and Airasian, P W (1982). *The Effects of Standardized Testing*. Kluwer-Nijhoff Publishing. Printed in U.S.A.
- McNamara, T (2000). *Language Testing*. Oxford University Press. Printed in China.
- McMillan , J H (2007). *Classroom Assessment: Principles and Practice for Effective Standards – based Instruction*. Pearson Education, Inc . Printed in U.S.A.
- Oosterhof, A (2003). *Developing and Using Classroom Assessments*. 3rd ed. Florida State University .Pearson Education , Inc.
- Pedulla. J J, Abrams. LM, Madaus. G F, Russell. M K, Ramos. M A and Miao J (2003). *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings From a National Survey of Teachers*. Lynch School of Education. Boston College.
- Pinter, A (2006) .*Teaching Young Language Learners*. Oxford University Press. Printed in China.
- Seifert, K and Sutton, R (2009) .*Educational Psychology*. 2nded. Kanada.
- Shiel, G. Kellaghan, T and Moran, G (2010). *Standardized Testing in Lower Secondary Education*. Dublin
- Spratt, M .Pulverness , A and Williams, M (2005) .*The Teaching Knowledge Test Course (TKT)*. Cambridge University Press, Printed in U.K.

Abstract

The present research aimed at “Investigating Washback Effects on the Students’ Performance in Standardized Tests”. For that purpose, the following main research questions are addressed: What is the nature and the scope of washback effect on teachers and students? What are the perceptions of the instructors and the students related to the effects of BAC examination as an example of such standardized tests? What do instructors think about the changes to the examination that might be useful? Based on the above research questions, we hypothesize that, firstly: If curricular goals align well with the material instructors actually teach, and what students actually wish to learn as well as what is tested, positive washback effects will tend to be strong. Secondly: If test content does not match well with these components, washback is apt to be either ineffectual or negative. According to descriptive methodology of two questionnaires directed to both teachers and students, our concluding findings are the main levels affected by standardized tests (teachers, students, curricula and schools) influence each other at the same time. Such influences can be positive and /or negative. There is a strong need for balancing and matching adequately these levels, and controlling, factors (such as motivation, self-confidence, anxiety, stress, mixed ability, teaching experience and learning strategies) that influence the effects of tests.

ملخص

البحوث الحالية تهدف إلى "التحقيق في الآثار الانعكاسية للاختبارات على أداء الطلبة". لهذا الغرض، نتناول الأسئلة البحثية الرئيسية التالية: ما هي طبيعة ونطاق الآثار الانعكاسية على المعلمين والطلاب؟ ما هي تصورات المعلمين والطلاب المتصلة بالآثار الانعكاسية للاختبارات البكالوريا؟ ما يعتقد المدربون القيام به حيال التغييرات التي قد تكون مفيدة؟ بناء على أسئلة البحث أعلاه، فإننا نفترض أن، أولاً: إذا كانت أهداف المناهج الدراسية موازية بشكل جيد مع مايقوم به المدربون في مواد التعليم في الواقع، و ما يرغب فيه الطلاب للتعلم وكذلك ما يتم اختباره، فسوف تكون الآثار قوية و إيجابية. ثانياً: إذا لم يكن هناك تطابق أو موازنة بين هذه المكونات بشكل جيد ، فإن الآثار الانعكاسية تكون غير فعالة أو سلبية. ووفقاً لمنهجية وصف اثنين من الاستبيانات الموجهة إلى كل من المعلمين والطلاب، والنتائج الختامية لدينا. وجدنا أن المستويات الرئيسية (المعلمين والطلاب والمناهج والمدارس) تؤثر على بعضها البعض في نفس الوقت. ويمكن لهذه التأثيرات أن تكون إيجابية و / أو سلبية. إذن نستنتج أن هناك حاجة قوية لتحقيق التوازن والتوفيق بين هذه المستويات، والسيطرة على مثل هذه العوامل (الثقة بالنفس، والقلق، والإجهاد، الخبرة في مجال التدريس واستراتيجيات التعلم) لأن كل هذه العوامل تزيد في قوة الآثار الانعكاسية للاختبارات أو تحد منها.

Chapter Two : A Probe on Standardized Tests

Introduction

This chapter is devoted to probe the nature of standardized testing ; in which we try to investigate the notion of standardized tests , their types and characteristics. This chapter consists of four main sections. The first section tries to explore the terms of evaluation, assessment and test, specially the differences between these terms, and what is special about test. The second section tries to investigate the nature of standardized tests by discussing the different views and / or definitions of the researchers after laying them out in order to identify the best and the more practical among these views. The third one tries to identify the different types, such as: norm-referenced tests and criterion- referenced test, their nature and what differs one from the other. Finally, the fourth section identifies the high-qualities of a good test, especially of standardized test, which are validity, reliability and practicality. In addition of the conditions that promote these qualities, and the factors that influence them negatively.

II.1.Differences Between Evaluation, Assessment, and Test

The terms evaluation, assessment, and test are often confused and used interchangeably. However, they do not mean the same thing.

According to McMillan, evaluation is “the making of judgments about quality –how good the behavior or performance is” (McMillan, 2007:10).He explains that evaluation “involves an interpretation of what has been gathered through measurement, in which value judgments are made about performance” (*ibid.*). But, according to Bachman, evaluation is defined as “the systematic gathering of information for the purpose of making decisions “ (Bachman ,1990:22).

However, assessment is “the collection, evaluation, and use of information to help teachers make decisions that improve student learning” (McMillan, 2007:8). Assessment is also defined as “the process of gathering ,evaluating , and using information “ (*ibid.*:430). According to Seifert and Sutton, “assessment is an integrated process of gaining information about students’ learning and making value judgments about their progress” (2009:232). They add that “ information about students’ progress can be obtained from a variety of sources including projects , portfolios, performances, observations and tests” (*ibid.*).

Whereas testing “refers to the specific procedures that teachers and examiners employ to try to measure ability in the language , using what learners show they know as an indicator of their ability” (Hedge, 2000:378). Also, McMillan defines a test as “a formal, systematic procedure for assessment in which students respond to a standard set of questions” . (2007:434). From Bachman’s view, a test is“ a measurement instrument designed to elicit a specific sample of an individual’s behavior ”(1990:20),but from Brown’s view, it is “a method of measuring a person’s ability ,knowledge, or performance in a given domain” (2004:3).

From all these definitions, we conclude that testing is just one part of assessment, but assessment and evaluation are more general and more global processes because we can evaluate teaching, teaching materials, and even tests, as well as learning. In addition, evaluation is a sort of interpretation of what has been gathered through measurement, but assessment is an ongoing process because the teacher assesses his student whenever he responds to a question, offers a comment, or tries out a new word or structure.

Moreover, Brown’s definition highlights more the main components of a good test which are:

*First, it is a method: in the sense that it encloses techniques , procedures and activities and requires performance on the part of the testee , or the tester, or both.

*Second, it is supposed to measure :in order to formulate a judgement .

*Third, it tests a person whose previous acquisitions should be known; testers need to understand who the test –takers are ?what is their previous experience and background?

*Fourth, it measures ability and knowledge: the testee produces a performance which is interpreted by the tester in terms of competence.

*Finally, it measures a given domain “desired area”: it tests what is supposed to test and not something else. This view leads to assert that constructing a good test is a complex task and, in fact, it is.

II.2.The Nature of Standardized Tests

In this section, we try to find out what is a standardized tests .Brown defines this type of testing as “an assessment instrument for which there are uniform procedures for administration, design, scoring, and reporting”(Brown,2004 :104).Moreover,a standardized test is largely defined and more detailed by Shiel, Kellaghan and Moran :

As a procedure designed to assess the abilities, knowledge, or skills of individuals under clearly specified and controlled conditions relating to construction, administration, and scoring, to provide scores that derive their meaning from an interpretative framework that is provided with the test (2010 :22)

However, a standardized test is simply defined by Oosterhof as “a test designed to be administered consistently across a variety of settings” (2003 :240) . Nearly similar to what is said, Seifert and Sutton claim that standardized test are type of tests “developed by a team of experts and are administered in standard ways” (2009:289).

From Wikipedia, “standardized tests are designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent and are administered and scored in a predetermined, standard manner.”

In other words, Brown, in his book ‘Teaching by Principles’, states that “standardized tests are almost by definition highly practical, reliable instruments. They are designed to minimize time and money on the part of test designer and test-taker, and to be painstakingly accurate in their scoring” (2007:480)

We conclude that the test is named standardized test when it is administered under uniform conditions and graded according to a fixed set of rules, such as:

- The same questions and / or tasks are required of each test-taker.
- The same information is provided to each test- taker immediately before and during the test.
- Each test-taker has the same amount of time allowed to take the test.
- Responses are scored in the same way.

What is important to standardized testing is whether all students are asked equivalent questions, under equivalent circumstances, and graded equally. In standardized test, if a given answer is correct for one student, it is correct for all students. Graders do not accept an answer as good enough for one student but reject the same answer as inadequate for another student.

II.3.Types of Standardized Tests:

Test specialists have identified two types of standardized tests: Norm-Referenced Tests (N R T) and Criterion –Referenced Tests (C R T).

II.3.1. Norm-Referenced Tests

According to McMillan, this type of tests allows “to compare performance to a well-defined norming or reference group and to determine relative strengths and weaknesses of students”. (Mc Millan,2007: 408). In the same sense, McNamara says:“Norm-Referenced measurement adopts a framework of comparison between individuals for understanding the significance of any single score” (2000:62). He explains more by saying “each score is seen in the light of other scores”(ibid.).

Brown views that “each test- taker’s score is interpreted in relation to a mean (average score), median (middle score), standard deviation (extent of variance in scores), and/or percentile rank”. (2004: 7). This means that such tests purpose to place test-takers along a mathematical continuum in rank order.

Similar to what is mentioned above by McMillan and McNamara , Oosterhof argues that norm-referenced interpretation “is provided by comparing the student’s performance with the performance of others” (2003:11).Nearly to that, Seifert and Sutton argue that “ norm-referenced standardized tests report students’ performance relative to others” . (2009:267).

We notice, from all what is said, that there is an agreement between the researchers that norm-referenced testing is used to compare a test-taker’s results to the results of a reference group that has taken the same test. This means that the scores are ranked and compared to each other, for example, the tests organized to hire new teachers in middle and high schools or in Magister contests. Candidates, in this type of tests, are chosen from the top list no matter what they score and in accordance with the number required to hire.

II.3.2. Criterion-Referenced Tests

This type of tests allows to compare student performance to established standards rather than to other students. However, from Brown’s view, criterion-referenced tests “are designed to give tests-takers feedback, usually in the form of grades, on specific course or lesson objectives” (2004: 7).

McNamara explains that more by saying: “An alternative approach which does not use a comparison between individuals as its frame of reference is known as criterion-referenced measurement” (2000:64). He argues that, in this type, “Individual performances

are evaluated against a verbal description of a satisfactory performance at a given level” (McNamara, 2000: 64). Therefore, there is no a reference group, but a preference level.

According to Oosterhof, criterion-referenced interpretation “is provided by describing what the student can and cannot do” (2003:11). In the same way, Seifert and Sutton claim that “criterion –referenced standardized tests measure student performance against a specific standard or criterion ...and provide information about what students can and cannot do” (2009:267). This leads to say that those views agree that this type of testing measures student performance according to specific standards not specific groups.

We conclude that criterion-referenced testing is used to measure a level of mastery according to a specific set of performance standards. Now, we know that this is the type in which performance is compared to levels of established criteria, or judged according to certain criteria, for example, the Baccalaureate examination.

Hughes explains more the purpose of criterion-referenced test as “to classify people according to whether or not they are able to perform some task or set of tasks satisfactorily”(2003:21). Whereas, norm-referenced test relates one candidate’s performance to that of other candidates.

Further, Seifert and Sutton (2009:267) add something new and very important when they claimed that standardized tests can incorporate both criterion-referenced and norm-referenced elements into the same test by providing not only information on mastery of a content standard but also the percentage of students who attained that level of mastery.

To sum up, norm-referenced testing is based on comparing a student’s score to that of other students. Whereas, criterion-referenced testing is based on how well the student has mastered the content of the test. This means that criterion-referenced tests are specifically designed to assess how much of the content in a course or programme is being learned by the students. For that, these kinds of tests are good measures of student’s strengths and weaknesses considering the goals and objectives of a particular course or programme. This is why they are most useful to classroom teachers and curriculum developers.

We can also claim that when the principal interest is in ranking all learners, norm-referenced tests are preferred, and when the issue is whether the learner has met a

particular standard, criterion-referenced tests are more appropriate, and of course if it is needed to incorporate between the two types, it will be better.

II.4. Characteristics of Good Tests

In order to judge the effectiveness of any test, especially standardized test, it is sensible to lay down some criteria, such as: Validity, Reliability, and Practicality.

II.4.1- Validity

Validity is the most complex criterion and the most important of a good test. According to Harmer, “a test is valid if it tests what is supposed to test” (2001:322). In the same way, Hughes writes that “a test is said to be valid if it measures accurately what it is intended to measure” (2003:26). Also, Oosterhof claims that validity refers “to the degree to which a test measures what it is supposed to measure” (2003:34). Nearly to that, McMillan defines validity as “the extent to which a test measures what it is supposed to measure” (2007:64). This means that validity refers to the appropriateness of the inferences that result from the assessment.

To support this view, Seifert and Sutton define validity as “the evaluation of the adequacy and appropriateness of the interpretations and use of assessment results for a given group of individuals” (2009:234). In addition Brown refers this term to “the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment”(2004:22). Moreover, validity-according to Shiel, Kellaghan and Moran- is “the interpretation of test scores required by proposed uses that are evaluated, not the test itself” (2010:27). In the same sense, Bachman refers this term to “the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores” (1990:243). He explains: “the process of validation, therefore, starts with the inferences that are drawn and the use that are made of scores ...” (*ibid.*).

We come to notice that all these views refer the term of validity: to the degree to which a test measures what is intended to measure. This presupposes that when we write a test we have an intention to measure something specific, and that validity concerns finding out whether this test actually measures what is intended. For example, it is not valid to test writing ability with an essay question that requires specialist knowledge of history or biology unless it is known that all students share this knowledge before they do the test.

Also, it is important to understand that validity refers to the interpretations and uses made of the results of an assessment procedure not of the assessment procedure itself. As what is mentioned in the above definitions, the criterion of validity involves making an overall judgment of the degree to which the interpretations and uses of the assessment results are justified. Finally, we discover that validity is a matter of “degree” (high, moderate, low); not all or none.

II.4.2. Reliability

As an essential characteristic of good test, reliability is defined by a big number of researchers as follows:

According to Brown ,“a reliable test is consistent and dependable”(2004:20). This means that if we give the same test to the same student or matched students on two different occasions, we obtain the same results. In the same way , Harmer demonstrates that “a good test should give consistent results”(2001:322), and he explains more by saying : “If the same group of students took the same test twice within two days- without reflecting on the first test before they sat it again- they should get the same results on each occasion”(*ibid.*). Davies and Pearse simply state that “Reliability is a matter of how far we can believe or trust the results of a test’”(2000:173).

In other words, reliability- according to McMillan-“is concerned with the consistency, stability, and dependability of the scores” (2007:69).The same thing declared by Seifert and Sutton when they said that “reliability refers to the consistency of the measurement”(2009:236). Shiel, Kellaghan and Moran add that this criterion“refers to consistency of the measurement when a testing procedure is repeated on a population of individuals or groups”(2010:31). In different way, the more similar the scores would have been , the more reliable the test is said to be.

We conclude that reliability concerns the stability or consistency of the scores; which means that the same test- taker would get the same results in the test if taken in different occasions.

However, the researchers (like: McMillan , Hughes, Seifert and Sutton) strongly believe that there are factors that influence reliability , such as:

- *The greater the number of items , the greater the reliability ,
- *The best reliability is obtained when items are not too easy or too hard.

- *The higher the number of students ,the stronger the reliability.
- *The more objective the scoring ,the greater the reliability.
- *Unclear directions and tasks lead to poor reliability.
- *The greater the differences between one administration of a test and another ,the greater the differences between a candidate’s performance on the two occasions.

II.4.3.Practicality

One of the desirable qualities of good tests after Validity and reliability is that of practicality , which is must be balanced with previously mentioned criteria.

According to Brown, “practicality is determined by the teacher’ s (and the students’) time constraints, costs, and administrative details , and to some extent by what occurs before and after the test”(2004:31).

From Brown’s view , a test is practical when it requires an appropriate time and cost, and when it is relatively easy to administer and to score . In the same sense, McMillan argues that tests “need to take into consideration the teacher’s familiarity with the method, the time required , the complexity of administration, the ease of scoring and interpretation, and the cost...”(2007:89).

Not far from these thoughts, Hughes states: “We should not forget that testing costs time and money ...”(2003:56). This means that a test should be easy and cheap to construct, administer, score and interpret.

From these views , we find out that practicality covers a wide range of issues , such as :

- *The time required :which means how long it takes to construct the test ,how much time is needed for students to provide answers, and how long it takes to score the results.
- *Ease of administration: when the direction and the procedures are clear.
- *Ease of scoring :for example ,objective tests are easy to score , but performance tests and essays are difficult to score because more time is needed to ensure reliability.
- *Ease of interpretation :interpretation is easier if you are able to plan, before the test, how to use the results.
- *Cost :it is best to use the most appropriate and economical test ,but it would be unwise to use a more unreliable or invalid test just because it costs less.

In short, a test should not require high financial support, long period of time to cover and to correct , sophisticated aids and large human resources.

Conclusion

In this chapter, we have discussed differences between evaluation, assessment and test, the nature of standardized tests, their types and their high qualities or characteristics. The above studies lead to determine clearly and carefully the differences between evaluation, assessment and test, the nature of standardized tests; their fixed standards and their main types, such as: norm-referenced tests and criterion-referenced tests; this includes their definitions , differences , and what makes one type more important and appropriate than the other in certain circumstances. Finally, we have discussed the main characteristics of standardized tests in a clear and detailed manner, in order to help both teachers and students to construct, administer, interpret, score and pass these tests successfully. Thus, we understand that validity refers to the degree to which a test measures what it is intended to measure. However, reliability concerns the stability and the consistency of the scores, while practicality means that a test should be cheap and easy to construct, administer, score and interpret.

ChapterOne :Investigating the Concept of Washback

Introduction

In this chapter, which shapes and guides the whole research, we try to investigate one of the most famous and complicated phenomena in language testing which is the washback concept. This chapter consists of three main sections. The first section lays out the origin and the different facets of this concept and how it derives by identifying various terms referred to this concept and showing the large quantitative and qualitative studies that discussed chronologically this phenomenon. The second section tries to investigate its complex nature and vague scope by discussing and analysing the huge number of definitions after exposing them. Thus, we notice that there is a sort of agreement between the researchers that the concept and the study of the washback phenomenon is much more important and complex at the same time. The third one identifies and explores its two types, which are the positive and the negative washback, what makes differences between them and how to promote the positive and inhibit the negative.

I.1. The Origin of Washback

The washback concept is largely discussed by a big number of researchers. The term washback, backwash, test impact, consequential validity and other several terms are referred to the same phenomenon. This what is obtained after a long and careful investigation and which is argued by the following:

- Washback or backwash (Alderson and Wall, 1993)
- Washback (Mc Namara, 2000; Cheng, 2005; Pinter, 2006; Fulcher and Davidson, 2007)
- Test impact (Mc Namara, 2000; Cheng, 2005)
- Consequential validity (Messick, 1996; Mc Namara, 2000).

In addition, Alderson and Wall asserted that the notion of washback or backwash is common in the educational and applied linguistics literature, but backwash is the older term (1993:2). Thus, we think that the term washback is preferred in the most of studies and also in this study.

Moreover, Cheng claims that “the concept and study of washback has also been derived from recent developments in language testing and measurement-driven reform in the areas of general educational assessment” (2005:25) and he states: “Perhaps the single most important theoretical development in language testing since the 1980s, was the realization that a language test score represents a complexity of multiple influences” (2005:25).

Before that, Wiseman (1961) says that “good examinations are useful and desirable: without them education would be poorer and much less effective” (Alderson and Wall,1993:3). Also, Morris (1972) “considers examinations necessary to ensure that the curriculum is put into effect” (*ibid*: 3). Davies (1985) “asks whether tests should necessary follow the curriculum, and suggests that perhaps tests ought to lead and influence curriculum”’(Cheng, 2005:26).

However, Alderson (1986) turns the attention to an additional area called washback effects and comes up with his famous declaration: innovations in language curriculum through innovations in language testing (Cheng, 2005:25-26).

Fulcher and Davidson stated that :

Alderson and Wall were the first to critically investigate the concept of washback . Before 1993 it had merely been assumed that (1) tests did have an effect on teaching , and (2) the effect was largely negative. By framing the ‘washback hypothesis’, Alderson and Wall made it possible for washback to be studied empirically , and the simplistic nature of the original concept was soon turned into a conceptually rich source of theory and research (2007:222).

More than that , Morrow (1986) coined the term ‘washback validity’ to describe the relationship between testing, teaching and learning , and he asserted that “this test is valid when it has good washback and ...this test is invalid when it has negative washback” (Alderson and Wall,1993:4) .However, this form of validity, like Alderson says ,has never been demonstrated because of the existence of other forces within society, education and schools that might prevent washback from occurring, and we think that is the truth.

Messick (1989,1992,1994,1996) has placed washback effect within a broader concept of construct validity (consequential validity).(Cheng, 2005:26).

We notice that all those studies stress the notion that tests should drive teaching and enhance learning. In order to achieve this, it should be a match between the test’s content and format, and the curriculum’s content and format. Thus, the notion of curriculum alignment intervenes, as lot of researchers claim .

We think that we need more than that, we need a total alignment between the curriculum, syllabus, students objectives and tests. Moreover, we need for that many more studies before we can assert that we understand the nature and mechanisms of washback.

We conclude that washback is an educational phenomenon derived from a big number of studies into the relationship between teaching, learning and testing and into the relationship between the curriculum, syllabus, students objectives and tests. Further, those studies demonstrate to us that washback ,in language testing, is a powerful concept that all the main participants, like tests designers and classroom teachers, should seriously pay attention to it.

I.2. Nature and the Scope of Washback.

For many years, It was assumed that good tests would produce good Washback and bad tests would produce bad washback until Alderson and Wall come with their ideas that "any test, good or bad, can be said to be having beneficial washback if it increases activity or motivation ..." (1993: 6).In addition, they argue that there are other forces exist within society, education and schools that might prevent washback from occurring. Also, they declare many times that this phenomenon is much more complex than it appears .

Thus, we try to identify this phenomenon carefully by exposing its various and important definitions then we discuss them later on.

Some researchers agree that “washback is a term commonly used in language testing, yet it is rarely found in dictionaries. However, the word backwash can be found in certain dictionaries and is defined as the unwelcome repercussions of some social action, and-unpleasant after- effects of an event or situation” (Cheng, 2005:27). Also, Hughes writes:“where washback came from I do not know, what I do know is that I can find backwash in dictionaries, but not washback”(2003: 57).

For Alderson and Wall, “the notion of washback or backwash- the influence of test on teaching-is commonplace in the educational and applied linguistics literature”(1993:2), but backwash is the older term. According to them this notion is referred to as ' backwash ' in general educational circles, but it is known as ‘washback’ in British applied linguistics, and thus, there is no reason for preferring either term.

This notion refers to the extent to which a test influences language teachers and learners to do things they would not necessarily otherwise do because of the test. According to Biggs (1995), the term backwash refers “to the fact that testing drives not only the curriculum, but teaching methods and students’ approaches to learning”(Cheng, 2005:27). Messick (1996) points out that washback “refers to the extent to which the introduction and the use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning” (*ibid.*).

McNamara simply defines washback by “the influence that testing has on teaching” (2000:72), also by “the effect of tests on teaching and learning”(2000:73), or “the effect of a test on teaching and learning leading up to it”(2000:138). He adds, in order to distinguish between ‘washback’ and ‘Impact’, that “tests can also have effects beyond the classroom. The wider effect of tests on the community as a whole, including the school, is referred to as test impact” (2000:74). Thus, he refers ‘test impact’ to the total effect of a test on the educational process and on the wider community.

In the same way, Hughes defines backwash by “ the effect that tests have on learning and teaching” (2003:53), and he argues that the term impact “ is not limited to the effects of assessment on learning and teaching but extends to the way in which assessment affects society as a whole” (2003:2). Therefore, we notice that there is a small difference between washback and test impact; the latter is wider than the former. While impact deals with wider influences and includes the ‘macro contexts’ tests and examinations in society, washback is an aspect of impact related to the ‘micro contexts’ of the classroom and the school.

According to Fulcher and Davidson, “the term washback is used to talk about the effect that tests have on what goes on in the classroom” (2007:74). They add “washback, sometimes referred to as backwash, the effect of a test learning and teaching. Washback studies focus on practices or behaviour that would not be present if it were not for the test”(2007:377). However, impact is “a term used to denote the effect of the use of a test on society, institutions, stakeholders and individual test takers. It may also include the effect of the test on the classroom, but this is more usually referred to by the more specific term washback”(*ibid* : 372).

From the above definitions, we conclude that washback is generally referred to the effects that tests have on teaching, this means that the existence of tests bring about some

changes in behaviors, attitudes and motivation of teachers, learners and parents. For that, some researchers argue that “if the concept of washback is to have any meaning, it is necessary to identify what changes in learning or teaching can directly attributed to the use of the test in that context” (Fulcher and Davidson , 2007 : 221). In addition, we discover that this phenomenon is really much more complex matter because of the existences of many forces and variables that might intervene and influence, both positively and negatively, teaching and learning. Thus, we should know how to promote the former and inhibit the latter.

Alderson and Wall set out some of those variables that might affect the washback of a test; they called them ‘washback hypotheses’:(1993:120-121)

- *A test will influence teaching
- *A test will influence learning
- *A test will influence what teachers teach; and
- *A test will influence how teachers teach
- *A test will influence what learners learn
- *A test will influence how learners learn
- *A test will influence the rate and sequence of teaching; and
- *A test will influence the rate and sequence of learning
- *A test will influence the degree and depth of teaching
- * A test will influence the degree and depth of learning
- *A test will influence attitude towards the content, method, etc., of teaching and learning
- *Tests that have important consequences will have washback; conversely,
- *Tests that do not have important consequences will have no washback
- *Tests will have washback on all learners and teachers
- *Tests will have washback effects for some learners and some teachers, but not for others.

The above study reveals the effects of tests on various aspects of the classroom such as: curriculum (by narrowing it to those areas most likely to be tested), materials, teaching methods (approaches or techniques), feelings and attitudes (anxiety and fear of test results among teachers and students), and learning.

I.3. Types of Washback

Most of researchers affirm that tests can have, both positive and negative, influence on teaching and learning. We try, in this section , to investigate those two types , know their nature and how to promote the positive and inhibit the negative.

I.3.1.Negative Washback

It is assumed that test influence teaching and learning .This influence is totally seen as negative by some researchers, and it is referred to the so- called negative washback. For example, Vernon (1956) claimed that examinations “distort the curriculum” because teachers tended to neglect all “subjects and activities which did not contribute directly to passing the exam”(Alderson and Wall,1993:115). For that, Fulcher and Davidson demonstrate that before 1993 it had been assumed that the effect of tests was largely negative (2007:222), and thus, Alderson writes that washback“is a hugely complex matter, and very far from being a simple case of tests having negative impact on teaching” (*ibid* :228).

Negative washback, according to Alderson and Wall, refers as “the negative or undesirable effect on teaching and learning of a particular... test” (1993 :5). Moreover, Davies indicates that “all too often the washback effect has been bad ;designed as testing devices, examinations have become teaching devices...”(Cheng,2005 :29). Noble and Smith add “that high- stakes testing affected teachers directly and negatively...” (*ibid.*).

Supporting the above ideas, Smith pointes out that “testing programs substantially reduce the time available for instruction, narrow curricular offerings and modes of instruction, and potentially reduce the capacities of teachers to teach content and to use methods and materials that are incompatible with standardized testing formats” (*ibid.*). Wiseman, also, argues “that those coaching classes were not a good use of time because the students were practising exam techniques rather than language learning activities” (Cheng,2005 :30).

We conclude that tests can have strong negative effects such as:

- *Narrowing the curriculum (restriction of content).
- *Teachers ignore subjects and activities that are not directly related to passing the examination.
- *Teachers are under anxiety, fear and pressure to cover the material.

*Increasing number of paid coaching classes to prepare students for examination.

*Learning test-taking skills rather than language learning activities.

To maintain some of those point of views, under certain circumstances like the fear of poor results, shame or embarrassment, teachers look for whatever way seems possible to achieve high scores for their students. This might lead to ‘teaching to the test’ with an undesirable ‘narrowing of the curriculum’.

I.3.2. Positive Washback

Other researchers, however, see that the influence of tests on teaching and learning is more positive and they declare that changing examinations can bring about beneficial change in teaching and learning. This sort of change refers to so- called positive washback. For example, Cheng (2005) titles one of his famous books ‘changing language teaching through language testing’. While Alderson (1986) calls , in his book, for ‘innovations in language curriculum through innovations in language testing’.

According to Alderson and Wall (1993), the positive washback “refers to tests and examinations that influence teaching and learning beneficially” (Cheng, 2005 :30). Morris (1972) also “considers examinations necessary to ensure that the curriculum is put into effect” (Alderson and Wall 1993 :3).A good test ,according to Davies(1985) , should be “an obedient servant of teaching ,and this is especially true in the case of achievement testing” (Cheng,2005 :30). He adds that “creative and innovative testing ...can, quite successfully, attract to itself a syllabus change or a new syllabus which effectively makes it into an achievement test” (*ibid.*).

In order to promote a positive washback, teachers have to comment generously and specifically on their students’ test performance, give praise for strengths as well as constructive criticism of weaknesses, ask their students to use test results as a guide to setting goals for their future effort and help them to discuss the give feedback and evaluation and to discover their own mistakes.

According to Bachman, “positive washback would result when the testing procedure reflects the skills and abilities that are taught in the course” (1990:283) .Cohen adds “that if students have a chance to practice types of tests and test items over time, their performance on such tests will improve ...” (1980:60). In addition of that, Pinter advices “the more teachers know about the classroom complexities and the learners they are

working with, the better their chance for success ” (2006:152). In order to know how successful or unsuccessful their teaching is, teachers must assess their students. For that Davies states “Assessing pupil outcomes is also vitally important as it informs good lesson preparation” (2006: 167).

To sum up, the washback effect of any test will be negative when it fails to reflect the learning principles and /or course objectives. However, it will be more positive when it encourages more the desired changes, for example, in textbooks or curricula, and when it motivates teachers and students to fulfil their teaching and learning goals. Moreover, Alderson and Wall (1993) come with the new idea that the quality of the washback effect might be independent of the quality of test. Which means that any, test good or bad, can have positive or negative washback. This is because of certain forces that exist within society, education, and schools that might intervene and affect the nature of washback .

Conclusion

In this chapter, the origin of washback, the nature and the scope, its types and function have been discussed. Washback, as phenomenon used in language testing, refers to the impact and power of testing on teaching and learning in schools. However, whatever the impact of testing is positive or negative it needs to be more studied further. The above studies help to determine the nature of washback and how this phenomenon works in education , to show how different levels (schools, teachers, learners, and parents) react to the desired changes, to determine possible areas of washback intensity in teaching and learning in high schools and to define the interrelationship between who changes , what , how, when, where and why.

Dedication

To my love ALLAH and the PROPHET MOHAMED peace be upon him.

To Mother and Father whose patience is beyond limit.

Thank you so much for your care,
help and support.

To my beloved sisters and brothers.

To Fatima Bdirina ,Faiza , Meriem,
Amina, Hanan and Imy.

To Yamma, Amel, Soad ,Saida, Asia, Houria, Sakina, Zeineb,
Radia,Zahra and Fairouz.

To Shems, Najwa, Asma, Rania ,Hanan,
Souhir,Wafa and Anfel

To my relatives, colleagues, and friends.

Acknowledgements

First of all, thanks to ALLAH the helpful and the merciful .

I would like to thank my supervisor Mr. BACHER AHMED for his help, his kind supervision, constructive comments and sincere pieces of advice....

Thanks to Mr. CHALLA and all members of English Branch : Mrs. BOUDIAF , Mrs. CHELLI , Mrs . HASSINA, and Mrs. RABEHI.

I would like to thank Mrs. SOAD BOUKHARI for her extensive most helpful comments and pieces of advice. Thanks also to Mr. MOHAMED BOUKHARI , Mr. DJAAFAR HAMZA, Mr. BERKEM MOHAMED TAHER, Mrs.YAMNA and Mrs. NAIMA CHALLA for their greatest help.

Thanks also to my brother BILAL for his immense help and support.

Thanks to my helpful inspector of Islamic Sciences, my best friends and colleagues , especially teachers of Islamic Sciences ,and all members of Smatti Mohamed El Abed high school(with special thanks to the Director)for their help and support

I am indebted to my teachers all along my studies for their kindness patience and encouragements, especially Miss. ANGAR , Miss DOUNIA , Mr. TOURKI, Mr. MEHERI, Mr. SIGNI, and Mr. BOULEGROUN

Thanks to teachers(TOBNI ABD ELRAHMEN and NAJET, NAIMA,ILHEM , DALILA, FATIHA and NOUR ELHOUDA) and students(scintific classes). Who collaborate with me in answering the questionnaires, with a great patience .

To all the students who will pass the Baccalaureate next month, I wish you a good luck

Finally, special thanks are due to all my family members, especially my parents, for their immense care, help, patience, encouragements, spiritual and financial support.

General Conclusion

The current study served as an attempt to investigate the washback effects on students' performance on standardized tests. Taking the Baccalaureate examination as an example of such tests, the main objectives of this research are : to explore the nature and the scope of washback effects on teachers and students , identify areas of washback intensity in teaching and learning , and understand how the main participants react to changes made in the BAC examination .

To achieve the above aims,it is necessary to include specific research questions, such as : what are the perceptions of teachers and students related to the effects of the BAC examination ? What do instructors think about the changes to the examination that might be useful? What are the learning and teaching practices employed by teachers and their students during the preparation processes? What is the nature and the scope of washback effect on teachers and students? Based on these questions, we hypothesize that: Tests can affect, both positively and negatively, what happens in classrooms .

To answer the research questions and to judge whether what we hypothesize is rejected or accepted, both teachers and students are asked to answer qualitative questions about the subject area. After collecting and analysing data, this study concludes by highlighting that testing, specially standardized testing exerts stronger effects on schools, teaching and learning processes and curricula. Moreover, these effects can be both positive and negative.

Our findings are divided into three level effects .Firstly, teachers' level effects, we find out that :

-Teachers, in fact, perceive strong changes in the BAC examination, and they react positively to these changes. Consequently, they make some changes in their teaching (like using more Communicative Approach).

-Teachers are under pressures such as, covering all the curriculum required, devoting appropriate time to test preparation activities, and improving the quality of students' learning .

-The more teacher talk dominate the class, the more lessons come to be bored .Thus, the more reducing students' interest .

-teaching an exam class is easier than teaching a non-exam class because the learners are more motivated.

Secondly, students' level effects such as :

- The more using motivational strategies (like Audio-Visual Aids), the more improving students' learning .
- The over number of students in class influences negatively learning and teaching processes .
- The lack of using audio-visual aids, organizing group work and discussion and using language games/ riddles make students not involved and not interested.
- Students react easily, both positively or negatively, under certain conditions (such as: motivation, large classes, mixed ability and anxiety).Thus, teachers should be aware of such conditions to increase positive washback effects and decrease negative ones.

Thirdly, curricula and schools' level effects such as:

- The more devoting time to the test preparation activities, the more narrowing curriculum required and reducing students' higher –order- thinking abilities
- Increasing time spent on subject areas that are tested and less time on those that are not.
- Curricula change to improve students' test scores.
- Schools' principals can intervene and influence negatively students' scores; under reputation for example, they are more interested in increasing test scores than in improving overall students learning.

Finally, the most important finding is that to align well curricular goals, with what instructors teach, what students wish to learn, and what is tested to promote positive washback effects and prevent negatives ones. Moreover, we cannot affirm that good tests lead to positive effects because of the existence of a huge number of variables (like: motivation, anxiety, teaching experience and learning strategies) that might increase and /or decrease the effects of tests, and they are so difficult to control them.

General Introduction

Testing has always a crucial aspect of language teaching and learning in expanding the knowledge and the progress of the students' achievement and growth. It is commonly asserted that tests affect teachers and learners and thereby affect teaching and learning; teachers and students work harder when they are approaching exams than when they do not. Also tests are thought to be a major determinant of course designs and classroom practices. Supporting these ideas, Wong writes: "The examination dictates the activities in schools... Interpretation of the syllabus is carried out chiefly by reference to past examination papers which ... tend to carry questions similar in type and content year after year"(Alderson and Wall , 1993: 2). Thus, our study is specifically for investigating Washback effects on the performance of students in standardized Tests.

Background of the Study

This current study is designed to examine one of major public standardized tests in Algeria, such as the Baccalaureate Examination, which affect a huge number of students, their teachers and families every year. This examination is administered once a year by the National Ministry of Education, its main objective is to select and place the students with a high academic potential. By investigating the student and instructor perceptions related to the Baccalaureate Examinations, this study aimed at shedding light on students' study practices for the exam during the preparation processes, and the effects of this exam on teaching and learning.

***The Origin of the “ Baccalaureate”:**

The term “ Baccalaureate” means the last secondary school examination in France.(Oxford advanced learners' dictionary).

According to Elatia, the Baccalaureate “is a comprehensive exam given originally in France” (History of the Baccalaureate). She dates back this exam to 1808 when Napoleon Bonaparte was looking for an exam that will enable the empire to select civil servants. Since that time, the Baccalaureate spread in many countries around the world, mostly francophone countries that were ex-colonies of France. Along this time, the Baccalaureate underwent several changes and is still undergoing change. The same researcher adds that, in the beginning, the Baccalaureate was an oral exam, candidates were tested in several

languages like, Latin, Greek and other important languages of Europe such as: English, German and Italian.

In the modern Baccalaureate, languages are still very important in all the specialties, whether it is mathematical sciences, natural sciences, or literature, there is a foreign language examination.

In the case of Algeria education, the government has introduced English language as a compulsory subject at young ages. So that the second foreign language after the French language, since it becomes the most global language.

Purpose of the Study

The main purposes of this study are to :

- 1-Study the phenomena of the washback effects on teaching and learning .
- 2-Explore the nature and the scope of the washback effect on aspects of teachers' and students' perceptions and teachers' behaviours .
- 3-Identify areas of washback intensity in teaching and learning
- 4-Understand how the main participants reacted to changes made in the Baccalaureate Examinations.

Significance of the Study

To my best knowledge, this investigation is the first study of washback in Ouled-Djellal high schools .It is one of the few washback studies that provides hard data and evidence of the washback effects in a specific educational context , and contribute to the general understanding of washback in education .

Main Questions

To achieve the above mentioned aims , the following main research questions were addressed :

- 1- What are the learning and teaching practices employed by students and their teachers during the preparation processes ?
- 2-What are the perceptions of the instructors and the students related to the effects of the Baccalaureate Examination?
- 3-What do instructors think about the changes to the examination that might be useful ?

4-What is the nature and the scope of washback effects on teachers' and students' perceptions of various aspects of teaching towards the Examination?

Research Hypotheses

Based on the research questions, it was hypothesized that: Tests can affect, both positively and negatively, what happens in classrooms.

To be precise and concise, two major hypotheses were proposed as follows:

- 1- If curricular goals align well with the material instructors actually teach, and what students actually wish to learn as well as what is tested, washback will tend to be strong.
- 2- If test content does not match well with these components, the washback is apt to be either ineffectual or negative.

As teachers, we have a strong belief that we can promote a positive washback from standardized tests if we align well between these four components: [what we design, what we teach, what our students wish to learn and what is being tested].

Methodology and Research Tools

Based on descriptive methodology, this study relies on two questionnaires:

- 1-Teachers' Questionnaire: It consists of twenty questions directed to the active members: Teachers of English, who teach at Smatti high school in Ouled-Djellal.
- 2-Students' Questionnaire: It consists, also, of twenty questions directed to the passive members: 3rd year scientific classes, who study at the same school.

The data gathered will be analysed and discussed through tables and graphs followed by comments for each

* Population:- 3rd year high school students.

*Sample:-Scientific classes.

Research Limitations

To perform this work there were big difficulties. We will state some of them:

- * The first thing is the large number of students in 3rd year (14 classes \approx 517 students). So, we took just a sample (scientific classes). They are our classes.
- * The second is that we have to take the distance between Biskra and Ouled-Djellal many times per week, if not every day.

* The third is the great effort that we have to exert with nine (09) classes, seven (07) of them have to practise and prepare to pass the Baccaureate Examination.

*The fourth is the lack of main sources (books) of study, especially Washback sources.

Structure of the Study

Our study has been structured into three chapters: The first chapter is devoted to the investigation of the concept of Washback by laying out the origin of this concept, identifying its complex nature, and exploring its types. The second is devoted to probe the nature of Standardized Tests by investigating the differences between evaluation, assessment and test, exploring the notion of Standardized Tests, discussing their types and their high-qualities. The third and the last chapter is devoted to the analysis of teachers' and students' questionnaires of the case study of third year scientific classes at high school of Smatti Mohamed El Abed in Ouled-Djellal.

Omnibus Recommendations

After a careful analysis of the two questionnaires, we come to find out some of recommendations such as :

1- Teachers assert that there is a need of matching well between curricular goals, what instructors teach, what students wish to learn and what is tested in order to promote the positive effects of tests and reduce the negative ones.

2- Teachers should not devote a large time to test preparation activities because this leads to narrow curriculum and reduce Higher-Order-Thinking of students abilities.

3 - Teachers need to make sure that their students have good test- taking skills.

4- Teachers should attract their students' attention by Audio-Visual Aids, organizing group work or discussion, and language games / riddles in order to be involved and interested.

5- Teachers should be responsible and thoughtful vis-a-vis their students.

6-Instructors should control factors such as (stress, anxiety, mixed ability, teaching experience, and learning strategies) that intervene and lead to negative effects .

7-Students need a careful direction of motivation to the positive side.

8- Students need a strong will to balance, as possible, between different modules in coefficient, time, and program length .

9-Students ought to understand that the role of teacher is to guide , to instruct , and to help not to hate or to revenge , even she /he gives low marks in order to refine / reduce the tension between student and his teacher .

10-Students need some motivational strategies (like: Audio-Visual Aids) for learning English successfully .

11- Students should work hard to build a strong background and self- -confidence .

12-Students must be involved in asking for clarification, making requests, expressing their own ideas .

13- Students must be interested in learning English because it comes to be International Language.

14- Principals should understand that the number of students in class has a strong and negative effect on teaching and learning.

15- Social and principal's sides can intervene and influence students scores , but teachers should be more interested in improving student learning than in increasing scores .

16-Schools should introducepsychologists to reduce negative effects of mixed ability, stress, anxiety and tension between students and their teachers, and help students to perceive the importance of examinations and tests.

List of Content

Dedication	i
Acknowledgements.....	ii
Abstract.....	iii
List of Tables.....	iv
List of Graphs.....	vi
List of Content.....	viii
General Introduction.....	1
Background of the Study.....	1
The Origin of the “ Baccalaureate”.....	1
Purpose of the Study.....	2
Significance of the Study.....	2
Main Questions.....	2
Research Hypotheses	3
Methodology and Research Tools.....	3
Research Limitations.....	3
Structure of the Study.....	4
Chapter One : Investigating the Concept of Washback.....	6
Introduction.....	6
I.1. The Origin of Washback.....	6
I.2.The Nature and the Scope of Washback.....	8
I.3. Types of Washback.....	11
I.3.1.Negative Washback.....	11
I.3.2. Positive Washback.....	12
Conclusion.....	13
Chapter Two : A Probe on Standardized Tests.....	15
Introduction.....	15
II.1.Differences Between Evaluation , Assessment , and Test.....	15
II.2.The Nature of Standardized Tests.....	16
II.3.Types of Standardized Tests.....	17
II.3.1. Norm-Referenced Tests	18
II.3.2.Criterion-Referenced Tests.....	18
II.4. Characteristics of Good Tests.....	20

II.4.1- Validity.....	20
II.4.2.Reliability	21
II.4.3.Practicality	22
Conclusion.....	23
Chapter Three : Data Analysis.....	25
Introduction.....	25
Structure of the Questionnaires	25
Catchment Area (population and sample).....	26
III.1.The Analysis of Teachers' Questionnaire.....	27
Section One : Background Information.	27
Section Two : Investigating the Concept of Washback	32
Section Three: A Probe on Standardized Tests.(e,g:BAC examination).....	44
III .2 The Analysis of Students' Questionnaire.....	51
Section One : Background Information.....	51
Section Two : Investigating the Concept of Washback.....	57
Section Three: A Probe on Standardized Tests.(eg: BAC examination).....	65
Conclusion	72
Omnibus Recommendations.....	74
General Conclusion	76
References	78
Appendices	
Questionnaire for Teachers.....	81
Questionnaire for Students	87
Questionnaire for Students (Translated).....	92
Arabic Abstract	97

List of Graphs :

III.1.The Analysis of Teachers' Questionnaire.

Graph1. Teachers' gender.....	27
Graph 2. Teachers' age.....	28
Graph 3. Teachers' grade.....	28
Graph4. Years of teaching English.....	29
Graph 5. Years of teaching as an exam teacher.....	30
Graph 6. The medium of instruction used.	31
Graph 7. Teaching an examination class.....	33
Graph 8. The major changes in the BAC examination.....	34
Graph 9. The major reasons for the changes of the English BAC examination.....	36
Graph 10. Extra work or pressure that these changes will put on you.....	37
Graph 11. The likely changes in teaching	39
Graph 12. The most difficult aspects of teaching	40
Graph13. Factors that influence most teaching	42
Graph14. Promoting positive washback effects	43
Graph15. Test preparation time.	45
Graph16. Characteristics changes through your teaching process.....	46
Graph17. Attitudes towards teaching and testing.....	48
Graph 18. The effects of standardized tests	49
III .2 The Analysis of Students' Questionnaire	
Graph 19. Students' age.....	51
Graph 20 . Students' gender.....	52
Graph21. Years of studying English.....	53
Graph 22. Students' interest in learning English.....	53
Graph 23. Students' reasons for learning English.....	54
Graph 24. Suitable medium of instructions.....	55

Graph 25. Students confident to pass BAC examination.....	56
Graph 26. Reference to the BAC examination.....	57
Graph 27 . Students’ reaction when reference to the BAC examination.....	58
Graph28. Students’ preparation for the BAC examination.....	59
Graph 29. Teacher talk in English sessions	60
Graph 30. Teacher activities in English sessions.....	61
Graph 31. Students’ activities in English sessions.....	62
Graph 32. Students’ preferred strategies for learning English.....	64
Graph33. The effects of exam scores.....	65
Graph 34. Aspects affected by exam scores.....	66
Graph35. Students’opinions towards the examinations.....	68
Graph 36. If I want I can get good grades.....	69
Graph37. When I really try hard I can do well.....	70
Graph 38, I am a good student.....	70
Graph 39. Students’ performance in the BAC examination.....	71

List of Tables :

III.1.The Analysis of Teachers’ Questionnaire.....	27
Table1.Teachers’ gender.....	27
Table 2.Teachers’ age.....	27
Table3.Teachers’ grade.....	28
Table 4. Years of Teaching English.....	29
Table 5.Years of teaching as an exam teacher.....	29
Table6. The medium of instruction used.....	30
Table 7.Teaching an examination class.....	32
Table 8. The major changes in the BAC examination.....	34
Table 9.The major reasons for the changes of the English BAC examination.....	35
Table 10. Extra work or pressure that these changes will put on you.....	37
Table 11. The likely changes in teaching	38
Table 12. The most difficult aspects of teaching.....	40
Table 13. Factors that influence most teaching.....	41
Table 14. Promoting positive washback effects.....	43
Table 15. Test preparation time.	44
Table 16.Characteristics changes through your teaching process.....	46
Table17.Attitudes towards teaching and testing.....	47
Table 18. The effects of standardized tests	49
III .2 The Analysis of Students’ Questionnaire.....	51
Table 19.Students’ age.....	51
Table 20 .Students’ gender.....	52
Table 21.Years of studying English.....	52
Table 22.Students’interest in learning English.....	53
Table 23.Students’ reasons for learning English.....	54
Table 24.Suitable medium of instructions.....	55

Table 25. Students confident to pass BAC examination.....	56
Table 26. Reference to the BAC examination.....	57
Table 27 . Students’ reaction when reference to the BAC examination.....	58
Table 28. Students’ preparation for the BAC examination.....	58
Table 29. Teacher talk in English sessions.....	59
Table 30. Teacher activities in English sessions.....	61
Table 31. Students’ activities in English sessions.....	62
Table 32. Students’ preferred strategies for learning English.....	63
Table 33. The effects of exam scores.....	65
Table 34. Aspects affected by exam scores.....	66
Table 35. Students’ opinions towards the examinations	67
Table 36. If I want I can get good grades.....	69
Table 37. When I really try hard I can do well.....	69
Table 38. I am a good student.....	70
Table 39. Students’ performance in the BAC examination.....	71