

Review of Empirical Studies of Washback in Language Testing

Raju Ahmmed¹ and Kh. Atikur Rahman²

Abstract

Washback is one of the very few areas of English language assessment research in the last 25 years that has gained substantial attention. Commencing with the phenomenal work of Alderson and Wall (1993), a considerable body of empirical studies of washback has been carried out throughout the world. This paper reviews ten recent empirical studies of washback in language teaching between 2011 and 2018. The studies have been collected from several databases like ScienceDirect, ERIC, ResearchGate, Google Scholar and the peer-reviewed journals and university websites. This review demonstrates the research findings and theoretical underpinnings of the washback of assessments and tests in language teaching and testing. This review research finds that washback of high stakes test has both positive, negative and mixed (both positive and negative) impacts depending on the specific contexts and learners' levels. The negative washback occurred when the focus shifted from learning the English language to test-taking strategies whereas the positive washback affected teachers, teaching methodologies, teaching contents, materials, learners and learner strategies.

Keywords: Washback, Teaching Methodologies, Teaching Contents, Learner Strategies.

Introduction

Madaus (1988: 83) opined that “it is testing, not the ‘official’ stated curriculum, that is increasingly determining what is taught, how it is taught, what is learned, and how it is learned.” This research article investigates recent studies (2011-2018) of washback to examine if they demonstrate this to be the case. The term, ‘Washback’ refers to the influence or impact of assessment and testing practices on the teaching-learning process. 1993 is the year which marks the beginning of the washback effect because this was the time when Alderson and Wall got their article titled ‘Does washback exist?’ published. The concept of the effect and influence of examinations, especially high-stakes examinations on language teaching and learning is, however, indicated by several terms such as ‘backwash’ (Hughes,

¹Lecturer, Department of English, Bangabandhu Sheikh Mujibur Rahman Maritime University, Bangladesh

²Assistant Professor, Department of English, Bangabandhu Sheikh Mujibur Rahman Maritime University, Bangladesh

1989), ‘washback’ (Alderson & Wall, 1993) and ‘impact’ (Baker, 1991; Wall, 1997), ‘consequential validity’ (Messick, 1989, 1996), ‘systemic validity’ (Frederiksen & Collins, 1989), ‘measurement-driven instruction’ (Popham, 1987) or ‘curricular alignment’ (Madaus, 1988; Smith, 1991). Below, the definitions given by various researchers are placed under the groupings of (a) backwash, (b) washback and (c) test impact.

Backwash

Biggs (1995) opines that backwash denotes that testing controls not only the curriculum but also teachers’ teaching methods and students’ learning strategies. Spolsky (1994:2) defines ‘backwash’ as a term that deals with the unforeseen side-effects of testing and not to the intended effects when the primary goal of the testing is the control of curricula. Hughes (1989:1) very precisely specifies backwash as “the effect of testing on teaching and learning”.

Washback

“Public examinations influence the attitudes, behaviours, and motivation of teachers, learners, and parents, and because examinations often come at the end of a course, this influence is seen working in a backward direction, hence the term, washback” (Pearson, 1988:98). Quite similar to Hughes (1989:1), Bailey (1996:259) defined washback as the “influence of testing on teaching and learning.” Alderson and Wall (1993) described washback as a phenomenon that forces teachers and learners to do certain things because of the test. Messick (1996:1) gave a similar definition of washback. He said washback is “the extent to which the introduction and the use of a test influence language and teachers to do things they would not otherwise do that promote or inhibit language learning”. Shohamy, et al. (1996:6) delineated washback as “the connections between testing and learning”. For Cheng (2005:8), washback indicates “an intended or unintended (accidental) direction and function of curriculum change on aspects of teaching and learning by means of a change of public examinations”.

Test Impact

Some researchers opine that tests have far-reaching effects in the educational world than in the language classroom. For example, Bachman and Palmer (1996:12) used the term “test impact” to refer to the effects that tests have on individuals (teachers and students) or educational systems and on the society at large. Wall (1997:11) held a similar view by stating that “Test impact refers to any of the effects that a test may have on individuals, policies or practices within the classroom, the school, the educational system, and society as a whole”. McNamara (2004:10) claimed 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”. Andrews (2004:9) used “test impact” to describe “the effects of tests on teaching and learning, the educational system, and the various stakeholders in the education process”.

Theoretical Framework

Alderson and Wall (1993:120-121), the pioneer, developed 15 washback hypotheses

according to what is influenced: teaching, learning, content, rate, sequence, degree, depth, attitudes and the number of teachers or learners affected by a test. The 15 hypotheses are: A test will influence 1) teaching; 2) learning; 3) what teachers teach; and 4) how teachers teach; 5) what learners learn; and 6) how learners learn; 7) the rate and sequence of teaching; and 8) the degree and the depth of learning; 9) the degree and the depth of teaching 10) the degree of learning; 11) attitudes towards the content and method of teaching and learning; 12) Tests that have important consequences will have washback; and conversely; 13) Tests that do not have important consequences will have no washback; 14) Tests will have washback on learners and teachers; 15) Tests will have washback effects for some learners and some teachers, but not for others. When studying washback, focus can be put on participants (teachers, students, material developers, publishers), process (actions by participants towards learning), and products (what is learned and the quality of learning), as suggested in Hughes's trichotomy model (Hughes, 1993 as cited in Bailey, 1996). The 3Ps introduced by Saville (2009) mentions these three principal aspects of washback, namely participants, processes and products (see Figure 1).

On the flip side, washback is conceptualised by Watanabe (2004) in terms of dimension (specificity, intensity, length, intentionality and value of the washback), aspects of learning and teaching that may be affected by the test, and the factors mediating the process of washback being generated (test factors, prestige factors, personal factors, macro-context-factors).

As far as the types are concerned, two types of washback can be observed: negative and positive. Negative washback occurs when test contents or format was based on a narrow

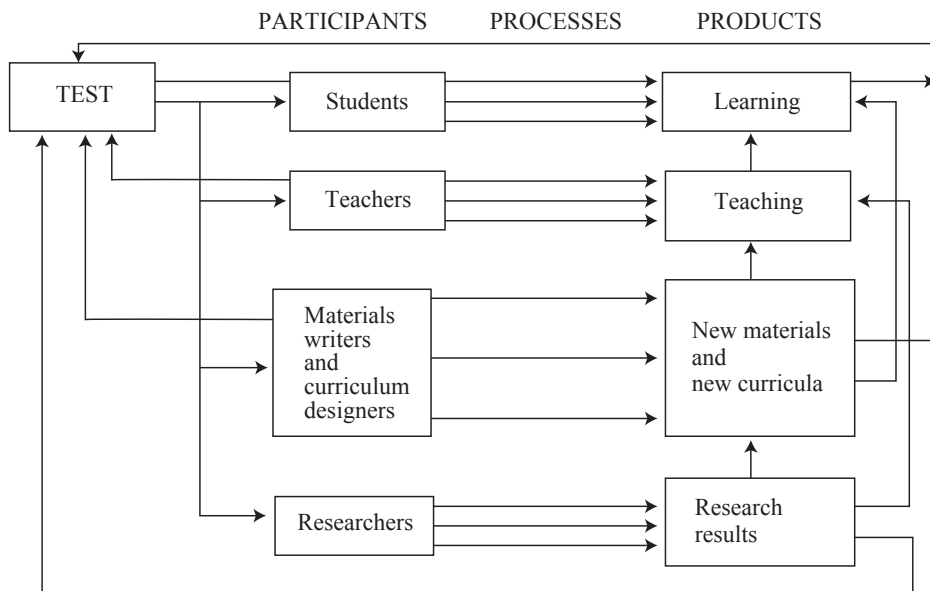


Figure 1: Basic Model of Washback (Bailey, 1996)

portrayal of language ability, and it confines the teaching-learning contexts. It refers to negative or undesirable influence on teaching and learning of a test, which means a poor test in which areas/activities that the teacher or student does not like to teach or learn and a discrepancy between the content (e.g., the material/abilities being taught) and the test (Alderson and Wall, 1993; Brown, 2004). Washback takes the form of negative washback when there is a discrepancy between construct definition and the test, or between content (e.g. the materials/abilities being taught). On the other hand, 'positive washback' refers to tests or examinations that influence teaching and learning beneficially (Alderson and Wall, 1993); where testing procedure promotes 'good' teaching practices (Taylor, 2005). Consequently, both teachers and students have a positive attitude towards the test and work voluntarily towards achieving its objectives.

The Rationale of the Study

The key to understanding and practically applying the findings of any investigation into test impact hinges upon the interpretation and analysis of the concept of washback. The goal of this review aims at playing that part. The review will help further understand the nature and complexity of washback which can facilitate teachers, test developers, curriculum specialists and other stakeholders to administer such tests which may create a positive effect.

The Methodology of the Study

This paper is a review based study on secondary data. Recent literature on washback studies from 2011 to 2018 has been reviewed. These reviewed studies have been collected from several databases like ScienceDirect, ERIC, ResearchGate, Google Scholar and peer-reviewed journals and university websites based on keywords e.g., washback studies, testing, and assessment etc. Data retrieved from various secondary sources are duly acknowledged.

Major Findings from the Studies Reviewed

This paper reviewed ten empirical studies of washback in language teaching conducted around the world between 2011 and 2018. The major finding of these studies are summarised and analysed below:

Cholis and Rizqi (2018) - Entrance Exam of Universities (EEU) in Indonesia

Cholis and Rizqi (2018) tested the effects of entrance exam in Indonesia which is a standardised test determining students' entry into universities. In the first part of this research, the researchers tried to find out how EEU affected teachers' attitudes in teaching. They found that EEU created some pressure or extra work (e.g. more lesson preparation, preparing and revising more materials, and preparing the students for the test and for the teachers). In the second part of this study, the researchers explored the teachers' point of views regarding the adoption of the teaching method. The respondents mentioned of more students' participation in class, putting more stress in the integrated skills (e.g. reading, writing, listening and speaking), using communicative approach and finally giving less emphasis on reading

comprehension. The result of the study reported positive washback since the teachers taught based on curriculum regardless of EEU format. So, the study of Cholis and Rizqi (2018) showed that the test did not bring any negative backwash as the teaching pedagogy was not completely dependent on the test contents and patterns. On the other hand, the majority of the respondent teachers in this research did not intend to adopt new teaching methods and employ real-life tasks which were examples of negative washback. At the same time, revising some of the existing materials, experiencing new challenges in teaching, and arranging new objectives of teaching, arranging additional examination practices for EEU test were examples of 'teaching to the test' and were evaluated as the negative effects of entrance examination in this research. However, as concluded in the study the positive influence of the entrance examination test outweighed the negative effects.

Khoshsima, Saed and Mousaei (2018) - Effect of Teaching Test-taking Strategies on Reading Section of IELTS in Iran

This study attempted to assess the impacts of teaching reading test strategies of IELTS candidates in Iran. The study was conducted on forty participants by an experimental research design in which the experimental group (20 participants) received test-taking instructions and the control group (the other 20) received general instructions. Before that, the researchers checked the participants' homogeneity (same level of competence). After the teaching sessions being done, the researchers took the reading test on the 40 candidates and scored the results. The result of this study showed that the mean score of the participants of the experimental group, who received test-taking instructions for their reading skill, was higher than those of control group, who were taught in a traditional method for their reading skills. So, the study of Khoshsima, Saed and Mousaei (2018) found that the teaching test-taking strategies had positive effects on Iranian IELTS candidates' performance on the reading section. They also took the perceptions of the experimental group participants regarding the teaching test-taking strategies and they found the same positive attitudes. This research finally recommended teaching strategies as 'beneficial' for the test-takers on a specific item in the IELTS test.

Munoz (2017) - Attitudes Towards Tests Scale (ATS) in Chile

Munoz (2017) studied the attitudes of the learners towards test in Chile. The study was conducted in experimental research design in which 25 learners were kept into the experimental group and 27 learners were in the control group. The researcher taught both of the groups for four hours of English instruction per week over a period of four months. Munoz (2017) taught the control group with the traditional practice of giving students' assessment results (just grades) whereas he taught the experimental group with washback techniques through which he graded the tests and written quizzes; handed them over to the students and discussed the answers. The result of this study showed that the experimental group had a better attitude towards tests regarding how tests helped them learn the contents; develop test-taking strategies such as time management or organisation skills; enhance self-confidence and motivation. This study finally suggested utilising washback techniques and indicated the improvement of respondents' attitude towards assessment and evaluation.

Khodabakhshzadeh, Zardkanloo, and Alipoor (2017) - Effect of Mock Tests on Iranian IELTS Candidates

This study attempted to find out whether mock IELTS exam (a preparatory test) is useful for IELTS test takers for increasing their scores. The study tried to see the role of IELTS preparatory courses and how its result could assist the material developers and curriculum designers to include or exclude these items in courses. Like the two previous reviewed studies, this study was also conducted in experimental research design in which 25 participants were put in Group 1 and 26 participants in Group 2. The researchers collected the data from their observations of the respondents who were from upper intermediate and advanced levels and did not participate in the IELTS course previously. In group 1, the researchers gave a mock test in each session and mailed the results to the participants before the next session. On the other hand, they taught group 2 with conventional deductive teachings like explaining the procedure, individual tasks in subsections and ways of

Table-1: Washback studies from 2016-2018

Studies	Cholis and Rizqi (2018)	Khoshsima, Saed and Mousaei (2018)	Munoz (2017)	Khodabakhshzadeh, Zardkanloo, and Alipoor (2017)	Barnes (2016)
Exams Studied	Entrance Exam of Universities (EEU) in Indonesia	Test-taking strategies on reading section of academic IELTS in Iran	Attitude towards Test Scale (ATS) in Chile	Mock tests in IELTS preparation courses in Iran	TOEFL iBT in Vietnam
Purposes	To explore the washback effect of a high-stakes test on teachers' attitude and teaching methods used	To investigate the impact of teaching reading test-taking strategies on the IELTS candidates' performance on the IELTS reading section	To improve learners' attitudes towards tests	To find out whether Mock IELTS exam could have any effect on IELTS candidates overall score on the IELTS exam	To investigate the washback effects of a high-stakes English language proficiency test, TOEFL iBT, in Vietnam.
Methodology	Survey design and questionnaire	Experimental design- experimental group received test taking instructions and the control group received general instructions	Action research- mixed method design, interview and questionnaire	Experimental design by dividing the participants in two groups and comparing their post-test scores	Observational data were collected
Collected Evidence	The respondent teachers showed positive attitudes. They showed a few negative issues in teaching methods	The score of experimental group was more than the general group. The respondents had a positive attitude	The experimental group scored higher than the control group. The experimental group showed positive attitudes towards tests	The group which practiced mock test outperformed the group that did not sit for mock test	The materials were dependent on TOEFL iBT text books. The teaching pedagogies did not change due to over dependence on the book
Conclusions	Positive influence outweighed the negative results	Teaching test-taking strategies had positive effects	Positive washback observed	Using Mock tests in the IELTS preparation courses can positively affect the participants scores on IELTS exam	Both the contents and pedagogies were influenced by the test. Negative washback found

answering certain types of questions. Khodabakhshzadeh, Zardkanloo, and Alipoor (2017) conducted ten sessions over a period of three months and after that, they compared the post-test scores of the participants of both groups. The results of the study revealed that the difference between the post-test scores of the two groups was significant and group 2 outperformed group 1. It became evident from this research that practising test-taking strategies were more effective than teaching course content in term of high stakes tests like IELTS. So, this research advocated the use of mock tests in IELTS preparation courses in Iran due to positive washback.

Barnes (2016) - Washback of TOEFL in Vietnam

Barnes' study (2016) examined the effects of TOEFL on the teaching contents and teaching pedagogy in Vietnam. The researcher chose two language centres in Vietnam which offered TOEFL preparation courses. The data was collected from the teaching materials and classroom observations. Barnes (2016) found teachers' overdependence on TOEFL iBT book and the existence of commercial materials. The researcher noticed that the majority of the materials collected in the class observations were not created by the teachers, but came directly from the classroom textbooks. Therefore, the teaching methods were not influenced much due to their reliance on TOEFL iBT textbook materials. Moreover, the majority of the practised tasks expected students to respond to questions individually. Barnes (2016) observed that the interactions in the class were from students to teachers, vice versa and students to students. Examples of the student to student or class interaction included oral presentations and class discussions. This study found that teaching activities took more time in class than student activities. All these findings of this study suggested that TOEFL iBT preparation courses in Vietnam emphasised the test skills but not the skills necessary for mastering a language. So, this study found the negative influence of the TOEFL iBT test on teaching methodologies and contents.

Mahmoudi (2015) - Washback of National University Entrance Exam in Iran

The study of Mahmoudi (2015) investigated the washback effect of Iranian National University Entrance Exam (INUEE) on the English learning process. This research was conducted on 218 female students at two pre-university schools in Iran and the instruments for collecting the data were students' questionnaire and an observation checklist. The study found that students' learning process and the teachers' teaching method were affected by the entrance exam. The INUEE affected learners' learning contents and learning strategies negatively whereas teachers' teaching method affected their learning in both positive and negative ways. As indicated in this research, the INUEE negatively affected the teaching and learning strategies through pushing the students towards mastering the test-tackling strategies rather than pursuing English for the sake of genuine learning. The majority of the respondent students studied their whole textbooks. Besides, most of the students did self-study to master the strategies for the entrance test. Mahmoudi (2015) identified students' most practice areas (grammar, vocabulary and reading) and the least practice areas (writing, speaking and listening). Students also changed their learning strategies to adjust to the entrance test and practised sample tests when the examination got closer. Therefore, Mahmoudi's study (2015)

suggested that students’ learning has been affected both positively as well as negatively. The washback was positive in the sense that students themselves developed some of the skills that were tested in the entrance exam. On the contrary, the washback was negative because only preparing for the test was not sufficient for learning the English language. However, this research concluded that the negative washback of the Iranian National University Entrance Exam outweighed the positive sides in the Iranian context.

Aftab, Qureshi and William (2014) - Intermediate English Examination in Pakistan

The study of Aftab, Qureshi and William (2014) investigated the nature and scope of the washback effect of the intermediate English examination in Pakistan which determines students’ entrance into university. The study used qualitative data collection method utilising interviews to take views of the teachers and students.

The respondent students demanded practice of examination related activities and relied on cramming of materials from guide (practice) books. As indicated in this research, students’

Table-2: Washback studies from 2011 – 2015

Studies	Mahmoudi (2015)	Aftab, Qureshi and William (2014)	Watanabe (2013)	Kirkpatrick (2012)	Sukyadi and Mardiani (2011)
Exams Studied	Iranian National University Entrance Exam (INUEE)	Intermediate English Examination in Pakistan	Center Test for University Admissions (an achievement test) in Japan	English Assessment System in Bhutan	English National Examination (ENE) in Indonesia
Purposes	To investigate the washback effect of INUEE on English learning process	To explore the nature and scope of the washback effect from the Pakistani Intermediate English examination	Explored the test validation and its effects on students’ achievement	To examine the washback effect of English assessment system	To evaluate the washback effect of ENE
Methodology	Questionnaire and Observations	Qualitative design-interviews	qualitative research method-exploratory analysis	Questionnaire	Qualitative method-observation, interviews, questionnaire and documents
Collected Evidence	Students’ learning process and teachers’ teaching methods were affected	Students crammed the answers. The teachers taught only the skills that were tested in the test	Measured the skills that were supposed to be measured. Students developed communicational skills which were tested in the test	Teachers’ willingness to accept new approaches to teaching, diagnostic feedback, use of rubrics. Lack of teacher training, lack of curriculum materials, crowded classes, and overemphasis on summative assessment	The school of high-level achievers had more positive washback. The school of moderate and low level achievers had more negative washback
Conclusion	Have both positive and negative washback.	The test had negative washback on teaching methodology, content and learning	The Centre test had positive washback	The new curriculum produced both positive and negative washback	Both positive and negative washback

tendency to memorising the answers for the test suggested that they treated English as a subject to be passed in the examination but not a language to be learnt. Simultaneously, the teachers heavily relied on practising examination related tasks. As a result, this study revealed that the test had negative washback on teaching methodology, content and learning since many of the students' language learning needs were neglected in the intermediate English examination.

Watanabe (2013) - Centre Test for University Admissions in Japan

Watanabe (2013) studied the national centre test for university admissions in Japan which is an achievement test held in different places nationally. It measures students' achievement in the last year of upper secondary level. Watanabe (2013) sought the validation and impact of the centre test. This centre test determines students' admission in national public and private universities and also attempts to improve the teaching-learning process for helping students develop their communicative skills. The test consists of both writing and listening components and it is designed and produced by the national centre for university entrance examinations. Watanabe (2013) mentioned that the test was fair to assess the respondent students and provided a valid measurement of students' competence to study in universities. Moreover, this study analysed the test scores obtained by the students and found that students reached the level of required achievement and the scores predicted the candidate's success at university. In addition, the test influenced the improvement of test preparation materials at the institutes which dealt pre-college- level education. Watanabe (2013) finally concluded that the test had positive washback on the learners and teaching materials.

Kirkpatrick (2012)- Washback Effect of the English Assessment System in Bhutan

The study of Kirkpatrick (2012) examined the washback effect of the English assessment system associated with the new curriculum of secondary schools in Bhutan which was modified in 2006. The data in this research was collected from 56 EFL secondary school teachers by means of questionnaire responses. The study suggested that the new curriculum and assessment system had positive as well as negative washback. The positive washback effects found in the study were teachers' willingness to accept new approaches to teaching, diagnostic feedback along with the grading given to the learners by the teachers, use of criteria, rubrics, and checklist samples for assessment. The negative washback as reported in this study occurred due to the mismatch between syllabus contents and time allocation, huge workloads on students and teachers, and the tendency of grade inflation. Besides, the factors like lack of teacher training, lack of curriculum materials, crowded classes, and overemphasis on summative assessment also led to the negative impacts. This study recommended for keeping the balance between summative and formative assessment. Kirkpatrick (2012) mentioned that the new assessment system brought some changes and improvement of teaching and learning in secondary schools in Bhutan but it did not meet the academic, workplace and development needs of contemporary Bhutan. In addition, Kirkpatrick (2012) referred that many of the teachers were being unable to implement the new assessment system but they could assess the students with proper guidance.

Sukyadi and Mardiani (2011) - English National Examination (ENE) in Indonesia

Sukyadi and Mardiani (2011) studied the washback effect of English National Examination (ENE) which is a high stake test held at the end of secondary schools in Indonesia. Three secondary schools were selected for this study from three different levels namely high-level achiever, moderate-level achiever and low-level achiever. The participants of this research were the students and teachers from the tenth, eleventh and twelve grades. The researchers found that ENE had mentionable impacts on teachers and students. Teachers' time arrangement, their teaching materials, contents, teaching methods, strategies, ways of assessment were affected by ENE. This study found that the ENE also affected the students' learning in the classroom in which teachers made the students practice the test and enhance their test-taking skills. Sukyadi and Mardiani (2011) found the effect of washback was only on the twelfth grade English teachers' classroom teaching, but not for lower grades like the tenth and the eleventh. The finding of this research indicated that the washback effect occurred only when the students and teachers felt that the ENE preparation is an obligation. However, Sukyadi and Mardiani (2011) referred that the school of high-level achievers encountered more positive washback in the teachers' teaching and students' learning practice, while the school of moderate and low-level achievers had more negative washback, particularly in teaching materials and teaching contents. The washback varied due to the condition of the students and so the study concluded that ENE had both positive and negative washback depending on the contexts.

Critical Commentary and Conclusion

This review study covered the empirical researches on washback effects in countries like Indonesia, Iran, Vietnam, Pakistan, Japan, Bhutan and Chile. All the studied examinations were of high stakes such as university entrance examination, national examinations, IELTS and TOEFL. All the reviewed studies sought to identify the nature and scope of washback, its effects on teaching methods, teachers' and learners' attitudes, and learners' test-taking strategies. The methodologies that these studies followed were experimental design, action research, survey design and qualitative research in which the data were collected through questionnaire, interview and observation. The evidence collected in these studies showed positive attitudes of the teachers, learners' better performance in the test. Conversely, the evidence was also meant to be negative when students crammed the answer for the test and the teachers taught only the skills tested in the test. Moreover, the washback effects were sometimes mixed (both positive and negative) and this happened due to the variance of learners' levels and their contexts.

The studies bring forth several findings with regard to washback. Such as, it is seen that rather than having a direct and simple effect, washback is quite complex and elusive. So, washback is very broad and it has many variables except the test itself. Some of the factors that affected washback include teacher and student factors (e.g. beliefs, attitudes, experience, education, training, personality, teaching and learning style, etc.), textbook writers and publishers (e.g. their interpretation of exam requirements), the status of the subject being tested, resources and classroom conditions, management of practices in the schools, communication between

test providers and test users, the socio-political context in which the test is put to use, etc. In addition, although external factors such as the teacher or student factors have been identified, insufficient research has been done to reveal how they function and interact with a test to bring about the results observed.

As far as pedagogical issues are concerned, this review reveals that teachers play an important role in bringing about washback effects (either positive or negative). Therefore, they should be a central factor of any empirical study conducted on washback. They play a vital role in the pedagogical implementations of bringing about positive washback envisaged by test developers and policymakers and expected by students and parents. Teachers can learn from these washback studies and (1) be aware of the existence of washback, (2) understand the influencing factors of washback, and (3) enhance their knowledge of educational theories related to washback. In this way, they can contribute to policy making and curriculum on which assessments are developed, whether and how teachers address students' learning styles and needs, how their students respond to their instructional activities, whether and how parents are involved in students' learning practices, and equally important, which kind of socio-cultural and educational context students are situated in. By doing these teachers can enhance their capacity to produce positive washback effects.

References

- Aftab, Asma, Sabeen Qureshi and Isabel William. *Investigating the washback effect of the Pakistani Intermediate English Examination*. International Journal of English and Literature, 5, no.7 (2014): 149-154
- Alderson, Charles and Wall, Dianne. *Does washback exist?* Applied Linguistics, 14 (1993): 115.
- Andrews, S. *Washback and Innovation*, In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in Language testing*, Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers, (2004): 37-50.
- Bailey, Kathleen M. *Language teacher supervision*. New York: Cambridge University Press, 1996.
- Bailey, Kathleen M. *Working for washback: A review of the washback concept in language testing*. Language Testing, 13 (1996): 257.
- Barnes, Melissa. *The Washback Effects of the TOEFL iBT in Vietnam*. Australian Journal of Teacher Education, 41, no.7 (2016): 158-174
- Biggs, John B. *Assumptions underlying new approaches to assessment*. Curriculum Forum, 4 (1995): 1.
- Brown, Douglas H. *Language assessment principles and classroom practices*. Longman: Pearson Education, 2004.
- Cheng, Liying. *The Washback Effect of Public Examination Change on Classroom Teaching*. Unpublished PhD thesis, The University of Hong Kong, Hong Kong, 1997.
- Cheng, Liying. *Changing Language Teaching Through Language Testing: A Washback Study*. Cambridge University Press, Cambridge, 2005.
- Cholis, Hafidz and Faricha Rizqi. *Senior High School English Teachers' Perceptions on a High-stakes*

Test (SBMPTN): A Washback Study. *International Journal of Education & Literacy Studies*, 6, no. 3 (2018): 47-52. doi: 10.7575/aiac.ijels.v.6n.3p.47

Frederiksen, J. R., & Collins, A. *A Systems Approach to Educational Testing*. *Educational Researcher*, 18, no. 9 (1989): 27-32.

Hughes, Arthur. *Testing for language teachers*. Cambridge: Cambridge University Press, 1989.

Khodabakhshzadeh, Hossein, Reza Zardkanloo and Iman Alipoor. *The Effect of Mock Tests on Iranian EFL learners' Test Scores*. *International Journal of Education & Literacy Studies*, 5, no.3 (2017): 47-51.

Khoshshima, Hooshang, Amin Saed and Fatemeh Mousaei. *Exploring the Effect of Teaching Test-Taking Strategies on Intermediate Level Learners on Reading Section of Ielts; Learners' Attitude in Focus*. *Advances in Language and Literary Studies*, 9, no.2 (2018): 4-9.

Kirkpatrick, Robert. *Washback Effects of the New English Assessment System on Secondary Schools in Bhutan*. *Language Testing in Asia*, 2, no.4 (2012): 5-32.

Madaus, George F. *The influence of testing on the curriculum*. In Tanner, L.N., editor, *Critical issues in curriculum: eighty-seventh yearbook of the National Society for the Study of Education (Part 1)*, Chicago: University of Chicago Press, 1988.

Mahmoudi, Leila. *The Washback Effect of the Iranian National University Entrance Exam (INUEE) on Pre-University Students' English Learning Process*. *British Journal of English Linguistics*, 3, no.3 (2015): 34-49.

McNamara, Tim. *Language Testing*, Oxford University Press, Oxford, 2000.

Messick, Samuel. *Validity and Washback in Language Testing*. *Language Testing*, 13(1996): 241-256.

Messick, Samuel. *Validity*. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.) New York: ACE and Macmillan. (1989):13-103.

Munoz, Rodrigo Esteban. *The Effect of Washback on EFL Learners' Attitudes toward Tests*. *Studies in English Language Teaching*, 5, no.3 (2017): 516-530.

Palmer, Adrian S. and Bachman, Lyle F. and. *Language Testing in Practice*. Oxford University Press, Oxford, 1996.

Pearson, Ian. *Tests as levers of change (or putting first things first)*. In D. Chamberlain & R. Baumgartner (Eds.), *ESP in the classroom: Practice and evaluation* *ELT Documents #128*, (pp. 98-107), Modern English Publication in association with the British Council, London, 1988.

Popham, James W. *The Merits of Measurement-Driven Instruction*. *Phi Delta Kappan*, 68 (1987): 679-682.

Saville, Nick. D. *Developing a Model for Investigating the Impact of Language Assessment with Educational Contexts by a Public Examination Provider*. Bedfordshire: University Bedfordshire, 2009.

Elana Shohamy, Smadar Donitsa-Schmidt, Irit Ferman. *Test impact revisited: Washback effect over time*. *Language Testing* 13 (1996): 298.

Smith, Mary. L. *Put to the test: The effects of external testing on teachers*. Educational Researchers, 20, no. 5, (1991): 8-11.

Spolsky, B. *The examination of classroom backwash cycle: Some historical cases*, in Nunan, D, Berry, V and Berry, R. (Eds) *Bringing about change in language education*, University of Hong Kong, Dept. of Curriculum Studies, Hong Kong, 1994.

Sukyadi, Didi and Ridha Mardiani. *The washback effect of the English National Examination (ENE) on English teachers' classroom teaching and students' learning*. Kata, 13, no.1 (2011): 96-111.

Taylor, Lynda. *Washback and impact*. English Language Teaching Journal, 59 (2005): 154-155.

Wall, Dianne. *Impact and washback in language testing*, In C. C. & D. Corson (Eds.), *Encyclopaedia of language and education, Language Testing and Assessment* 7 (1997): 291.

Watanabe, Yoshinori. *The National Center Test for University Admissions*. Language Testing, 30, no.4 (2013): 565-573.