SciencePG
Science Publishing Group

Implementing a New Testing System in the Algerian Secondary Schools: *A Washback Study*

Hoadjli Ahmed Chaouki

Department of Foreign Languages, University of Mohamed Kheider, Biskra, Algeria

Email address:
chaouki.hoadjli@gmail.com

Abstract: This study explored the issue of evaluation in general and language testing in particular in the Algerian educational system. In precise terms, the present research investigated how a particular type of language tests, achievement tests, should be designed and developed in order to make them fit their intended purposes. This aimed to provide EFL teachers at the secondary school level in Biskra region with theoretical and practical understanding, with the expectations to make these developed tests become part of the learning experience and not a mere operation of grading for the sake of passing or failing. The rationale for this study is twofold. Firstly, given that there is a multitude of washback studies of language tests on teaching and learning, no empirically study of this sort was available for this particular context. Secondly, this study stems from the students' low scores that do not actually reflect their performance by virtue of the non-pertinence between what these students learn and what they are tested in. Ultimately, the fundamental objective of the present exploration is to yield practical insights to overcome the array of anomalies present in the current testing system. In order to achieve this objective, an Alternative Testing Model was proposed. This model was first tried-out and then put into practice on a selected sample of EFL classes in Biskra region. In the investigation of the washback effects of the new testing system on participants' perceptions of, and attitudes towards, teaching and learning, a mixed-methods approach was employed, using four data collection methods to gather insightful feedback. From the analysis of these data, the findings show that the new testing policy had positive washback effects on teachers and students in this study. It succeeded to change to some extent and with different degrees these participants' attitudes and behaviours in teaching, learning, and assessment practices.

Keywords: Achievement Test, Algerian Secondary Schools, ATM, Test Design, Test Development, Washback Effect

1. Study Background

In this study, the Algerian educational context at the micro-level referred to the teachers and students at the Algerian secondary level was explored. These subjects aimed at teaching and learning English as a foreign language for three years that culminate the secondary education cycle which begin earlier at the primary and middle school levels. The teachers at this level are expected to design and develop the required testing instruments that are meant to evaluate and assess students' attainment and progress in accordance with the taught programmes and the subsequent instructional objectives set out in the correspondent syllabi.

The analysis of a preliminary investigation on EFL teachers and students in this context generated a number of interesting findings in relation to the impact of the current testing practices. The main finding is that of the influence of the 'Bac' examination on both teaching/learning and assessment practices. That is, a great deal of teachers believed that this examination influenced the content of their teaching and classroom in that this examination made them pay more attention to grammar and vocabulary and also some other language components, such as the written skills rather than the oral skills throughout the school year.

This practice would have encouraged teachers to use non-communicative methodology in ordinary classes, which markedly contrasted with the underlying principles of the adopted official methodology that prelude communicative teaching. According to the atmosphere that reigned in almost all the secondary schools in Algeria, whole class activities

dominated. This suggested a fairly traditional teaching method where the teacher is in charge of the events in the classroom. Whole class activities mainly consist of the teacher's interaction with one or several students, very rarely students get to lead the activity. Individual work is not very common in those classes, giving place to class time spent on whole class activities, and a total dominance of the teacher talk instead of the interaction between teacher and students, and students with students.

Equally important is the role of the 'Bac' examination and the intense pressure it seems to impose on students. The observations show that the students attitudes towards the use of that target language inside and outside the classroom is dominated by the use of some English with occasional Arabic. The right frequency of the use of a mixture of the native and target language could be explained by the assumption that teachers provided less practice opportunities for the students.

Meanwhile, their emphasis on to train the students to be fixed on some typology of activities seems to make these students to develop the skills that could help them to respond to questions that are recurrent in the tests that they usually construct. The above observations referring to the teachers and students towards the use of English language inside and outside the classroom revealed again the strong pressure the 'Bac' examination often exerted on both students and teachers.

Given the above conclusion, it is necessary to set some remedies that are expected to repair the anomalies that area current practice among EFL teachers and students in the Algerian secondary schools. These remedies would stand as practical solutions to problems that resulted from the misunderstanding of the practitioners between the tasks they are hoped to realize and the requirements of the reality, the misconceptions that have become ordinary with time, and therefore are accepted as a common sense, and the resistance to change which is common at that micro-level in the area under scrutiny. The tenets of these remedies are the core of the Alternative Testing Model (ATM) that the current researcher has suggested. The fundamental purpose of this innovative act seeks to replace the current testing system in order to improve and appropriate the assessment practices, and hence make them compatible with the teaching and learning processes.

2. Literature Review

2.1. Washback in Language Testing Revisited

An important assumption behind the numerous uses of language tests is that a language test can exert an influence on both teaching and learning. It can even transgress these levels to concern broader areas in the educational setting. On the assessment literature, there is a clear consensus that the impact of a test on teachers and learners is termed washback.

For many specialists, the concept of washback direction encapsulates the principles that some efforts of a language test may be beneficial, while some other efforts may be harmful. In this sense, washback is judged positive or negative according to how far it enhances or reduces forms of teaching and learning evaluated to be appropriate. Of course, what is considered appropriate will depend on the instructional objectives one espouses.

2.1.1. Positive Washback

This phenomenon refers to tests and examinations that influence teaching and learning positively (Alderson, Wall 1993: 15). In a broad sense, good tests can be utilized and designed as beneficial teaching-learning activities so as to encourage teaching-learning process (Pearson 1988:7). Andrews et al. (2002) suggest deliberately introducing innovation in the language curriculum through modifications in language testing. Davies (1985) considers that 'creative and innovation tests can advantageously result in syllabus alternation or even in a new syllabus' (18). In this way, a test no longer needs to be only an obedient servant; rather, it can also be a leader.

2.1.2. Negative Washback

Negative washback is seen by testing researchers as the negative influence of tests on teaching and learning. Alderson and Wall (1993:5) point out that:

A negative washback is defined as the undesirable effects on teaching and learning of a particular test. The test may fail to reflect the learning principles and /or the course objectives to which they are supposedly related.

In such a case, these tests will lead to the narrowing of content in the curriculum, instead of covering a definite content from what has been learnt in class. For Vernon (1956), 'teachers tend to ignore subjects in activities that are not directly related to passing examinations, and testing accordingly after the curriculum in a negative way' (18). Thus, it is logical that these tests may fail to create correspondence between the learning principles and /or the course objectives to which they should be related (Cheng 2005:8). More dangerous, negative washback can substantially reduce the time available for instruction, narrow curriculum offering and modes of instruction, and potentially reduce the capacities for teachers to teach content and use methods and materials that are incompatible with useful testing instruments (Smith 1991:120). In such a learning content, an atmosphere of high anxiety and fear of test results become common among teachers and learners (Shohamy et al. 1996:9). For teachers, they will feel that success or failure of their students is reflected on them, and they speak of pressure to cover the material for the examinations. When the students know that one single measure of performance can determine their levels, they will less likely take a positive attitude towards learning.

2.1.3. Washback: Lever of Innovation in Education

Enchanted by the power of the decisions they can provide; language tests are key components in projects to introduce in education through interventions. This simply means that a language test can stand as an effective instrument to reform or innovate in educational or language testing systems. In

line with this assumption, research on washback has demonstrated that a language test can also be a way by which teachers, and students' behaviours and perceptions of their abilities can change. In addition, a testing instrument may influence the content and methodology of teaching programmes, attitudes, towards the value of certain educational goals and activities. Eventually, it may even serve the needs of society as a whole. Given this potential power, a language test, as an innovative act, can thus become a useful source for contending participants who seek to bring about a new vision of what education in general and instruction should be.

2.1.4. A Washback Model for the Present Study

The model proposes that the nature of washback from language tests flows from overlap, the distance between the contents, and the instructional objectives set out in relevant syllabi. The greater there is a correspondence between the two, the likely positive washback becomes. Nonetheless, in this model, washback is not simply a matter of test design; it is realised through, and limited by, participants' characteristics, participants' perceptions of, attitudes towards, and reactions to, of test importance and difficulty and their ability to accommodate to test demands, will moderate the strength of any effort and certainly the evaluation of its value.

To provide a structure of this model, the researcher used an adapted model of washback that is based on the framework that was elicited by Hughes (1993). The choice or this model is dictated by the idea that it is appropriate to the nature of the present investigation. To describe it, it is important to distinguish between three components. The latter are: participants, process, and product, recognising that the three components may be affected by the nature of a test. In what follows, a brief identification of each one of these three components is displayed in below:

(i). Washback on Participants

Bailey (1999:12) views that the participants either teachers or students affected by washback may be influenced by information that a test undertakes prior to its administration. Besides, these participants may also be influenced by several sources of feedback following the administration of the test. These would include the actual test scores provided by teachers, feedback from students, and feedback from the teachers to the students' scores.

(ii). Washback on Process

The second component in the adopted model is washback on process. Hughes (1993:2) defines process as 'any actions taken by the participants which may contribute to the process of learning'. Specifically, the author includes processes, such as materials development, syllabus, teaching methods, contents of teaching, strategies, and assessment.

(iii). Washback on Products

The last component in the model is washback on products. Hughes (ibid) asserts it as, 'what is learnt'. This mainly concerns matters, such as facts and skills, and the quality of

learning. What is notable on this component is that it is sometimes difficult to entangle it from the two components that are participants and process. For Bailey (1989) much of the literature about participants and washback describes the various processes participants try to increase.

Drawing on the reviewed literature, dependent variables in this study include the effects of the testing model on participants' attitudes and beliefs, the content and methods of teaching and learning and the students' outcomes. Therefore, this study aims to explore how the different components in the Algerian educational system reacted when washback was strategically anticipated to determine the possible areas of washback intensity in teaching and learning English language, and to define the interrelationship between the 'who' changes, 'what', how, 'where, and 'why'.

2.2. The Alternative Testing Model: A Rationale

On the testing literature, a great deal of testing frameworks about how language tests should be developed are available. Each one of these is an attempt to yield the most practical and easiest process. At the same time, it is another way to provide useful tests that can lead to the best final results from which inferences for decisions to be drawn out.

In this study, we have thought to opt for Bachman and Palmer's (2000) model with an adaptation of this model to suit the requirements of this exploration. The choice is interpreted in terms of a set of reasons. First, Bachman and Palmer model is considered by most of language testing specialists to be the recent. It is a synthesis of a big number of other available models. Second, this model is practical and easy to put into practice. Third, an important advantage of this model is that it can enable test users to review, re-consider, and re-evaluate the decisions and activities that have been taken and performed so far. This flexibility allows any user of this model to innovate at any time there is a need to do that. In sum, Bachman and Palmer's model has been judged to be the most appropriate choice for the present research.

2.3. The Alternative Testing Model: A Conceptual Framework

Conceptually, the process of test design and development is organised into three stages: design, operationalisation, and administration. What is specific and worth to know is that the process is not sequential in its implementation. In practice, it is usually stated that test development is linear, with development process from one stage to another. Meanwhile, it is also believed that this process is iterative since the decisions that are made and the activities that are completed at one stage may lead teachers to reconsider and revise them.

2.4. The Alternative Testing Model: The Structure

After revising the shortcomings and contradictions found in the current testing system, a set of modifications, and additions were suggested from the conceptual stage to the operational. These mainly concern:
(1) The development of an achievement test formed of

three sub-tests.

(2) The three sub-tests were devised to assess the following skills and language components:

(3) Sub-test One: Testing Listening

(4) Sub-test Two: Testing Speaking

(5) Sub-test Three: Testing Writing, Reading, Grammar, and Vocabulary.

(6) A revision of the testing methods was proposed.

(7) New testing tasks were added.

(8) A redistribution of test-items was introduced. This was done in accordance with the new sectioning.

(9) A revision of test instructions was also made.

(10) A new scoring scale was proposed.

3. The Study

3.1. Research Questions and Purposes

The research in this study explored teachers' and students' attitudes towards the ATM. In precise words, it investigated whether the implementation of the new model had any washback effects on classroom teaching and learning in terms of teachers' and students' behaviours. The research questions were formulated as follows:

RQ1. What is the nature and scope of the washback effects on teachers' behaviours because of the ATM implementation?

RQ2. What is the nature and scope of the washback effects on students' behaviours because of the ATM implementation?

3.2. Methodology

Because of the nature of this study, a Mixed-methods Approach seemed the most appropriate. A Mixed-methods Approach which is both quantitative and qualitative driven data would allow the refinement and checking of the used data collection.

In this respect, 4 data collection methods were employed. These are stated as follows:

(1) Questionnaires

(2) Interviews

(3) Focus groups

(4) Observation

In order to verify the results and check for the validity of these results, the SPSS (Statistical Package for Social Sciences) method was used.

3.3. Population and Sample

6 secondary schools were sampled from the 15 that included EFL classes in the Biskra region (Algeria) where the study took place. The 6 schools comprised 33% of the school population. From these 6 schools, 8 classes were selected.

Based on the nature of the methodological considerations of the present study, the following criteria to identify potential schools were selected:

(1) Schools which comprised EFL classes;

(2) Schools which agreed to accept to take part in this study;

(3) Schools which agreed to accept that selected EFL classes would be observed;

(4) Schools which agreed to accept that teachers and students would be allowed to participate in questionnaires;

(5) Schools which agreed to collaborate with the researcher;

(6) School administrations who had a positive attitude towards the study.

4. Results and Discussion

The findings of this study obtained using four data collection methods have led to a better understanding of the participants' reactions at the micro-level of the washback effects of the Alternative Testing Model (ATM) on teaching, learning, and assessment practices. First of all, a review of the questionnaire findings from both teachers and students provided insightful feedback regarding classroom practices. With regard to the teacher's questionnaire exploring the teachers' perceptions of the reasons behind the intervention, it can be seen that those latter were consistent with the underlying theories on which the new testing policy rested. Teachers are aware that the main purposes of the introduction of new tests are in line with the idea that considers the students in the context under exploration as active agents who possess a high degree of competence of the learnt language. This, in itself, is the fundamental purpose of teaching English language in the Algerian secondary school, which is clearly identified in the three syllabi at this level, and that enables students to communicate to some extent fluently and accurately, or at least, with some ease, inside and outside the classroom. Concerning the alterations to teaching due to the intervention, evidence from the teachers indicated that these participants no longer are fixed on a given typology of activities that are currently included in the 'Bac' examination papers, as it used to be in the period of time that preceded the introduction of the new testing system; rather, there is a more positive attitude of these teachers towards the relative importance of classroom activities, and the subsequent use of new teaching methods. In this respect, it was found that the participant teachers in this study give a higher weighing to the activities of 'putting more stress on the co-operative learning', 'putting more stress on group discussion', and 'putting more stress on speaking and the integration of the four skills'. Certainly, these activities and perhaps other ones can increase students' communicative competence, indicating a degree of interest of these teachers to make changes in certain aspects of teaching. In short, there seemed to be a washback effect of the new policy on aspects related to teachers' reactions to the new system, the reasons behind introducing the ATM, and alterations to teaching methods due to the new policy.

Contrary to the positive indications obtained with the teacher' perceptions of, and reactions towards the new policy, the result in the area related to aspects of teaching materials revealed that the teachers still rely on the official textbook as

the major resource in their teaching. It appeared that the majority of the teachers relied mainly on the content and organisation of general English textbook as the source of their plans because they do not want to deviate from the official instructions recommended by the educational authorities, mainly those of the general inspectors, bearing in mind that one of the limitation of this study was that it is extremely difficult to secure the necessary support and involvement of all the teachers in this study because it is often perceived not to have the authority and status to compel the participants to apply all what it suggests. Given this reality, it is not surprising that the washback effect on classroom teaching resources appears to have been minimal since the teachers for the most part did not try to change so much the teaching materials, although in many cases they expressed in their answers they were not convinced of what they were doing, and they showed some interest to opt for other resources.

In terms of classroom teaching behaviours, the majority of the teachers used half English and half Arabic as the medium of instruction. According to the teachers, this was due to the weak general level of their students in English language.

Furthermore, although it was noticed in this study that teachers became aware to talk less to students, and talk more often to groups or to individuals aiming at providing the students more practice opportunities, talking to the whole class was still the main activity in the classroom. The results of the teacher questionnaire in the study showed that the pattern of teacher's talk to the whole class was the most frequent, though as discussed earlier, the new policy did appear to encourage teachers to be more learner-centred. In brief, there seemed to be a washback effect of the new testing system in the area of teacher, but it was not as considerable as it was expected.

Regarding the factors that influenced how teachers perceived the issue of assessment and evaluation, the results indicated that these factors proved to be significantly different from the time when the teachers' main expectations were to obtain good scores in the tests, and pass the 'Bac' examination to a new time when the teachers' expectations are to obtain satisfaction in teaching. Accordingly, there is an evidence that using the ATM as an agent of change could possibly make teachers pay more attention to their own expectations, as well as to their students.

When aspects of learning were investigated from the teachers' perspectives, it appeared that these teachers were more willing to encourage their students cope with the innovative act, rather than change their own methodology. Teachers' attitudes towards how students should study in the content of the new policy was positive to put more emphasis on aspects which suggest change; the same attitude was obtained towards learning strategies. It appeared that teachers' attitudes towards aspects of learning in terms of learning aims and learning strategies indicated that there is a washback effect, even though the teachers' reluctance to make change could be explained by the resistance to innovation and the short time the new policy has taken to be discovered and well understood by the teachers in the context under investigation.

Finally, in the teacher questionnaire, exploring teachers' reactions to the new testing model shows that generally there is a positive washback. A great deal of the participants welcomed the intervention. They were attracted by the new format and contents of the achievement tests based on the ATM. They declared that they are ready to abandon what they used to use as assessment instruments, and embrace the new testing system. Such a positive reaction towards the policy is also interpreted in terms of the fact the new model encompasses a new philosophy, approach, and methodology.

Besides washback on teaching, major issues regarding washback on learning explored in this study were the basic concerns of the student's questionnaire in this last phase. First, with respect to washback in relation to the student' attitudes towards teaching and learning activities inside and outside class, the results obtained of this aspect confirmed the findings from the teacher's questionnaire. The general atmosphere is characterised by a dominance of teacher control. Although one major aim of the introduction of the ATM was to encourage speaking and oral practice, and to enhance pair and group work, there was a lower frequency in carrying out these activities and practices. Again, this reality should not hide the fact that in comparison to what prevailed in the scene during the pre-implementation period of the new testing system, the situation has to some extent changed but not to what is expected. Perhaps, there is a need for more time to get positive feedback. Overall, the result indicated that there was actually a minimal washback effect on the activities that students carried out inside and outside class.

With respect to the students' attitudes towards aspects of learning, the analysis of the results showed that the students' expectations correlate with the new intentions of the new testing system implementation. The students' response claimed that to have better opportunities in the future and to be able to communicate with foreigners in English language, this language could be a ticket to the world beyond the context where they live and study. The students also revealed an acute consciousness of the fact that English is a language that they are compelled to learn to open new horizons and get a place in the world of today. This general mood corresponds to a great extent to the new policy in which the functional use of English language is encouraged to be considered as an important aspect of learning. In addition, in terms of the students' preferred learning strategies, it is certain that with regard to these students' perceptions of the new model, their answers reported a positive reaction, and influence to the new achievement tests since the most used strategies were the one that suggest a washback effect on oral communicative strategies.

With regard to students' attitudes towards the impact of the new policy on aspects of testing, the findings indicated a positive shift in the students' perceptions. It can be seen that before the intervention, all the students said they did not enjoy sitting for the test. This feeling of disgust was interpreted in how these students viewed tests. For most of them, a test used to be equivalent to a sanction. After the incorporation of the new model, the students turned to regard the new achievement

tests as a useful means to see how well or badly they have achieved in correspondence to what they learn in class. More importantly, this new perception considered tests as effective instruments to learn English language. In sum, these students viewed the new model in a positive light. In particular, this indicates that they have changed their attitudes towards the function of tests.

The analysis of the findings related to aspects of the ATM format and contents indicated that the students appreciated the new structure, considering that sectioning a test into these sub-tests allows a better assessment of the four skills and language components. Besides, the students believed that the variety and richness of the contents of the tests could be a good indication that they are representative to a large extent to their lessons. This new attitude towards this matter completely differs from that of the pre-implementation period. Finally, the findings in this section confirmed this positive attitude towards the new testing model since it seems that all the students welcomed the new scoring procedures and scales. They thought that those latter are comprehensive since they can provide them with actual final outcomes. All in all, it is ostensible that the changes that occurred to the format and contents of the new, proposed tests have brought about positive attitudes of the students towards both assessment and evaluation.

Secondly, besides to the findings and discussion of the results from the teacher's and student's questionnaires, a discussion of the classroom observation outcomes is indispensable to understand the nature of washback in the context under investigation. Because the major purpose of the new testing policy was to improve students' communicative skills, it was essential to explore a sample lesson in detail to see whether there are any discernible washback. In terms of participants' organisation, the patterns of teacher interaction in the classroom revealed some differences among the 4 observed teachers. For 3 of them, teacher talk dominated the lesson time, whereas with 1 teacher, there were some differences, especially, in allocation of time in teacher and student talk. Indeed, in T3's lessons, there was less teacher talk than in the other teachers' lessons. T3's lessons were much more learner-centred, and consequently showed the most amount of student activity. As a result, the introduction of the ATM seems that it has encouraged a more learner - controlled approach, even though this result was observed with only one teacher. This in itself is a positive light.

With regard to the activity types and content as percentage of the class, the classroom observations revealed that although in general the approaches to teaching did not vary, since it was observed that the teachers resisted to keep on using the same practices they were carrying out for a long period of time, there was an increase in the time devoted to student activities with serious attempts to make the lessons more interactive. It is apparent that because of the new policy, the teachers have become aware that it is crucial to change many of their classroom activities, and the content of their teaching to achieve more substantial results at the end. What is more

significant on these aspects of activity types and teaching content is that there were many signs of a change in the teachers' fundamental beliefs and attitudes about how to carry out teaching, and what should be included in lessons to make all the participants in the teaching operation get profit. In brief, what was notable from the study, even though it was a short period of time since the new policy had been in operation and short period of research, there existed an actual washback effect of the ATM, or a perceived effect, indicating the power of the existence of the new tests.

In terms of the final aspect in the observations, which mainly concerned the used teaching materials, it is actual that almost the same conclusion drawn out from the previous aspects of the activity types and contents of lessons was attained. The findings suggest that the observed teachers most of the time relied on the official textbook, but with a progress in the application of the new policy, the teachers began incorporating some other new teaching resources that pertain to the kind of teaching materials that correspond to where less teacher talk and more oral student practices were involved. Examples of these materials included CD's, data show, pictures, cards, and other audio-visuals. Overall, it can be seen that a washback effect of the ATM has motivated the teachers to review at least some of their teaching materials, and suggest new ones.

Thirdly, a further careful look at the interviewed teachers' and students' responses revealed how these participants perceived the new policy. All the interviewees agreed to say that the tests are completely different from the current. Essentially, they have found that the new crucial criterion these tests have brought is that they do not eschew any skills or language components; reasonably, they thought that because in one test, three sub-tests that assess the four skills, vocabulary and grammar make it possible to provide students with actual grades. In particular, a number of the teachers affirmed that this new model has made it clear for them what stages one has to follow to be systematic, and what to consider when time comes to develop their tests. This finding aided the teachers to avoid relying on the same tests that frequently looked similar to the 'Bac' exam papers. Finally, the findings from the interviews revealed that the intended washback, as it is expected by this study, was to an extent achieved, despite the fact that the new policy is still at its initial phase of diffusion and implementation.

5. Conclusion

In sum, one of the fundamental purposes of this study is to sensitise the secondary school English teachers in Algeria that the deleterious effects of the current testing system can be overcome by switching to the evidence that there should be permanent revision and revising of the ways that these teachers usually employ to check their students' achievement and progress in English language. The need to realise that testing is not a simple operation of assigning scores, but it is an ongoing process that requires test design and development stages and that the final outcomes involve building a logical

case in support of a particular interpretation. Because of this way of conceiving the issue under study in this investigation, we proposed an ATM that would be regarded as the essential basis for quality control throughout the entire test development process.

Therefore, we would argue that in this study all test development and use should be informed by this model. The most important consideration in it is the use for which it is intended, so that the most important quality to develop achievement tests based on this model is its usefulness. In this way, the proposed ATM is thought to make the teachers in the context under exploration aware of the possibility to have their tests closely linked to the goals of instruction as to be almost undistinguishable from them; it is asserted that teaching to these tests, as they are conceived by the model, would be considered a virtue.

To emphasise, the ATM had an impact on a great deal of participants in this study. It has increased their answers in the roles they can take in contributing to positive changes in assessment practices. It has provided them the opportunities to critically re-examine their past and current practices in assessment. It has also made them discover the problems of the weaknesses they were currently doing. Crucially, they have become aware of how they can make use of language assessment literacy and knowledge in the future. Another important gain that has emerged from this study for participants is that the new testing model has strengthened their confidence in making sound judgements on their students' achievement and progress. This would make the teachers more self-consistent in assessing the students' performance. In short, this testing system has managed to change the teachers' and students' attitudes and behaviours in teaching, learning, and assessment practices.

References

[1] Alderson, J. C., & Wall. D. (1993). Does Washback Exist? Applied linguistics, 14, 15-29.

[2] Bachman, L. F., & Palmer. A. S. (2000). Modern Language Testing at the Turn of the Century. Accusing What Counts. In J. C. Alderson, & B. editors, Language Testing 17 (1).

[3] Bailey, K. M. (1989). Washback in Language Testing. Princeton, New Jersey: Education Testing Service.

[4] Cheng, L. (2005). Changing Language Teaching through Language Testing: A Washback Study. Cambridge: Cambridge University Press.

[5] Davies, A. (1985) Follow Any Leader: Is that What Language Tests Do? In Y. P Lee, A. C. Y. Y. Forlord & G. Low (eds.) New Directions in Language Testing, 3 (12). Oxford: Pergamon Press.

[6] Hughes, A. (1993). Backwash and TOEFL 2000. Unpublished Manuscript, University of Reading.

[7] Pearson, I. (1988). Tests as Levers to Change. In D. Chamberlain, & R. J. Baumgardner (eds.) ESP in the classroom: Practice and Evaluation (98-107). UK: Modern English Publications.

[8] Shohamy, E., Donitsa- Schmidt, & Ferman. I. (1996). Test Impact Revised: Washback Effects of External Testing 13, 298, 317.

[9] Smith, M. C. (1991). Put the Test: The Effects of External Testing on Teachers. Educational Researcher, 20 (5), 8-11.

[10] Vernon, P. F. (1956). The Measurement of Abilities. London: University of London Press.