

Washback in Language Testing: An Exploration with a Focus on a Specific EFL Context in Oman

Mohammed Khamis Al Hinai¹ & Khalid Salim Al Jardani²

¹ Oman Tourism College, Sultanate of Oman, Oman

² A'Sharqiyah University, Sultanate of Oman, Oman

Correspondence: Khalid Salim Al Jardani, A'Sharqiyah University, Sultanate of Oman, Oman. E-mail: kaljardani@hotmail.com

Received: September 15, 2020 Accepted: October 14, 2020 Online Published: October 31, 2020

doi:10.5539/ijel.v11n1p68 URL: <https://doi.org/10.5539/ijel.v11n1p68>

Abstract

The notion of washback has been a widespread topic in language teaching and testing literature and has been discussed and looked at from various perspectives. The significance of the topic emerges from the fact that the consequences of tests greatly determine or shape the nature of language teaching, learning, and testing.

This paper aims to explore the notion of washback based on current perspectives in the literature and the needs of a specific teaching setting, which is the language Program at Oman Tourism College (OTC). The Literature Review section reviews the literature trying to shed light on the notion of washback as a core construct in language teaching and learning. The second section will discuss the influences of washback in ELT with a special reference to a teaching context. Finally, some useful implications are drawn to come up with recommendations for increasing the positive washback and reducing the negatively of it.

Keywords: washback, Oman, Oman Tourism College (OTC), ELT, EFL

1. Literature Review: Washback in Language Testing

1.1 Definition and Historical Perspective

Various definitions have been given to the term *washback* ranging from simple to very complex. Bailey (1996, p. 256), for example, defines washback as “the influence of testing on teaching and learning”. Messick (1996, p. 243) defines washback as “the extent to which the text influences language teachers and learners to do the things that they would not necessarily otherwise do”. A more complex and elaborated definition of washback is provided by Buck (1988) when talking about the influence of the entry tests of the Japanese Universities on English language learning in Japan:

“There is a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to the future of the students, and pass rates are used as a measure of teacher success. This influence of the test on the classroom (referred to as washback by language testers) is, of course, very important; this washback effect can be either beneficial or harmful” (p. 17).

In an attempt to describe *washback* based on different factors, Alderson and Wall (1993) have developed 15 hypotheses about the influences of tests which will be discussed later. Shohamy (1992) describes washback in terms of its influence on test-takers by arguing that it is the “utilization of external language tests to affect and drive foreign language learning in the school context” (p. 513). Bachman and Palmer (1996) describe washback based on the levels of the influences of tests. They discuss that the impact of tests works at two different levels; the micro level, which is the effect of tests on individuals and the Macro level, which is the impact on the society and the educational system.

Providing historical perspective of the notion of washback, Wall (2013) states that the discourse regarding the washback of high stakes testing only appeared in the field of language testing in the early 1990s. She adds that before this time, there were only general claims about the impact of tests on the curriculum and there was uncertainty whether this influence had really existed. Cheng and Curtis (2004) refer to an earlier time at which the concept started to emerge. They report the contributions of several scholars such as Alderson (1986), Davies

(1985) and Morrow (1986). They state that Alderson, for example, discussed the potential powerful impact of tests and called for innovations in the language curriculum by using the outcomes of language testing. They also claim that Davies had suggested that curriculum should be led and influenced by tests. As for Morrow, they discuss that he included the notion of *washback validity* to explain the relation between testing and both teaching and learning.

1.2 Validity and Washback

Much has been written on the relation between washback and validity. However, the common assumption among many scholars is that, according to Alderson and Wall (1993), a test is considered valid if it has a positive washback and invalid if it has a negative washback. Cited in Bailey (1999, p. 9), Boyle and Falvey (1994, p. 11) argued that validity is “one of the Big Four considerations in evaluating the worth of a test”. The relation between washback and validity, as discussed by Wall (1993), was intensively explored by Messick (1996), who reviewed the concept of construct validity and argued that washback was one of the significant aspects of validity, and hence establishing test validity requires investigating the potential washback. Explaining this further, Anderson and Wall (1993) advise that it is crucial to analyze the test design if there was no link observed between the test and the teaching or learning outcomes in order to see whether there are any adjustments that need to be made.

1.3 Positive and Negative Washback

We have seen so far that washback is the impact of tests on teaching and learning. Depending on the nature of test influences, washback can be either positive or negative. This is clearly explained by Alderson and Wall (1993, p. 117), who claim that “tests can be powerful determiners, both positively and negatively of what happens in classrooms”. However, the concept of positive washback is highlighted by Loumbourdi (2014):

“We could say that if the changes caused to methodology, the curriculum and the whole process of teaching by the washback are beneficial, then we have **positive** language learning in an authentic context and a less stressful environment, attention to students’ individual needs and learning for the sake of learning and not the test” (pp. 20–21).

Cheng (1998) cited in Loumbourdi (2014, p. 26), asserts that

“Generally, a test could theoretically produce positive washback if there is a correlation between the content of the test and the actual washback. When we refer to beneficial changes, we mean the promotion of world”.

Wall and Alderson (1993, p. 46) proposed some conditions under which beneficial washback could be achieved by stating that “If there were no conflicts in the aims, activities, or the marking criteria of the textbook and the exam, and if teachers accepted these and worked towards them, then a form of positive washback could be assumed to have occurred”. Cheng and Curtis (2004) stress that “...it is feasible and desirable to bring about beneficial changes in teaching by changing examinations, representing the “positive washback” scenario, which is closely related to “measurement-driven instruction” in general education” (p. 10). They add that this would enable teachers and learners to establish a positive attitude toward exams and would be more willing to work collaboratively toward their objectives.

Describing the negative washback, Loumbourdi (2014, p. 21), declares that “...if the effect of a particular test and the preparation for is considered to be harmful, leading to old-fashioned or ineffective techniques and narrowing of the curriculum, then we refer to negative washback”. Looking at the negative washback in terms of the anxiety tests bring to learners and teachers Alderson and Wall (1993) note:

“The most obvious such effect is anxiety in the learner brought about by having to take a test of whatever nature, and, if not anxiety, then at least concern in teachers, if they believe that some consequences will follow on poor performance by the pupils” (p. 117).

The way the negative influence of tests on teachers operates is discussed by Alderson and Wall (1993, p. 118) when they acknowledge that “the fear of poor results, and the associated guilt, shame, or embarrassment, might lead to the desire for their pupils to achieve score whatever way seems possible”. They add that in such situation, teachers will find themselves teaching toward the test or what is so called “test-driven instruction”.

We have seen so far that washback is a complex phenomenon and taking into account the huge range of integrated and complex factors surrounding washback, it seems quite difficult to determine how tests actually influence teaching and learning. This issue is raised by Wall (2013, p. 80), who argues that “The idea of washback takes on more complexity when we consider not only whether the effects of tests are positive or

negative but also whether they are immediate or delayed, direct or indirect, or apparent or not visible". The complex effect of tests could also be clearly seen when looking at Alderson and Walls' washback hypothesis.

2. Description of Teaching Context

English is the medium of instruction in most of the Oman's institutions. It is taught for both specific and general purposes (Al Jardani, 2017). This section describes teachers, learners and the English program utilized in one of Oman's higher education institutions, *Oman Tourism College*.

2.1 Teachers

Along with Omani teachers, the Language program at *Oman Tourism College* comprises teachers from different nationalities including Native English Speakers. Their educational backgrounds vary and most of them have long teaching experiences as well as high teaching qualifications. Coming from different teaching backgrounds, teachers have different beliefs about teaching and learning which are reflected in their own teaching practices. This is reinforced by the fact that there are not specific teaching techniques used as long as the pre-set learning outcomes are successfully achieved. Based on our experience as a *Program Coordinators* through which we had an opportunity to observe teachers in their classes, it could be noticed that there were different teaching strategies and techniques employed by teachers. For example, some teachers would mainly teach towards the tests, especially since the tests are written two weeks in advance. However, the heavy workload on teachers and the other ongoing teaching-related duties sometimes contribute to establishing an atmosphere of tension that, by away or another, would re-shape their teaching attitudes, beliefs and behaviors. This, for instance, is observed in their quick or superficial teaching, especially when exams period approaches.

2.2 Learners

The foundation students at *Oman Tourism College* are mainly secondary school graduates who have completed 12 years of EFL program at public schools. They are admitted to the language program and placed on either level 1 or level 2 based on their grades in the placement test. Generally, these students come with different learning abilities but many of them have insufficient language proficiency due to deeply rooted learning problems. Their speaking and writing skills, for instance, are far below the required levels. Their knowledge of the English grammar is also low and they tend to face difficulties in understanding and using it. Based on the Omani educational context, such situation seems to be an outcome of factors like the absence of a positive attitude among many students towards learning English. Furthermore, many students appear to be de-motivated, which is due to factors like the absence of a clear goal behind studying and being unable to adjust to the new learning setting, which requires much more effort and commitment than in the primary education environment.

The majority of students come from rural areas and from cultural backgrounds where most of the parents are not well educated. They are admitted to the college through fully covered governmental scholarships. Generally, the students are more likely to be more exam-focused and directed as their potential target is passing the tests without paying adequate attention to what they are actually expected to achieve in terms of language skills and knowledge. As a matter of fact, such mind-set seems to create some undesirable long-term effects as the students proceed to their undergraduate studies.

2.3 The General English Program

The general English program is designed to equip the students with the English language skills and strategies they require for their next undergraduate studies. As mentioned earlier, the program consists of two study levels: level 1 (semester 1) and level 2 (semester 2). The students are placed in either level based on their grades in the placement test. The program is delivered through in-house teaching materials that are regularly revised and amended. Supplementary teaching materials are used to support the students' learning. The language skills are taught separately where in each week there are certain pre-set learning outcomes to be achieved. The delivered teaching materials develop as the students' progress throughout the program.

Assessment takes place at pre-determined times throughout the program and conducted through formative and summative tests. Except for speaking, in level 1, there are two formative tests for each skill and one summative exam that takes place at the end of the program. A similar assessment procedure is applied for level 2 except for writing and ESP in which the formative assessment is conducted through two short tests and one assignment. Before each test, the students are given a *Mock Test* in order to familiarize them with the actual test questions and format which is an idea that has been recently introduced as a way of dealing with the students' poor performance in the exams. The students also can take re-sit test for each assessment component if they cannot obtain the pass grades.

3. Influence of Tests on Teaching and Learning

3.1 Educational Influence

This section explores the influence of tests on teaching and learning by focusing on a specific teaching context, the language program at *Oman Tourism College*. It should only focus on the most four common aspects influenced by washback.

3.1.1 On Teachers

The influence of tests on teachers can be seen from different perspectives. The most obvious aspect is related to the teachers' own teaching practices. This is emphasized by Baily (1996), who discusses that "the most visible participants in program washback are language teachers. It is they who are the "front-line" conduits for the washback process related to instruction" (p. 17). Among their 15 hypotheses about washback, Alderson and Wall proposed 6 hypotheses that involve teachers and teaching (see Alderson & Wall, 1993, pp. 120–121).

It is argued that most of Alderson and Walls's influences concerning teachers can be clearly observed in our own teaching context. In terms of *what* to teach, it is obvious that is difficult to talk about a change in the taught content as it is pre-determined. Based on personal observation, some teachers would adjust their teaching styles and techniques based on the students' performance in tests. For example, concerning my institution, OTC, in order to increase the students' understanding of grammar, teachers would provide more explanation and set more oral production tasks. The test influence on the degree and depth of teaching is noticeably present in the increased amount of explanation given to OTC foundation students, especially in grammar which seems to represent a common challenge to many students. Teachers sometimes tend to teach towards the test to ensure that the materials intended to be tested are covered.

The influence of tests on the teachers' feelings is examined in terms of the potentials fear, anxiety, shame and embarrassment they could possibly bring to teachers (Shohamy, 1996; Spratt, 2005). The consequences of such feelings are explained by Shohamy (1996), who asserts that "an atmosphere of high anxiety and fear of test results among teachers and students. Teachers feel that the success or failure of their students reflects on them and they speak of pressure to cover the materials for the exam" (pp. 309–310). To examine these influences within our teaching context, it is argued that the above-mentioned atmosphere of fear and anxiety takes place before and after conducting the tests. The need to submit the test to the program administration as per the stipulated deadlines is sometimes a source of tension over teachers. After the test, there is a concern among teachers that their students may not achieve the desirable results. This concern is strengthened by the deeply rooted assumption that teachers are mainly to be blamed for the students' failure or low performance. This whole situation may lead to the teachers being demotivated or losing self-confidence which I could notice when, for example, some teachers would request to be exempted from teaching certain skills.

3.1.2 On Learners

As its discussed in 2.2, students are mainly exam focused as they assume that successful learning simply means passing exams. This is aligned with the situation Loumbourdi (2014, p. 24) describes as "stressful environment and tension" emerges. Such environment seems to exist throughout the course, especially during the exams period and naturally leads to anxiety among the students. Closely watching our students' reaction to exams, the most frequent anxiety-related behaviors and attitudes that it could be noticed was:

- 1) Skipping classes to study for the exams
- 2) Approaching teachers for a last-minute tutorial
- 3) Asking teachers about what will be included in the test
- 4) Providing excuses for not being ready for the test such as sickness or family matters
- 5) Putting the blame on teachers for their failure or poor grades

Concerning *what* and *how* the students learn, it sounds difficult to make a definite claim. However, based on our observation of OTC students, it could be said that there were generally signs of positive washback as some students started to study harder and apply better learning strategies such as establishing study groups among themselves.

3.1.3 On Curriculum/Teaching Materials

It is natural that the curriculum and teaching materials in any language teaching setting needs to be regularly revised and amended based on the learners' needs which are addressed by assessment. The impact of tests on curriculum and the content of teaching materials is highlighted by Spratt (2005), who reports the findings of

various studies of washback effects. For example, he refers to the study of Cheng (1997) which concluded that the content of teaching was changed as a result of introducing revised exams in which reading aloud activities were replaced by role play and discussion activities. He also discusses the findings of Lam (1994) which indicated that there was more emphasis on teaching the areas of exam that carried the most mark. Spratt (2005, p. 10) talks about how tests influence the choice of teaching material which he describes as “highly exam techniques oriented” which may include previous test books and papers.

Since 2011, the English language curriculum at *Oman Tourism College* and all associated teaching materials have undergone a series of changes of which many were due to the needs addressed by exams. The most significant change was adjusting the teaching materials in terms of depth and intensity. The content of certain skills has been either reduced or increased which required as well adjusting the number of weekly teaching hours allocated for each skill. Supplementary teaching materials have been also introduced as a way of reinforcing the students’ learning in the skills in which they did not appear to achieve good grades. Furthermore, special remedial classes have been introduced to support week students based on their attainment in the first summative assessment (short test 1), which also entailed designing additional teaching materials.

3.1.4 On Assessment

The effect of tests on assessment mainly operates in indirect ways. Taking, for example, into account washback hypotheses proposed by Alderson and Wall, we can easily understand this indirect influence. For instance, the hypothesis which says: “A test will influence what teachers teach” (p. 121), assumes that a test can result in making changes to the existing teaching materials. Consequently, the upcoming tests will be modified based on the new materials. So, we can say that it is an ongoing cyclical process.

Alongside the indirect influence of test on assessment, there is also the direct influence through which the test’s process and procedures are regularly reviewed and modified which could be observed in our teaching context. Aspects such as content, length, mark allocation and test weights are regularly revised based on the students’ results. The assessment plan discussed in 2.2 is the latest version that shows some slight changes that have been recently made. For example, as it is mentioned in 2.2 mock tests were introduced to familiarize the students with the content of the actual tests. Also, tests weights have been changed by increasing the weight of the final examination from 60% up to 70%. Marking is also another aspect influenced by tests. In order to ensure objectivity and fairness and avoid any potential bias, the students’ test papers are no more marked by their respective skill teachers.

3.2 Political Influence

This is another aspect of washback effect which occurs within what Bachman and Palmer (1996) refers to as a “macro level”. When we talk about the political influence, we mainly refer to the decisions made by the policy makers based on the significance of issues addressed by exams. Shohamy (1996, p. 299) affirms that “the power and authority of tests enable policy makers to use them as effective tools for controlling educational systems and prescribing the behavior of those who are affected by their results-administrators, teachers and students”. As a matter of fact, many of high stakes decisions taken by the policy makers come as a response to various needs addressed by exams such as curriculum innovations and recruitment or replacement of teaching staff and providing resources.

Concerning the educational context, the aspect of washback occurs within three levels of authority: the program management, the academic management, and the top management. However, being the direct supervising authority of the language program, the program management is the primary source of any potential change in the language program. Given Shohamy’s claim above, some political influences of tests could be observed. These are reflected in our teaching context through various decisions such as excluding some teachers from teaching specific skills or terminating the service of certain teachers. Moreover, there are other influences concerned with the program structure such as curriculum development, review of program length, review of entry requirements and review of the program delivery plan. However, although these procedures are considered academic in nature, they have a political dimension as implementing them requires the availability of human and financial resources which, in most cases, cannot happen without an approval from the top educational authorities.

3.3 Social Influence

No doubt that tests have social impacts that exceed the institutional environment affecting parents, families, and the whole society. Within the society, as discussed by Shohamy (2001, p. 39), test “...symbolize order and discipline and are perceived as an indication of effective learning”. Hence, they are considered to be a source of tension and anxiety for parents and families who have certain expectations about their children (Hamp & Lyons,

2000), and consequently they tend to put some pressure on them. However, the strength and degree of test impact on parents is mainly determined by the significance of the test. High stakes tests, for example, normally generate a huge atmosphere of tension among parents as they represent important milestones in their children's educational journey.

Examining the social influence of tests within our educational context, it could be claimed that no tangible influences of tests on the students' parents have actually been observed which should assumingly be reflected in their regular visits to the college asking about their children's progress in their studies. This could be due to the lack of adequate education among many parents that it is discussed in 2.2 and which probably makes them uninformed of the importance of exams in determining the future of their children.

3.4 A Considerable Issue

Having explored the areas through which tests affect teaching and learning in my educational context, OTC, the significant inquiry is: *Is washback possibly determined, shaped or influenced by certain factors?* As a matter of fact, the answer would definitely be *YES*. No doubt that there is a strong mutual effect between assessment and teaching and learning. If we claim, for example, that tests influence the way learners approach their studies either positively or negatively, then it is inevitable that their performance in future exams will either improve or decline.

Factors affecting washback are highlighted by Spratt (2005) in his review of some empirical studies of washback. He puts factors that affect washback within three main categories: teacher-related factors, resources, exams and school. Regarding our teaching context, it could be argued that these factors are strongly present in addition, of course, to learner-related factors, which play a crucial role in affecting washback. One example of how these factors influence washback is what was discussed in 2.1 concerning the fact that some teachers tend to teach towards the test as of providing support to the students. This, however, would result in positive washback as many students would probably perform well in the test. It could be also considered negative since some of the content materials will be skipped.

4. Implications and Recommendations

We have already seen that washback is a complex phenomenon and occurs within highly interconnected variables (Spratt, 2005). Given all this, it sounds irrational to talk about successful solutions in dealing with the potential drawbacks of tests, especially since every single teaching and learning setting has certain variables that shape its own identity. However, the primary action should focus on maximizing the positive washback and minimizing the negative one through careful diagnosis of the characteristics of the most influencing factors. This should involve looking at each factor from a holistic perspective rather than treating it as an independent construct.

Since learners are the key aspect in any learning or teaching setting, learner-related factors should probably be the starting point towards any desired change. The primary focus should be on re-shaping their attitudes and feelings about learning (Shohamy, 2001) and reducing fear and anxiety that they are likely to develop. This involves considering the relation with other dimensions: teachers, curriculum, teaching content, and assessment. Alderson and Wall (1993, p. 118), advises that "We may also wish to consider the possibility of a test *reinforcing* some behavior or attitude rather than bringing about an otherwise unlikely behavior". Harlen and Crick (2003) propose some suggestions for increasing the positive washback of tests such as promoting more learner-centered teaching approaches and enhancing the students' understanding of the objectives of their learning. Furthermore, Bailey (1996) and Loumbourdi (2014) stress on the importance of using authentic teaching and assessment tasks in promoting more beneficial washback. However, it is worth mentioning that the options for increasing the desirable influences of tests remain unlimited. What really matters is the extent to which changes could be implemented based on the interference of the existing contextual variables.

5. Conclusion

This paper has sought to explore washback as a crucial aspect of language teaching and testing by bringing about both the discussion in the literature and addressing the needs of a specific EFL context. Several aspects of washback influences were examined in relation to certain factors. Finally, some useful implications for prompting positive washback were made.

It is obvious that washback will remain a complex phenomenon and will continue to stir inquiries about the highly complicated relation between testing and teaching and learning. Hence, any future investigation of this phenomenon should take into account the special characteristics of the target teaching and learning context with its all variables including the local society. Furthermore, suggestions for promoting the beneficial washback should be made based on a clear understanding of the relation between the different existing contextual factors.

References

- Al-Jardani, K. S. (2017). English education policy in Oman. In *English language education policy in the Middle East and North Africa* (pp. 133–146). Springer, Cham. https://doi.org/10.1007/978-3-319-46778-8_9
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115–129. <https://doi.org/10.1093/applin/14.2.115>
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests* (vol. 1). Oxford University Press.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257–279. <https://doi.org/10.1177/026553229601300303>
- Bailey, K. M. (1999). *Washback in language testing*. Princeton, NJ: Educational Testing Service.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10(1), 15–42.
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38–54. <https://doi.org/10.1080/09500789708666717>
- Cheng, L. E., Watanabe, Y. E., & Curtis, A. E. (2004). *Washback in language testing: Research contexts and methods*. Lawrence Erlbaum Associates Publishers. <https://doi.org/10.4324/9781410609731>
- Hamp-Lyons, L. (2000). Fairness in language testing. Fairness and validation in language assessment. *Studies in Language Testing*, 9, 30–34.
- Harlen, W., & Deakin, C. R. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice*, 10(2), 169–207. <https://doi.org/10.1080/0969594032000121270>
- Loumbourdi, L. (2014). *The power and impact of standardised tests: Investigating the washback of language exams in Greece*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften. <https://doi.org/10.3726/978-3-653-03958-0>
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241–256. <https://doi.org/10.1177/026553229601300302>
- Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76(4), 513–521. <https://doi.org/10.1111/j.1540-4781.1992.tb05402.x>
- Shohamy, E. G. (2001). *The power of tests: A critical perspective on the uses of language tests*. Pearson Education.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298–317. <https://doi.org/10.1177/026553229601300305>
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29. <https://doi.org/10.1191/1362168805lr152oa>
- Wall, D. (2013). Washback. In *The Routledge handbook of language testing* (pp. 93–106). Routledge.

Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).