

Original Paper

Demystifying the Complexity of Washback Effect on Learners in the IELTS Academic Writing Test

Masoomah Estaji^{1*}¹ Allameh Tabataba'i University, Iran

* Masoomah Estaji, E-mail: mestaji74@gmail.com

Abstract

Test-takers' perceptions of the tests can often provide abundant information for the test designers. This study sought to explore the perspectives of learners involved in the IELTS Academic Writing Module (AWM) courses. In particular, the study explored learner perspectives of motivation, test-taking anxiety, test-taking strategies and the expectations students bring to their courses. This study adopted a mixed methodology and collected data through questionnaires, observations, and interviews. The results didn't indicate a statistically significant difference for the within-subjects variable of learners' perspectives, meaning that the learner perspectives' mean and change from one time to another was not noticeably significant. Further, it indicated a complex relationship between exams and learners' perspectives. However, Analysis of Covariance revealed significant effects for IELTS Writing Preparation course and the learners' improvements in their Writing scores. These findings led to the proposal of a new model for washback which took into account learners' perspectives and provided what should prove to be a useful framework for future studies in the emerging field of washback. The picture that emerged from this study is that the IELTS tests and preparation courses, powerful as they are, cannot be efficient agents for profound changes in learners' perspectives.

Keywords

washback, IELTS AWM, writing assessment, test preparation, learner washback, learner perspective

1. Introduction

Tests are frequently used to spur educational change and innovation, in addition to establishing accountability. This practice results in a phenomenon referred to as “washback,” a term used to describe the impact (negative or positive) that testing has on various aspects of a teaching program, including teachers, classrooms, and students. It is important to note that tests influence teaching and learning. Where a test is used for selection, as is IELTS, those who seek access will attempt to gain the skills they believe necessary to succeed on the test. Some of these skills are generally considered to be desirable, as they are required in the target language use domain. However, as all tests are limited in

how much of the domain they can sample and involve a certain amount of measurement error, there is inevitably scope for the misrepresentation of test-takers' abilities (Green, 2007). The skills required to pass a test are not necessarily or comprehensively the skills required in a target language use domain (Bachman & Palmer, 1996). Washback is thus grounded in the relationship between preparation for success on a test and preparation for success beyond the test, in the domain to which the test is intended to generalize and to which it may control access.

Language learners are the key participants whose lives are most directly influenced by language testing washback. However, there is relatively little research that documents their point of view or their washback-related behavior before and after tests. Some researchers (see, e.g., Cohen, 1984) have reported on what students say about actually taking tests, but more information is needed about learner washback. Furthermore, the majority of the studies on washback have focused on TOEFL contexts. Despite washback researches in language programs, washback in IELTS writing preparation classrooms in Iran has not been significantly researched, nor has it been researched with focus on the learner perspectives. Therefore, a major void in our understanding of learner washback in language pedagogy exists.

2. Review of the Related Literature

It is now widely recognized that tests (and particularly tests that are associated with important decisions, such as university admissions) have a major impact on educational systems and on the societies in which they operate. Alderson and Wall (1993) suggested that the term "washback" provides a useful metaphor to help us explore the role of language tests in teaching and learning, i.e. in relation to factors such as the individual learner, the teacher's attitudes and behavior, the classroom environment, the choice and use of teaching/learning materials. "Washback" allows for the possibility of effects of tests on learning and teaching to be viewed on a continuum – stretching from negative (harmful) at one end, through neutral, and into positive (beneficial) at the other end.

Education relies heavily on testing to make predictions about learner achievement, skill level, and future success. Writing is also a key skill for international students at university as it is most often the basis for assessing their work and so plays a key role in academic success. This study explores the influence of perspectives of learners preparing for a test of academic writing, the IELTS Academic Writing Module. IELTS is a high-stakes gate-keeping test used by universities to screen applicants for language ability. Performance on the test may have serious implications for the life chances of test takers. Hence, IELTS might be expected to exert a strong influence on learner behavior.

Research in language testing has centered on questions about whether or not and how we assess the specific characteristics of a given group of test takers and whether and how we should incorporate these characteristics into the design of language tests. Perhaps the single most important theoretical development in language testing since the 1980s was the realization that a language test score

represents a complexity of multiple influences. Language test scores cannot be interpreted simplistically as an indicator of the particular language ability to be measured. They are also affected by the characteristics and content of the test tasks, the characteristics of the test taker, the strategies the test taker employs in attempting to complete the test task, and the inferences we wish to draw from them. What makes the interpretation of test scores particularly difficult is that these factors undoubtedly interact with each other.

As mentioned above, washback may affect learners' actions and/or their perceptions, and such perceptions may have wide-ranging consequences. Sturman (2003) used a combination of qualitative and quantitative data to investigate students' reactions to registration and placement procedures at two English-language schools in Japan. The placement procedures included a written test and an interview. He found that the students' perceptions of the accuracy of the placement process (i.e., the face validity of the results) were statistically associated with their later satisfaction with the school, the teachers, and the lessons (1996, p. 347).

As Hughes (1993) has pointed out, the key question about the products of washback is whether or not it leads to learning (i.e., language learning). Alderson and Hamp-Lyons (1996), in a study of TOEFL preparation courses in the United States, interviewed students in groups of 3 to 12 people at three different institutions. The language learners were asked for their ideas about how they would like TOEFL preparation classes to be conducted, compared to what they had already experienced. In the preliminary findings reported by Alderson and Hamp-Lyons (1996), the students suggested "having a placement test before a TOEFL preparation course, more opportunities for student participation and student questioning; diagnosis of individual student weaknesses, and the combination of self-study with revision in class" (p. 285).

Alderson and Hamp-Lyons acknowledged, however, that their study would not be able to answer questions about the actual "effects of TOEFL on learners and learning" (p. 284). In fact, only one of the language testing washback studies has documented any demonstrable gains in student learning that can be tied to the use of a test. Hughes (1988) was able to show that students' performance on the Michigan Test (a different, widely recognized measure of English proficiency) increased following the introduction of a new exam and subsequent changes in the English program at a Turkish university.

On the other hand, empirical studies of IELTS washback on learners and the learning process are scarce. The preliminary efforts made by some researchers in this area (e.g. Archibald, 2001; Brown, 1998; Deakin, 1996; & Geranpayeh, 1994) have confirmed some of Alderson and Wall's (1993) Washback Hypotheses on learning (for example, a test influences students' learning content and strategies and the influence varies from student to student) and some of Bailey's (1996) assumptions on the learning processes students would take up. However, these empirical studies have some weaknesses in methodology. Moreover, there is little in the research literature about what actually happens in test preparation classrooms. This gap is particularly troubling in relation to large international and

high-stakes tests such as IELTS, where there are many commercial providers of test preparation courses, especially in Iran. Therefore, a picture of IELTS washback on learners and the learning process is still incomplete and vague.

This study addresses that void by researching washback to the learner in the testing environment of IELTS AWM preparation course within a large IELTS program at a language institute. The research questions provided the framework necessary to uncover the learner perspectives that relate to washback through a mix of both qualitative and quantitative methods. These questions were as follows:

- Is there any washback effect of IELTS writing tests on learners' perspectives of motivation, anxiety, test-taking strategies, and expectations of IELTS writing courses both in IELTS and non-IELTS courses?
- What is the nature and scope of the washback effect on learners' perspectives?

3. Method

3.1 Participants and Research Settings

The participants in this study were selected from IELTS Academic Writing preparation and Advanced Writing classes held at a language institute in Tehran. To control for differences attributable to nationality and first language, all participants in this study were Iranians and L1 speakers of Persian. The participants of the main study (n=79) were both male and female (mostly aged 19-35). IELTS preparation participants were the youngest with an average age of 22 years, while advanced course participants averaged 25. Overall, 41 of participants were female, 33 male, and 5 participants didn't respond to the question. There were 42 learners in non-IELTS courses (i.e. Advanced Writing classes) and 37 studying in IELTS preparation courses. IELTS preparation classes included fewer students on average with a lower proportion of learners; hence, the non-IELTS group was substantially larger than the IELTS preparation group. Most IELTS participants were studying with the aim of entry to local or international universities at Bachelor or Master levels.

The IELTS preparation courses included in this study ranged from 8 to 10 weeks in length and the non-IELTS courses ranged from 5 to 10 weeks. Although the courses varied in length, we didn't expect this variable to predict differences in responses to the questionnaire items and so is not included in the analyses that follow. The non-IELTS courses were all advanced and post-advanced writing courses. Students on these courses were assessed through a combination of teacher assessments and locally developed tests. These features of the context should be kept in mind in interpreting the results.

3.2 Instrumentation

The instruments used in the study consisted of:

- A TOEFL test (for determining the homogeneity of participants)
- Two linked forms of the IELTS Academic Writing Module (AWM). All IELTS tasks were scored by two independent raters (including the researcher) using the official IELTS Writing

Assessment Guide (IELTS, 2000). The scoring scale was initially “pilot-tested” (Weigle, 2002, p. 89) with the test scripts of 58 EFL students at KEI (Kish English Institute) and Tehran Institute of Technology. There were satisfactory inter-rater and intra-rater agreement and reliability (inter-rater) $r = .80$ and (intra-rater) $r = .93$ (using Pearson product-moment correlation coefficient).

- Four student questionnaires: (a) The learners’ motivation toward the IELTS writing tests: *Student Questionnaire A*, (b) The learners’ anxiety toward the IELTS writing tests: *Student Questionnaire B*, (c) The learners’ use of test-taking strategies in the IELTS writing tests: *Student Questionnaire C*, and (d) The expectations students bring to their courses: *Student Questionnaire D*). Items comprised a sentence accompanied by a five-point Likert scale attached to descriptors ranging from *I definitely disagree* to *I definitely agree*. Internal consistency reliability of the questionnaires’ items was measured by the Cronbach Alpha coefficient. The reliability of the items in Questionnaire A was estimated $\alpha = .57$. However, the reliability estimates of the other questionnaires, B, C, and D, enjoyed higher degrees of internal consistency or reliability (Questionnaire B, $\alpha = .88$; Questionnaire C, $\alpha = .74$; and Questionnaire D, $\alpha = .83$).

- Semi-structured interviews

3.3 Data Collection Procedures

In order to carry out this study, the following steps were taken.

Stage I. First, the homogeneity of the participants across the groups was determined in terms of their L2 proficiency by administering an official version of the TOEFL to 140 participants. The homogeneity of the participants was proved based on the scores of the testees in their TOEFL. Based on their scores, those participants who obtained scores within the range of one standard deviation above and below the mean participated in this study. Out of 90 subjects who had obtained scores within that range, 79 subjects were selected and were randomly assigned to Experimental and Control Groups. They comprised 42 learners in non-IELTS courses (i.e. Advanced Writing classes) and 37 studying in IELTS preparation courses.

Stage II. The first interview was conducted prior to the IELTS preparation course. This interview focused on the learners’ plan for the course, and what their perceptions were of the differences between IELTS and non-IELTS courses that they had previously attended. In this stage, learners took an Academic Writing Module of IELTS as their pretest. Furthermore, the learners were asked to complete 4 questionnaires to collect information about their motivation for the study, test anxiety, test-taking strategies, and expectations of the IELTS preparation course. They were also given these questionnaires at the end of the course to record any changes in their perspectives or knowledge of the test.

Stage III. In this stage, treatment was conducted. The treatment conditions of the study were operationalized for the IELTS Academic Writing preparation class.

IELTS Academic Writing Preparation: A syllabus was designed for use in IELTS AWM course as their treatment sessions. IELTS test practice materials had focused on the requirements of the IELTS Academic Writing tests and had been targeted at problem areas with hints for improvement. Moreover, in each session after presenting the IELTS test practice materials (while focusing on IELTS Academic Writing Test Tasks and Strategies), the Experimental Group (i.e. IELTS learners) was assigned to take an IELTS sample test. Six IELTS sample tests were taken during this preparation course. These sample tests were used to help students practice under test conditions and develop their understanding of IELTS Academic Writing Test.

Stage IV. In this stage, after treatment was conducted, both the first and the second test forms of the IELTS AWM were administered as their post-test to statistically equate the test forms for their level of difficulty. All IELTS tasks were scored by two independent raters (including the researcher) using the official IELTS Writing Assessment Guide (IELTS, 2000). The raters employed for the study were all IELTS instructors and experts in rating IELTS scripts and writing instruction. To preclude any bias resulting from expectations of gain following instruction, the rater was given no indication of whether any given script had been written at course entry or at the course exit. There were satisfactory inter-rater and intra-rater agreement and reliability (inter-rater) $r = .80$ and (intra-rater) $r = .93$ respectively (using Pearson product-moment correlation coefficient).

Furthermore, the students were given the questionnaires at the end of the course to record any changes in their perspectives or knowledge of the test. A final interview was held after students had finished their IELTS preparation course. This final interview elicited responses about changes that had occurred in the learner's perspectives as a result of their increased knowledge of the test and the testing procedures for IELTS AWM.

4. Results and Discussion

A Multivariate Test for Repeated Measures was run to compare group means (i.e., the learners' perspectives of both Experimental and Control Groups) of this study at two various points in time (i.e., pretest vs. posttest) with an interval of eight weeks in between. The results of the Multivariate Test for Repeated Measures for the within-subjects effects displayed in Table 1 don't indicate a statistically significant difference for the within-subjects variable of learners' perspectives, meaning that the learners' perspectives' mean and change from one time (pretest) to another (posttest) was not noticeably significant ($F=1.319, p>.05$). More importantly, a statistically significant effect was not found for the interaction of learners' perspectives and writing gain ($F=0.104, p>.05$), showing that the statistically significant development, change, or achievement in learners' perspectives did not occur with respect to the levels of their writing gain scores. Furthermore, a statistically significant effect was not found for the interaction of the learners' perspectives and the groups ($F=0.357, p>.05$), meaning that no significant difference of change or improvement in the learners' perspectives occurred between

the groups. That is, the statistically significant development, change or achievement in learners' perspectives did not occur in any of the groups influenced by IELTS Academic Writing preparation course employed in their treatment settings or academic writing course.

Table 1. The results of multivariate test for repeated measures

Effect	Multivariate Tests for Repeated Measures					
	Test	Value	F	Effect df	Error df	<i>p</i>
R1	Wilks	0.949	1.319	3	74	0.274
	Pillai's	0.050	1.319	3	74	0.274
	Hotelling	0.053	1.319	3	74	0.274
R1*Writing-Gain	Wilks	0.995	0.104	3	74	0.957
	Pillai's	0.004	0.104	3	74	0.957
	Hotelling	0.004	0.104	3	74	0.957
R1*Group	Wilks	0.985	0.357	3	74	0.784
	Pillai's	0.014	0.357	3	74	0.784
	Hotelling	0.014	0.357	3	74	0.784

Although a statistically significant effect was not found for interactions, Univariate Tests of Significance for gain were run to see whether this analysis might reveal a significant effect or not. The Univariate Tests of Significance for gain showed no significant ($p>.05$) effects for either Motivation-Gain or Anxiety-Gain. Similarly, Univariate Tests of Significance for gains with Test-taking strategies or Expectations indicated no significant ($p>.05$) effects. However, the results of the Univariate Tests of Significance for writing gain shown by Table 2 indicate the occurrence of a statistically significant difference between the mean performances that is the writing gain ($F=8.255$, $p<.05$) of the experimental and control groups in this study. Table 2 displays the results of the Univariate Tests of Significance for writing gain.

Table 2. The results of univariate tests of significance for gain

Effect	Univariate Tests of Significance for Gain				
	SS	Degr. of Freedom	MS	F	<i>p</i>
Motivation-Gain	0.050	1	0.050	0.077	0.780
Anxiety-Gain	0.040	1	0.040	0.062	0.802
Test Stra-Gain	0.017	1	0.017	0.026	0.870
Expectations-Gain	0.149	1	0.149	0.232	0.631
Group	5.314	1	5.314	8.255	0.005

That IELTS Academic writing preparation course influences on the group's writing performance seems clear, but how this course influences learners' perspectives is much less clear.

As noted, the results of the Table 1 demonstrate that the treatment presented in this study did not have a significant effect on the learners' perspectives of the IELTS Academic writing preparation course. The research findings revealed that the IELTS Academic Writing course influenced learners' writing gain

scores; however, learners' perspectives toward the exam remained largely unchanged. Based on these results, it can be concluded that the first null hypothesis of this study is not rejected.

In sum, the results of the data collected in the form of pretests and posttests did not find a significant effect for learners' perspectives in any of the groups influenced by their preparation courses. Interestingly, the noticeable improvement that the learners achieved in their scores or performance was found to be statistically significant. Overall, the results showed that IELTS tests, powerful as they are, might not be efficient agents for profound changes in learners' perspectives.

Of even more interest were the findings of the qualitative analysis that provided further evidence for learners' perspectives which were disparate and too mixed, probing the complex and manifold mechanisms of the learner washback studies. The major conclusion reached through examination of the interview data appeared to be that change had occurred for some learners, but not for all, and to different degrees. It also appeared that this change might differ with the passage of time. In addition, learners had mixed feelings toward the preparation course and the exam, recognizing on the one hand that this course made them work hard to achieve good scores but at the same time they thought that this course could not satisfy all their needs and that exams were not an accurate reflection of all aspects of their study. More specifically, learners' perspectives to test preparation varied. Some students tended to rely on fulfilling writing tasks rather than motivating themselves to learn.

What helps me in carrying out these tasks is that I learned how to do it. However, it is harder for me because it seems that I want to include everything presented in tables and figures but that seems impossible.

For another learner in academic writing course, constructing a different attitude toward writing rested in part on her reflectively juxtaposing aspects of writing with other subjects she had experienced, such as logic and mathematics. She stated,

Writing is, to some extent, more logical than thinking...I'm getting to like academic writing now. Just a bit... Since I don't dislike logic. I am taking math as well now, which has been doing logically, too. And I realize that the problem that I have is I'm not good at brainstorming.

Various factors were cited by learners becoming motivated and interested in these courses. For instance, one learner said,

I know that getting involved in different tasks and motivation are important factors in learning, perhaps... the most important factors; however, I enjoy doing different writing tasks. It is because in the course of IELTS writing practices, I can learn to write in a variety of styles and organize my ideas carefully. This course helped to build my knowledge of the test with spending much time on test practices.

In contrast, some students seemed to have problems with different approaches to writing, as they said,

I feel not so exciting. Some writing tasks have really become a grinding thing without pleasure. Why do we have to restrict our ideas in topic sentences? It just makes me lose ideas on how to argue in an interesting way. I am used to this way to write directly from heart, from... intuition not skills, which makes me feel glee and sparks more ideas.

The most difficult and frustrating point is when you are given a topic to write an essay about and you do not have a clear image of what you need to write and include; however, we were taught some test-taking strategies, such as "It is very hard to delineate the meaning of, but as I figure it out it...."

For analyzing tables, charts, and figures you need to practice and think fast. Sometimes, you get baffled and it is so hard to keep everything in mind and focus on the most remarkable points.

While the study reiterated the complexity of investigating washback to the learner, it also provided an indication as to the sources of this complexity that can be traced both inside and outside the classroom context. Thus, these findings and the discovery of the patterns in the learner responses led the researcher, finally, to a proposed new model for washback to provide more insights into the nature and scope of this educational phenomenon across different factors and research contexts.

5. A Proposed Dynamic Model of Test Washback

A proposed dynamic washback model is presented in Figure 1 in light of the previous analysis of washback studies, the major washback models, and current leading theories such as Alderson and Wall's (1993) fifteen washback hypotheses, Green's (2003) predictive model of test washback, Saville's (2000) model of test impact, Messick's (1996) perspective of test validity, Bailey's (1996) basic model of washback, Hughes' (1994) trichotomy of washback, and Bachman and Palmer's (1996) micro-and-macro washback model.

This model incorporates ideas from Hughes (1993, cited in Bailey, 1999) in describing a trichotomy of test effects in terms of "participants," "process," and "product." Tests could affect teachers, students, administrators, materials writers, and publishers in terms of their perceptions, activities they engaged in, as well as the amount and quality of learning outcomes. Alderson and Wall (1993) propose fifteen washback hypotheses and illustrate some of their effects, from the most basic to the more specific, that tests might have on teaching and learning. For example, "A test will influence teaching/learning" (p. 120) and "Tests will have washback effects for some learners and some teachers, but not for others" (p. 121). Bailey (1996) combined the fifteen hypotheses from Alderson and Wall (1993) within the trichotomy of the backwash model proposed by Hughes (1993), and created the "basic model of washback." Bailey distinguishes between "washback to the learner" (what and how learners learn and the rate/sequence and degree/depth of learning) and "washback to the program" (what and how teachers teach and the rate/sequence and degree/depth of teaching) to illustrate the mechanism by

which washback works in actual teaching and learning contexts.

A common characteristic of these washback models is that they tend to highlight what washback looks like and who is affected, but do little to address the factors that contribute to the phenomenon. In other words, “process” is less understood than “participants” and “products.” Besides, the products in these three models/hypotheses refer mainly to teaching and learning washback, not to the aspects of washback context which highly influence the learners and teachers in education. Furthermore, these models address three key factors in contemporary educational systems which need to be accounted for:

- the nature of complex dynamic systems
- the roles that stakeholders play within such systems
- the need to see assessment systems as educational innovations within the systems and to manage change (i.e., dynamism) effectively

The proposed model in Figure 1 aims to represent the test washback effects on students and the interrelationship among a lot of factors that influence students’ learning. These interrelations form a circle of the causal links. However, these interrelations indicate non-directionality of the influence from the determining factors to the dependent one in various contexts (in both micro- and macro-contexts). The more it departs from the micro-context, the more variable it becomes. Washback at the micro-context is postulated to consist of teaching, learning, teaching material, and score gain effects, while washback at the macro- context is postulated to consist of innovation and social dimension feature “washback variability” and “washback intensity,” in Green's (2003) term. Green’s model relates design issues to the contexts of test use, including the extent to which participants (including material writers, teachers, learners, and course providers) are aware of and are equipped to address the demands of the test and are willing to embrace beliefs about learning embodied therein. These features are most closely related to washback *variability* (differences between participants in how they are affected by a test) and washback *intensity*. In other words, washback intensity varies in relation to participants’ perceptions of test stakes and test difficulty (Hughes, 1993). In sum, to understand how these aspects in both micro- and macro-contexts evolve, an investigation of how participants themselves react toward tests must be conducted. For example, to better understand why teachers change what they teach but not necessarily their methodology (Cheng, 1999, 2004, 2005) following the introduction of a test, their beliefs, perceptions of the test, and their levels of participation in its implementation may help us understand the phenomenon of washback.

Apart from the test itself, there are many factors within a society, particularly the educational environment with its typical conditions, which all influence the behaviors of learners. The flowchart in figure 1 displays the circle of testing effects on learners.

**NON-DIRECTIONALITY
OF WASHBACK**

**WASHBACK VARIABILITY
WASHBACK INTENSITY**

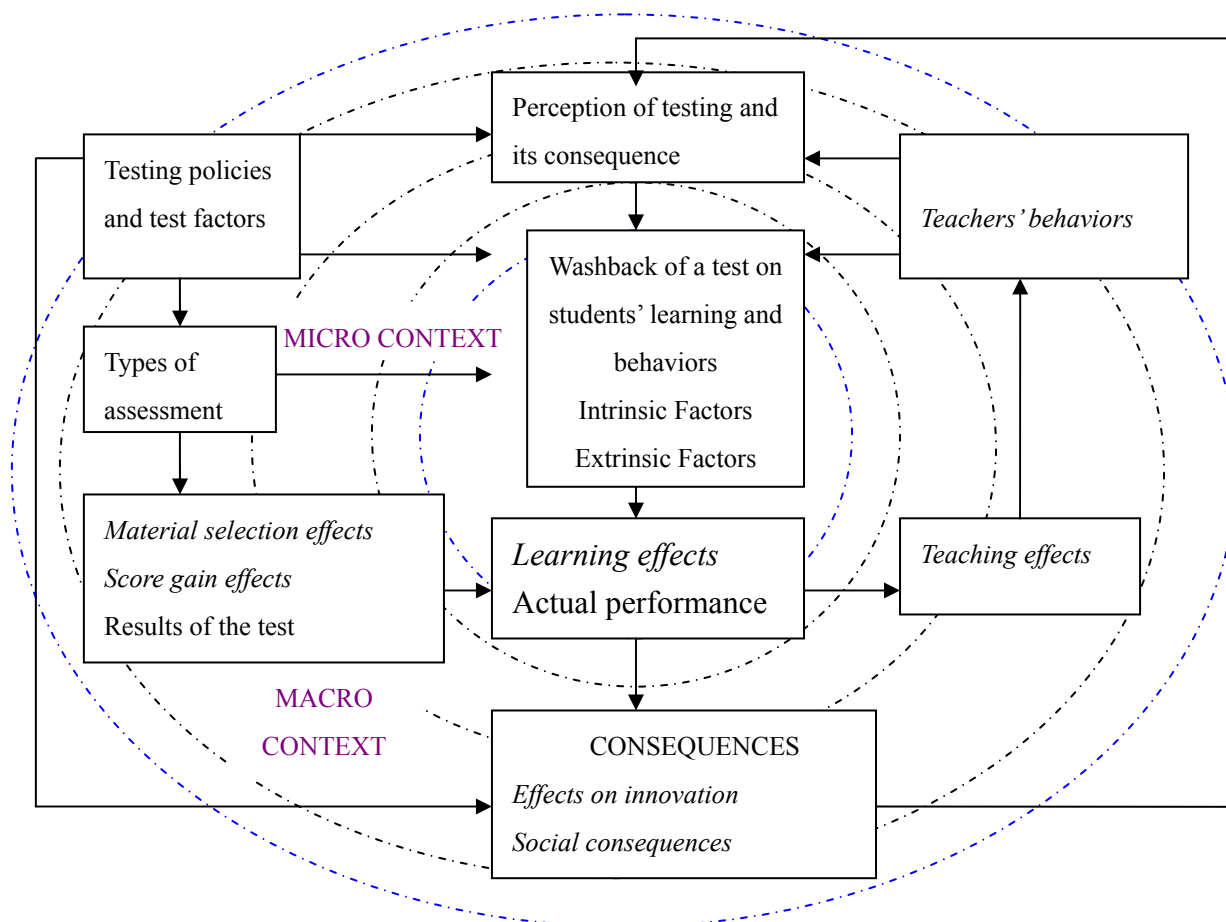


Figure 1. A proposed dynamic model of washback

Based on ideas of Bachman and Palmer (1996), Green (2003), Hughes (1993), and Saville (2006)

Examining the figure from left to right, testing policies and types of assessment play a very important role that together with results of the test influence learners' perspectives of testing. Test types thus enhance changes in learners' behaviors that lead to changes in attitudes and motivation and learners' learning method. Consequently, the outcomes of these changes and interactions are the change in learners' behaviors and perspectives thus lead to change in learners' actual performance then consequences. However, the complexity of the phenomenon lies in the multi-directionality which often entails variance in washback specificity, intensity, intentionality, length, and value.

The model presented in Figure 1 investigates how these factors interact on both macro- and micro-contexts. In addition, this model advocates a well-rounded investigation of washback that focuses not only on a given educational context but also society at large. As Cheng (2001), Alderson and Wall (1993), and Watanabe (1997) suggest, washback is a dynamic and complex phenomenon that

involves a variety of intervening variables such as tests, test-related teaching, learning, and the perspectives of stakeholders. The hugely complex nature of washback is due to the fact that it is an interactive multi-directional process involving a constant interplay of different degrees of complexity among the different washback components. Testing and assessment are located in both a narrower educational and a broader social context, and the involvement of not only direct participants, but also indirect participants including those from the broader social context adds both greater significance to the washback phenomenon and different degrees of complexity. The researcher feels that it is not possible to illustrate the phenomenon completely within a single model such as the figure above. The major question is how these considerations can be combined to produce a comprehensive, integrated model.

6. Conclusion

Data analysis revealed that tests affected learning, and learners could profit, in terms of writing score gains, from giving attention to IELTS preparation tasks and activities, but the additional benefit was surprisingly limited. However, there was little evidence of dramatic increases in scores on the part of the learners as a result of preparation in their Academic Writing courses. In contrast, the data showed that the washback effect of this exam seems to be limited in the sense that it did not appear to have a fundamental effect on learners' perspectives. In other words, learners' motivation for the study, test anxiety, test-taking strategies, and their expectations remained largely unchanged. The behaviors suggesting washback exhibited during this study were disparate and mixed. The results revealed that tests affected learning in IELTS Academic Writing Preparation classes, but they affected different learners in different ways. That is, the effect was not the same in degree or in kind from learner to learner and no significant change was observed regarding aspects of examination influence on the learners' perspectives. In other words, the contribution of test preparation to learners' perspectives appeared to be minimal in this setting. Learners pursuing a test-preparation course did not obtain a significant advantage in their perspectives. However, learners intending to take the test, both the learners of IELTS Academic Writing course and the Academic Writing course, did take a significant advantage in their test performance. It is important to reiterate that the extent to which learner perspectives reflect only test expectations was limited due to the variety of other factors influencing learner perspectives that included but were not limited to, learner motivation for study, test anxiety, test-taking strategies, and their expectations of the IELTS preparation course. There may be individual differences among learners in the way they perceive and react to exams. Therefore, in this context, it seems to be washback to the program, rather than washback to the learner, which has the greater relevance to outcomes.

In sum, this study indicated that rather than being a direct automatic effect, washback is actually complex and elusive; while this study showed that there was washback from the exam onto a variety of

learning areas, it also indicated that washback to the learner was not present and it varied in form and intensity. There are many independent and intervening variables such as teacher factors, the stakes of the test, the design of the test, textbooks, resources, classroom conditions, management of classroom practices, and many other factors which seem to be important variables influencing learners and their learning.

Hence, several discrepant findings from this study further support the argument that washback is quite context-oriented and complex. Simply examining one or some factors or examining the phenomenon in one context is not capable of explaining critical washback issues, such as how and why washback phenomenon influences some learners but not others. Previous washback studies conducted by Cheng (1998), Ferman (2004), Green (2007), Read and Hayes (2003), and Shohamy et al. (1996), have shown that affecting learner perspectives is challenging and complex and requires an attentive focus on various affective, cognitive, and social factors. Focusing on these variables can help provide learners with the critical perspective needed for improvement, as well as the impetus for change in behaviors when needed. In addition, learner perspectives concerning the relationship between teaching, learning, and high-stakes exams needs to be studied longitudinally and directly using research methods that will capture the aspects under investigation more clearly. However, it can also be argued that even more methods could be employed to help researchers probe deeper into the less observable factors related to the individuals involved. Without this type of focus or attention, learners will often continue to learn in the same manner that they are used to learn, and will continue to emphasize what they believe are the important aspects of language learning, whether or not they are based on skills, and/or included on the tests. Future research into washback, by taking learner perspectives into account, will provide more grounded accounts of test washback and its implications for test validity. In addition, to get a more comprehensive picture of test washback, it is desirable to conduct studies which look at washback of a specific test from different perspectives using different research methods (including at least the two central participants involved: teachers and students) in order to investigate the influence it exerts on classroom learning and teaching in depth.

Tests will never be eliminated from educational institutions; therefore, it is best to embrace them and their power. What is needed in such a situation is a rational argument based on empirical evidence indicating the actual power of the examinations, whether negative or positive.

References

- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129.
- Archibald, A. (2001). Managing L2 writing proficiencies: Areas of change in students' writing over time. *International Journal of English Studies*, 1(2), 153-174.

- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: OUP.
- Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- Bailey, K. M. (1999). *Washback in language testing*. TOEFL Monograph Series. Report Number: RM-99-04, TOEFL-MS-15. Princeton, NJ: Educational Testing Service.
- Brown, J. D. H. (1998). An investigation into approaches to IELTS preparation, with particular focus on the Academic Writing component of the test. In S. Wood (Ed.), *IELTS research reports* (Vol. 1, pp. 20-37). Sydney: ELICOS/IELTS.
- Cheng, L. (1998). Impact of a public English examination change on students' perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24(3), 279-301.
- Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teacher and Teacher Education*, 15, 253-271.
- Cheng, L. (2001). Washback studies: Methodological considerations. *Curriculum Forum*, 10(2), 17-32.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 147-170). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge, UK: CUP.
- Cohen, A. D. (1984). On taking tests: What the students report. *Language Testing*, 1(1), 70-81.
- Deakin, G. (1996). IELTS in context: Issues in EAP for overseas students, *EA Journal*, 15(2), 7-28.
- Ferman, I. (2004). The washback of an EFL national oral matriculation test to teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 191-210). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Geranpayeh, A. (1994). Are score comparisons across language proficiency test batteries justified? An IELTS-TOEFL comparability study. *Edinburgh Working Papers in Applied Linguistics*, 5, 50-65.
- Green, A. (2003). Test impact and EAP: A comparative study in backwash between IELTS preparation and university pre-session courses. Unpublished doctoral dissertation. Roehampton, UK: University of Surrey.
- Green, A. B. (2007). *IELTS washback in context: Preparation for academic writing in higher education*, *Studies in Language Testing* 25. Cambridge: Cambridge University Press and Cambridge ESOL.
- Hughes, A. (1988). Introducing a needs-based test of English language proficiency into an English medium University in Turkey. In A. Hughes (Ed.), *Testing English for University Study (ELT Documents #127)* (pp. 134-146). London: Modern English Publications in association with the British Council.

- Hughes, A. (1993). Testing the ability to infer when reading in a second or foreign language. *Journal of English and Foreign Languages*, 10(11), 13-20.
- Hughes, A. (1994). Backwash and TOEFL 2000. Unpublished manuscript, commissioned by Educational Testing Service (ETS). University of Reading.
- IELTS (2000). *Writing assessment guide*. Cambridge: UCLES, The British Council, IELTS Australia.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241-256.
- Read, J., & Hayes, B. (2003). The impact of IELTS on preparation for academic study in New Zealand. *IELTS International English Language Testing System Research Reports*, 4, 153-206.
- Saville, N. (2000). Investigating the impact of international language examinations. *Research Notes*, 2, 4-7.
- Saville, N. (2006). Criterial features of English across the CEFR levels: Evidence from the English profile program. *Research Notes*, 37, 34-40.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298-317.
- Sturman, L. (2003). Teaching to the test: Science or intuition? *Educational Researcher*, 45(3), 261-273.
- Watanabe, Y. (1997). The washback effects of the Japanese University entrance examinations of English-classroom-based research. Unpublished PhD thesis, Department of Linguistics and Modern English Language, Lancaster University, Lancaster, England.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Questionnaire A (Students' Motivation Questionnaire)

In this questionnaire, we would like to find out about your reasons for taking this course, your anxiety for taking a test, your test-taking strategies and what you expect to study on this course.

Your responses to this questionnaire will be treated in confidence, and only used for the stated purpose of the study, but we do need your name to help us to organise the information.

The questions usually take about 20 minutes to answer.

Thank you very much for your time and co-operation.

Section 1

Full Name: _____

Nationality: _____

Age: _____

Gender: _____

University Degree _____

Have you taken any IELTS test before? If yes, where, when, how many times and which module?

Have you attended any IELTS preparation classes in the past? If yes, where and for how long?

In this section, we would like to find out about your reasons for taking this course and for studying in an English speaking country.

Please grade the following on a 5-point scale where

1= I definitely disagree

2= I tend to disagree

3= I don't know/I can not

answer this question

4= I tend to agree

5= I definitely agree

and tick the appropriate column for each item.

		I definitely agree	I tend to agree	I tend to disagree	I definitely disagree	I don't know
RE01	I am taking this course because I want to get a good grade on IELTS (or other test/assessment called: _____).	5	4	2	1	3
RE02	I am taking this course because I want to learn useful skills for studying at university.	5	4	2	1	3
RE03	I am studying on this course because I want to improve my general ability to use English.	5	4	2	1	3
RE04	I am required to take the course by my employer, my parents, or other authority. (_____)	5	4	2	1	3
RE05	This course will help me acquire new ideas and broaden my outlook.	5	4	2	1	3
RE06	I want to do well in this course because it is important to show my ability to my family/friends/supervisors/others.	5	4	2	1	3
RE07	I have a different reason for taking this course (write your reason here): (my reason is: _____).	5	4	2	1	3
RE08	I am going to college/university in an English speaking country to improve my English.	5	4	2	1	3
RE09	I am going to college/university in an English speaking country to help me get a good job in the future.	5	4	2	1	3
RE10	I am going to college/university in an English speaking country to study a subject that interests me.	5	4	2	1	3
RE11	I am required to attend university/ college by my employer, my parents, or other authority. (_____)	5	4	2	1	3
RE12	I have a different reason for going to university/college in an English speaking country: (my reason is: _____).	5	4	2	1	3