

The Effect of Washback on EFL Learners' Attitudes toward Tests

Rodrigo Esteban Muñoz^{1*}

¹ Facultad de Ciencias de la Educación, Universidad San Sebastián, Puerto Montt, Chile

* Rodrigo Esteban Muñoz, E-mail: rodrigomunoz.1983@gmail.com

Received: July 18, 2017

Accepted: July 25, 2017

Online Published: July 31, 2017

doi:10.22158/selt.v5n3p516

URL: <http://dx.doi.org/10.22158/selt.v5n3p516>

Abstract

The purpose of this action research was to improve learners' attitudes towards tests through the implementation of positive washback. This article compares the results obtained by two groups of learners in a secondary school in Southern Chile on the Attitude toward Test Scale (Dodeen, 2008). During the intervention, tests were followed by washback activities for the experimental group. Lessons in which positive washback was sought were observed using a focused observation table in order to track the techniques used. Finally, an in-depth follow-up interview was conducted in order to obtain deeper insights. After implementation, results indicated that the experimental group evidenced a statistically significant improvement in their attitude towards tests.

Keywords

Attitudes toward tests, EFL, tests, washback

1. Introduction

Feelings of anxiety and self-doubt seem to be inherent in learners of a second language during the process of taking a test (Horwitz, 2001; Würde, 2003; Cohen, 2001). This relates to Hughes' (2003, p. 1) argument that "tests have a harmful effect on teaching and learning and fail to measure accurately whatever it is they are intended to measure". Despite this, tests do not need to be negative experiences but they can "build a person's confidence and become learning experiences" (Brown, 2004, p. 2). According to Brown (2004), as learning experiences, tests should not only define how students prepare for the assessment, but also identify weaknesses and strengths on the part of learners, that is, to diagnose them. These desired effects were documented in his five principles of language assessment under the name of *Washback* and are aligned with Cheng and Curtis' (2008) vision of good tests as learning activities. The effects of negative or unintended washback have been reported in several studies (Cheng, Watanabe, & Curtis, 2008; Andrews, 2008). Additionally a quick review of the available studies on the concept seem to indicate there is enough evidence of high-stakes examinations

influencing teaching (Barnes, 2016); this leads to the conclusion that there is apparently a lack of documented evidence on the impact of positive or intended washback on low-stakes examinations, a void for which this study should supply more information. The objective of this action research was to observe if emphasizing the positive washback functions of summative assessments would improve learners' attitudes toward testing itself in the Chilean context. In order to do so 52 students from a secondary school in Southern Chile underwent the following process: 25 learners belonging to the experimental group and 27 learners belonging to the control group answered the Attitude towards Test Scale (ATS) (Dodeen, 2008) at the beginning and at the end of the instruction period. Between both implementations of the ATS, summative assessments were followed by positive washback activities during the intervention phase of the study.

The research questions underlying this action research are the following:

- a) How do learners score on the ATS prior to an instruction period using washback techniques after summative assessment?
- b) How do learners score on the ATS at the end of an instruction period using washback techniques after summative assessment?
- c) What is the relationship between the use of washback techniques and the learners' attitudes toward summative assessment?

2. Review of Literature

2.1 EFL

In the Chilean context, English is taught as a Foreign Language (EFL). According to Thornbury (2006), EFL takes place in societies where "English is not the usual language for communication" (p. 74). For the students involved in this action research, English is a subject at school and all of them speak Chilean Spanish as their first language. This situation makes testing a normal occurrence in their involvement with English. As opposed to contexts where English is taught as a Second Language (ESL), learners' exposure to the target language is limited mainly to the classroom. In this sense, the teacher's role becomes crucial in order to compensate for the aforementioned disadvantages. According to Brown (2000), among the many ways in which a teacher can enhance the EFL context is to "play down the role of tests and emphasize more intrinsic factors" (p. 117) which can be interpreted as raising positive washback.

2.2 Language Anxiety

Horwitz (2001) contends that learners' anxiety when acquiring a second language usually has a negative impact on achievement. Language anxiety can be divided into trait anxiety (stable) and state anxiety (situation-specific) (Salehi & Marefat, 2014). Language learning anxiety may originate from three sources: communication apprehension, test anxiety and fear of negative evaluation (Horwitz, E. K., Horwitz, M. B., & Cope, 1986). As the objectives of this study are focused on students' attitudes in connection with summative assessment, only two of the previously mentioned sources of anxiety will

be addressed: test anxiety and fear of negative evaluation. In terms of test anxiety, studies have shown that “individuals who are test-anxious become more obsessed with the implications and consequences of failure to meet situational challenges rather than rationally focusing on completing the task in an orderly manner” (Arezou, Rusnani, Habibah, & Maznah, 2012, p. 4).

2.3 Washback

Washback has been defined as “the influence of testing in teaching and learning” (Gates, 1995, p. 101) or “the way [tests] affect the classroom teaching that leads up to them” (Thornbury, 2006, p. 228). As stated by Zhang (2016), “There is a general consensus that language testing can exert an impact or influence on all test stakeholders” (p. 779). Such an influence may be positive or negative. The negative influence of washback on teaching and learning has been widely documented. For instance, Prodromou (1995) mentions how teachers may feel “trapped in the examination preparation cycle” (p. 14) and how such situations may sacrifice communicative or humanistic methodologies in favor of the proper covering of the syllabus. Additionally, Pan (2009) highlights how activities that are not directly connected with the passing of an exam may be dismissed by teachers who end up focusing on the teaching of test-taking skills rather than language learning activities. Similarly, Barnes (2016) reports on the influence of high-stakes exams on teaching and learning strengthened by the use of test preparation materials. The situations previously described clearly have a negative effect on the teaching and learning process as the need to teach to the test may also cause a lack of follow-up or lead-in activities in connection to tests (Prodromou, 1995).

On the other hand, washback may also have a positive effect on teaching and learning. According to Brown (2002), positive washback may be fostered by altering factors of test design, changing factors of test contents, adjusting factors of test logistics and modifying factors of test interpretation. Additionally, Brown (2004) mentions another form of positive washback as the “information that ‘washes back’ to students in the form of useful diagnoses of strengths and weaknesses” (p. 29). In this sense, good tests will affect the teaching-learning process in a positive way, motivate teachers as well as learners and encourage the idea of lifelong learning (Pan, 2009).

Examining Brown (2004) and Prodromou (1995), one can conclude some techniques in order to achieve positive washback inside the classroom: praising correct answers, suggesting strategies for success, commenting generously on test performance, criticizing constructively for weaknesses, being accessible for discussion on evaluation and giving learners the chance to ask questions about their tests. All of the aforementioned techniques were implemented in the sessions in which students received their marks.

As previously stated, the intention of this action research was to identify learners’ attitudes towards tests before and after being exposed to these washback techniques. The importance of learners’ attitudes regarding tests has been emphasized as they may affect achievement (Dodeen, Abdelfattah, & Alshumrani, 2014). This acquires even more relevance if we take into account Cheng and Curtis’ (2008) position, as test results are undoubtedly influenced by the strategies learners use and the conclusions

that may be drawn from test results. As washback helps enhance factors such as “intrinsic motivation, autonomy, self-confidence, language ego, interlanguage, and strategic investment” (Brown, 2004, p. 29) it can be concluded that learners’ attitudes towards tests should be enhanced if positive washback is sought alongside the implementation of tests.

3. Methodology

While planning this action research, attention was placed on Watanabe’s (2008) guidelines regarding the design of washback studies. In this respect, this action research can be considered “specific” (p. 20) in the sense that it is limited to classroom low-stakes tests of EFL and one specific aspect, that is, the activities carried out after the graded tests had been returned to the learners. Additionally, data was also gathered in connection with participants’ feelings “about the effects of examinations” (p. 23). Furthermore, classroom observation was not the sole method of data gathering to be implemented, since a group follow-up interview and the ATS replaced the recommended “interviews and questionnaires” (p. 23). Finally, the observation table used during the classroom observations was adapted from an existing instrument called “Classroom observation table B” (Tanner & Green, 1998, p. 52), something supported by Watanabe’s (2008) guidelines as he foresaw the necessity to “modify an instrument that is available” (p. 30). The modification previously mentioned was reviewed by experts in order to ensure the validity and reliability of the instrument.

3.1 Design

This action research relied on a mixed approach. Quantitative data was collected from the tabulation of learners’ answers on the ATS in the Pre and Post-Test. It was expected that learners’ attitudes measured by the ATS would improve in the interim due to the implementation of the washback techniques. Additionally, qualitative data was collected through the observation and recording of the lessons in which the washback techniques took place as well as a follow-up group interview that was carried out after learners took the Post-Test. The reason for the use of this mixed approach was to allow a more accurate triangulation of the data during data analysis.

3.2 Participants

Two sections of 4th year high school students from a semi-public school in southern Chile were selected for this study. Students were between the ages of 17 and 18. The experimental group consisted of 15 females and 10 males; whereas, the control group consisted of 15 females and 12 males. Both groups received four hours of English instruction per week and have had the same instructor during the last two years prior to this action research.

This school was selected because it does not usually consider washback or feedback activities after applying summative assessments. This problem was perceived by the researcher and supported by learners attending the follow-up group interview at the end of the research. The school’s vision is aimed at high-stakes tests with little emphasis on feedback in most of its subjects and that implies that teachers usually need to cover extensive amounts of content that are subsequently assessed through

close-ended tests. This context means that learners are generally unfamiliar with washback and feedback activities and consider tests as judgments and not learning experiences. These claims were evidenced in the follow-up group interview carried out at the end of the research.

3.3 Instruments

In order to measure learners' attitudes regarding tests, both the control and the experimental group answered the Attitude towards Test Scale (ATS) (Dodeen, 2008) (see Appendix A) prior to the instructional period and at its conclusion. This survey consists of 17 questions that are answered using a Likert scale from "Strongly Disagree" to "Strongly Agree". Some examples of the questions included on the ATS are "Tests motivate me to study hard", "Tests help me identify my academic weaknesses", "During tests, I learn many useful skills" and "I think taking tests is a useful experience for me". Numerical results on the ATS reflect learners' attitudes in relation to tests as "a high score on ATS suggests positive attitudes towards tests" (Dodeen, Abdelfattah, & Alshumrani, 2014). This instrument was selected because it has already been applied in at least two studies (Dodeen, 2008; Dodeen, Abdelfattah, & Alshumrani, 2014) measuring learners' attitudes. Despite its use in studies related with mathematics, it also has application across the disciplines. The instrument was originally implemented in Arabic, although there is also an English version that was translated by Dodeen for the purpose of this study. This version was subsequently translated into Spanish (see Appendix B), and validated by experts in the field, in order to be implemented for this action research. Additionally, question 13 was adapted for the particular context of this study and was rendered: "I try to avoid courses that have many tests". However, the study's two cohorts have very limited choices when it comes to course selection. Therefore, question 13 was replaced by "Si tuviese la oportunidad, evitar á las asignaturas en las que se rinden muchas pruebas". (If I had the chance, I would avoid taking courses that have many tests) as it did not change the intention of the question and better suited the learners' actual context.

The objective of this action research is rooted in Watanabe's (2008) perspective that specific washback should make learners emphasize certain aspects of a test in their learning process. In this sense, if different aspects of tests and the strategies needed to succeed in them are made explicit for learners, then their perception of the educative use of the tests should consequently improve.

Additionally, lessons in which the washback techniques were implemented were videotaped and observed by another teacher. The observation was carried out through the use of a chart adapted from Tanner and Green's (1998) "Classroom observation table B" (p. 52). This instrument was originally designed to focus on the use of transitions within lessons of English. Due to a lack of washback specific charts, an adaptation was made in order to focus on the use of washback techniques (see Appendix C).

3.4 Procedures

Over a four month period both groups received distinct treatment in terms of the activities carried out once their tests were graded and returned. The control group experienced the traditional practice of receiving their assessment results with little or no student input beyond score clarification. The

experimental group experienced the methodology carried out in the research, that is, the instructor implemented the washback techniques during lessons in which learners were handed back their tests and written quizzes. The aforementioned assessment instruments were projected on the board and expected answers were discussed with learners. In doing so, the researcher carried out the strategies previously specified: comments on test performance, praising correct answers, suggestions of strategies for success, constructive criticism of weaknesses, and giving learners the chance to ask questions about their own performance. As these sessions were taking place, another teacher observed the use of the strategies and recorded the performance by using the Class Observation Table on Appendix C. Prior to the first session of activities, learners of both groups responded to the ATS. This was then at the close of the intervention. Results on the Pre and Post test were analyzed through the use of SPSS software in order to obtain averages, standard deviation, graphs and charts.

4. Results and Discussion

Results of this action research were analyzed from the research questions posed at the beginning of this study. In terms of the first research question, “*How do learners score on the Attitude towards Test Scale prior to an instruction period using washback techniques after summative assessment?*” The results are as follow:

Scores ranged from 0 (negative attitude towards tests) to 5 (positive attitude towards tests) in each question. As the ATS survey consisted of 17 questions, a score closer to 85 indicated a positive attitude towards tests. Both groups scored similarly on the Pre-test. This is evidenced by the total average and standard deviation obtained by both groups. This appears to imply that both groups had a similar attitude towards tests at the beginning of the intervention period. These scores (55.9 and 57.4) can be interpreted as a moderately positive attitude. Regarding their Standard Deviation, both groups’ results can be interpreted as highly dispersed attitudes within each cohort.

Table 1. Total Average and Standard Deviation in the Pre-Test

	Mean	Standard Deviation
Control Group	55.9	7.8
Experimental Group	57.4	10.8

Additionally, both groups scored similarly in most of the questions included in the survey. The most significant difference was present in questions 5 and 8 (Appendix A). In most of the questions, the experimental group scored higher than the control group, though not significantly. The experimental group scored lower than the control group on questions 9, 11, 12, 14 and 16. However, the difference in the scores for these questions was no greater than .3.

Both groups displayed a more positive attitude towards testing on questions 4, 7 and 10 as their scores

on these questions were closer to 5. At the same time, the most negative attitudes towards tests were reflected in both groups in their scores for questions 12, 13, 14 and 16. The aforementioned scores seemed to indicate that at the beginning of the instruction period the experimental group showed a significantly better attitude towards tests regarding how tests help learners visualize the content they are studying at the moment and develop test taking strategies such as time management or organization skills. Additionally, both groups expressed a more positive attitude towards tests in connection with their relevance in focusing on what is important, identifying their own weaknesses and knowing how to behave during a test. On the contrary, both groups revealed a more negative attitude towards tests regarding the fact that tests do not seem to increase their self-confidence, their own attitude towards subjects that need many tests in order to pass them and enjoying the experience of test taking.

Table 2. Average Answers per Question in the Pre-Test

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Control Group	3.1	3.5	3.2	3.9	3.7	3.4	3.9	3.2	3.5	4.6	3.1	2.9	2.5	2.5	3.6	2.1	3.1
Experimental Group	3.4	3.8	3.5	3.9	4.1	3.6	4	3.6	3.3	4.7	3	2.6	2.5	2.3	3.7	2	3.3

Regarding the second research question, “*How do learners score on the Attitude towards Test Scale at the end of an instruction period using washback techniques after summative assessment?*” The results were as follows:

The table below indicates the results of both groups on the Post-test. From this data, one might conclude that the experimental group obtained a significantly ($p = .0239$) higher average score over the control group at the end of the semester (58.7 against 52.2) as opposed to their scores on the Pre-test that were similar. As previously stated, a Mean closer to 85 would indicate a highly positive attitude towards tests. From the results of the Post-test, it continues to suggest that the experimental group displayed a moderately positive attitude towards tests; however, such an attitude was improved over the beginning of the semester. On the other hand, the Mean obtained by the control group, indicated a decrease of 3.7 points in comparison with their results on the Pre-test. Regarding their Standard Deviation, both groups experienced a slight increase of less than one point. Hence, both groups can still be considered as highly dispersed. The score discrepancy comparing the Pre and Post tests was discussed with learners from the experimental group attending the follow-up group interview the end of the instructional period. As a result, students attributed a natural decrease in their attitude towards tests and schooling in general during their senior year due to its characteristics, which will be elaborated below. Such remarks may indicate that the decrease of positive attitude in the control group was not softened as they did not experience any washback technique; meanwhile, the experimental group experienced an increment, which could likely be attributed to the intervention.

Table 3. Total Average and Standard Deviation in the Post-Test

	Mean	Standard Deviation
Control Group	52.2	8.1
Experimental Group	58.7	11.5

Additionally, by analyzing the results on a question by question basis, additional differences can be identified. As opposed to the results in the Pre-test, the experimental group scored higher on all the questions except number 10. However, their score on this question remains closely aligned with the ideal score. The most significant differences were evidenced in questions 13 (.9), 12 (.8), 14 (.7), 1, 16 and 17 (.6). Considering the nature of the aforementioned questions, it can be stated that the experimental group displayed a significantly better attitude towards tests regarding feelings associated with subjects that place a great emphasis on tests (which is the case of English), self-confidence attained through test taking, feelings during test taking, and motivation enhanced by test taking. During the follow-up group interview, learners indicated that the implemented techniques were in fact significant in their increase in the mentioned areas. Washback techniques involved in this study were said to have strengthened self awareness of mistakes and their correction. Their self-confidence was also improved upon seeing that their answers were somehow closer to the ones given by classmates they considered to be good learners.

Table 4. Average Answers per Question in the Post-Test

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Control Group	3.1	3.3	3.3	3.4	3.7	3.2	3.7	3.4	3.1	4.8	2.7	2.4	2.1	2.1	3.3	1.7	3
Experimental Group	3.7	3.6	3.6	3.6	4.1	3.6	3.9	3.6	3.5	4.6	3.2	3.2	3	2.8	3.3	2.3	3.6

In connection with the third research question, “*What is the relationship between the use of washback techniques and the learners’ attitudes toward summative assessment?*” Tables 5 and 6 show results of the experimental group at the beginning and at the end of the instruction period in which washback techniques were implemented.

The mean obtained by learners belonging to the experimental group showed an increase of 1.3 after one semester. This increase proved to be statistically significant ($p = .0156$). These results receive additional relevance if we consider the fact that scores decreased for the control group. Taking this data into account, the results seem to suggest that not only did the methodology implemented cause a significant improvement in the learners’ attitude towards assessment, but it may have also softened its decrease. A natural drop in the learners’ attitude towards tests during the instructional period may be explained by the nature of learners’ final year at school (plethora of tests within a short period of time, frequent mock exams of the Chilean PSU test or decrease in motivation due to exhaustion, for instance).

This perspective was voiced by learners during the follow-up group interview.

Table 5. Total Average and Standard Deviation for the Experimental Group in the Pre and Post-Test

	Mean	Standard Deviation
Pre-Test	57.4	10.8
Post-Test	58.7	11.5

Regarding the experimental group's results on a question by question basis (Table 6), most answers fluctuated though not in a significant way. The most significant increase was connected with questions 12 (.6), 13 and 14 (.5). Taking into account the nature of the questions, this rise could be associated with an improvement in learners' attitude towards assessment in terms of feelings related to subjects that require several tests to pass, self-confidence attained through test taking and feelings associated with test taking. During the follow-up group interview, learners indicated that they considered tests as learning experiences thanks in part to the techniques implemented, and that more tests meant more opportunities to prove their knowledge and increase their grades. A small amount of tests was said to create more expectations and anxiety. Feelings associated with test taking were said to have improved due to the fact that seeing their own mistakes from a new perspective allowed them to correct those mistakes and not repeat them on future tests.

Table 6. Average Answers per Question for the Experimental Group in the Pre and Post-Test

Question	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Pre-Test	3.4	3.8	3.5	3.9	4.1	3.6	4	3.6	3.3	4.7	3	2.6	2.5	2.3	3.7	2	3.3
Post-Test	3.7	3.6	3.6	3.6	4.1	3.6	3.9	3.6	3.5	4.6	3.2	3.2	3	2.8	3.3	2.3	3.6

5. Limitations

The accuracy of these results may have been affected by time constraints. This action research was designed to be carried out in one semester for logistical reasons. As was pointed out by Watanabe (2008), research on washback needs to be carried out in a longitudinal study. That being said, one may assume that carrying this action research out during a longer period of time could have presented more accurate results or an even more significant improvement.

Another limitation of this action research was that the interventions were carried out only in one of the 13 subjects that the experimental group had. The questions on the survey asked about their attitude towards tests in general, so their attitude may have improved for the tests of English as the interventions were done with these assessment procedures, but they answered the Post-test with all tests in mind.

Finally, as evidenced by students participating in the follow-up group interview, subjects' answers on the Pre and Post-test may have also been influenced by the grade they were in at the moment of this research. The aforementioned context may have also negatively influenced students' attitudes towards tests.

6. Conclusions and Implications

This study seems to suggest that implementing washback techniques after learners have their graded tests returned improves their attitude towards assessment significantly. The impact of this study may have had an even greater impact had it been carried out for a full academic year.

Additionally, it is not possible to generalize the results obtained through this action research as the sample population was small. Replication of this study with a larger population may lead to more generalizable results.

Another factor influencing the results may have been the balance between the washback techniques used in every session. Analysis of the Class Observation Tables used in the aforementioned sessions indicates a large gap between the frequencies of use of the different washback techniques specified in the literature review: "Praising correct answers" was by far the most frequently used technique as it took place multiple times on every test item reviewed. "Comments on test performance" and "Suggestion of strategies for success" were also used frequently, although they did not take place in every item revision. This was supported by the remarks of students participating in the follow-up group interview as they immediately identified the aforementioned techniques and highlighted how useful they thought they were. "Criticizing weaknesses constructively" was seldom used during the implementation of the sessions. This was evidenced by students attending the follow-up group interview as they seemed unable to identify the aforementioned strategy. Finally, "Learners asking questions about teacher's feedback" hardly ever took place. This last technique was also recognized by learners participating in the follow-up group interview; however, they stated that although given the opportunity they thought it unnecessary to share their queries most of the time as they were able to identify their own mistakes. The results of the study may lead to conclude that a more balanced implementation of the aforementioned techniques may have produced a more significant increase in learners' attitudes towards assessment.

In conclusion, statistical data and learners' opinions raised by this action research seem to suggest that the implementation of washback techniques after tests improves students' attitudes towards evaluation. Therefore it is recommended for EFL teachers to implement the aforementioned techniques in order to reduce test anxiety on learners and thereby transform evaluation into a more positive learning experience for both students as well as teachers.

References

- Andrews, S. (2008). Washback and curriculum innovation. In L. Cheng., Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 37-50). Mahwah, NJ: Lawrence Erlbaum Associates.
- Arezou, A., Rusnani, A. K., Habibah, E., & Maznah, B. (2012). Test anxiety and its related concepts: A brief review. *GESJ: Education Science and Psychology*, 3(22), 3-8.
- Barnes, M. (2016). The Washback of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education*, 41(7), 158-174. <https://doi.org/10.14221/ajte.2016v41n7.10>
- Brindley, G. (2006). Assessment. In R. Carter, & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other languages*. Cambridge: Cambridge University Press.
- Brown, H. D. (2000). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). White Plains, NY: Pearson Education.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education.
- Brown, J. D. (2002). Extraneous variables and the washback effect. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 6(2), 10-13.
- Cheng, L., & Curtis, A. (2008). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3-17). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cheng, L., Watanabe, Y., & Curtis, A. (2008). *Washback in language testing: Research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cohen, A. (2001). Second Language Assessment. In M. Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language*. Cambridge: Cambridge University Press.
- Dodeen, H. (2008). Assessing test-taking strategies of university students: Developing a scale and estimating its psychometrics indices. *Assessment & Evaluation in Higher Education*, 33(4), 409-419. <https://doi.org/10.1080/02602930701562874>
- Dodeen, H., Abdelfattah, F., & Alshumrani, S. (2014). Test-taking skills of secondary students: The relationship with motivation, attitudes, anxiety and attitudes towards tests. *South African Journal of Education*, 34(2), 1-18. <https://doi.org/10.15700/201412071153>
- Gates, S. (1995). Exploiting washback from standardized tests. In J. D. Brown, & S. O. Yamashita (Eds.), *Language testing in Japan* (pp. 101-106). Tokyo: Japan Association for Language Teaching.
- Horwitz, E. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, 21, 127-149. <https://doi.org/10.1017/S0267190501000071>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70, 125-132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.

- Pan, Y. (2009). A review of washback and its pedagogical implications. *VNU Journal of Science, Foreign Languages*, 25, 257-263.
- Prodromou, L. (1995). The washback effect: From testing to teaching. *ELT Journal*, 49(1), 13-25. <https://doi.org/10.1093/elt/49.1.13>
- Salehi, M., & Marefat, F. (2014). The effects of foreign language anxiety and test anxiety on foreign language test performance. *Theory and Practice in Language Studies*, 4(5), 931-940. <https://doi.org/10.4304/tpls.4.5.931-940>
- Tanner, R., & Green, C. (1998). *Tasks for teacher education: A reflective approach. Trainer's book*. Essex: Longman.
- Thornbury, S. (2006). *An A-Z of ELT: A dictionary of terms and concepts*. Oxford: MacMillan Publishers Limited.
- von Würde, R. (2003). Students' perspectives on foreign language anxiety. *Inquiry*, 8(1), 21-40.
- Watanabe, Y. (2008). Methodology in washback studies. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 19-36). Mahwah, NJ: Lawrence Erlbaum Associates.
- Zhang, S. (2016). Washback Effects of CET4 on College English Learning. In *Proceedings of The Fifth Northeast Asia International Symposium on Language, Literature and Translation* (pp. 779-784).

Appendix A

Attitude toward Tests Scale (ATS). English version

Attitude toward Tests Scale—by Hamzeh Dodeen (Translated from Arabic to English by Hamzeh Dodeen)

No.	Item	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	Tests motivate me to study hard					
2	Tests are important to understand materials					
3	Tests help me organize studying time					
4	Tests help me focus on important materials					
5	Tests help me make a big picture of what I am studying					
6	Tests are useful to determine my achievement level					
7	Tests help me identify my academic weaknesses					
8	During tests, I learn many useful skills (e.g., organization, time management, working fast ...)					
9	For me, taking tests is a painful experience					
10	I know how to behave appropriately in tests					

11	I believe that tests should not be used in teaching					
12	Taking tests increases my self-confidence					
13	I try to avoid courses that have many tests					
14	I enjoy taking tests					
15	Tests give me opportunity to present my skills and knowledge I learned during the course					
16	I prefer courses that require many tests					
17	Overall, I think taking tests is a useful experience for me					

Appendix B

Attitude toward Tests Scale (ATS). Spanish version

Attitude toward Tests Scale—by Hamzeh Dodeen (Translated from English to Spanish and adapted by the researcher)

Instrucciones: Esta encuesta mide las actitudes de los estudiantes en relación a las evaluaciones sumativas (pruebas y controles acumulativos). No existen respuestas correctas o incorrectas, por lo que es importante que respondas honestamente.

Lee atentamente cada afirmación y luego rellena con una X el recuadro que mejor representa lo que piensas.

No.	Afirmación	Muy en desacuerdo	En desacuerdo	Neutral	De acuerdo	Muy de acuerdo
1	Las pruebas me motivan a estudiar.					
2	Las pruebas son importantes para entender los materiales y contenidos vistos en clases.					
3	Las pruebas me ayudan a organizar mi tiempo destinado al estudio.					
4	Las pruebas me ayudan a enfocarme en los materiales y contenidos que son importantes.					
5	Las pruebas me ayudan a tener una visión general de lo que estoy estudiando.					
6	Las pruebas son útiles para determinar mi nivel de logros.					
7	Las pruebas me ayudan a determinar mis debilidades académicas.					
8	Durante las pruebas, aprendo muchas habilidades de utilidad (por ejemplo, organización, manejo del					

	tiempo, a trabajar rápido, etc.),.					
9	Para mí rendir pruebas es una experiencia negativa.					
10	¿Cómo debo comportarme durante el transcurso de una prueba.					
11	Creo que las pruebas no debieran utilizarse en la educación.					
12	Rendir pruebas mejora mi confianza en mí mismo.					
13	Si tuviese la oportunidad, evitaría las asignaturas en las que se rinden muchas pruebas.					
14	Me gusta rendir pruebas.					
15	Las pruebas me dan la oportunidad de demostrar las habilidades y conocimientos adquiridos durante las clases.					
16	Prefiero las asignaturas que requieren muchas pruebas.					
17	En general, creo que rendir pruebas es una experiencia de mucha utilidad para mí					

Appendix C

Classroom observation table

CLASS OBSERVATION TABLE

USE OF WASHBACK TECHNIQUES

Class _____ No. of learners _____ Age of learners _____

Length of Lesson _____ Level _____ Date _____

Observer _____

Teacher Observed _____

Materials used _____

Instructions

As you observe the section of the lesson agreed with the observed teacher, put a tick (✓) in the appropriate boxes in the left-hand column every time the teacher uses one of the mentioned Washback Techniques.

If time permits, write down the words the teacher used or a brief description of the Washback Technique implemented.

Item I

Skill _____ Starting Time _____ Finishing Time _____

Washback Technique	Words/Description
Comments on test performance	
Praising correct answers	
Suggestion of strategies for success	
Criticizing weaknesses constructively	
Learners asking questions about Teacher's feedback	

Item II

Skill _____ Starting Time _____ Finishing Time _____

Washback Technique	Words/Description
Comments on test performance	
Praising correct answers	
Suggestion of strategies for success	
Criticizing weaknesses constructively	
Learners asking questions about Teacher's feedback	