



Can Exams Change How and What Teachers Teach? Investigating the Washback Effect of a University English Language Proficiency Test in the Turkish Context

Asli Lidice Gokturk Saglam^{a1} 

^a *Ozyegin University, School of Languages, Istanbul, Turkey*

Received 31 March 2018 | Received in revised form 26 June 2018 | Accepted 24 July 2018

APA Citation:

Gokturk Saglam, A., L. (2018). Can exams change how and what learners teach? Investigating the washback effect of a university English language proficiency test in the Turkish context. *Eurasian Journal of Applied Linguistics*, 4(2), 155-176. doi: 10.32601/ejal.464094

Abstract

This article reports a mixed-method study that examined the washback effect from a locally-produced, theme-based, high-stakes English language proficiency test in tertiary education in a Turkish EAP context. The aim was to explore the extent to which washback on teaching was induced by an integrated theme-based English proficiency test designed to reflect authentic language use in the tertiary education context in Turkey. The data collection involved classroom observations and focus group interviews with 14 instructors from the Preparatory English Language Program (PEP). Classroom observations were conducted using the Communicative Orientation of Language Teaching Observation Scheme (COLT) (Spada & Frohlich, 1995), and data was qualitatively and quantitatively analyzed. Inductive analysis of the transcribed interview data was also used. The findings indicated that both positive and negative test effects were exerted on teaching. In addition to positive washback on materials, this study also found negative washback in the form of narrowing of the curriculum. Findings also implied that although the test had varying amounts and types of washback depending on the particular teacher involved, both content and methodology in teaching are affected. The article concludes by interpreting these results in the light of recent studies on learner washback, discussing implications for teachers, and providing suggestions for further research.

© 2018 EJAL & the Authors. Published by *Eurasian Journal of Applied Linguistics (EJAL)*. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Integrated language proficiency test; theme-based language proficiency test; washback effect; washback on teaching

1. Introduction

Tests impact learning and teaching in ways generally acknowledged to be complex and multifaceted. Targeting a positive washback effect with respect to student learning, educators and test designers often attempt at bringing about educational change in English language teaching and learning educational policies by means of introducing new high-stakes exams or integrating modifications on existing exams. A prevalent conclusion that is drawn from the washback studies has framed the mediating role played by the teacher and the predictability of the effects of a testing innovation. It is

* Corresponding author. Tel.: +90-216-564-9000
E-mail address: aslilidice@gmail.com

often concurred that the crux of using washback to engineer a pedagogic change depended on teacher perception and on the content of teaching (Andrews, Fullilove & Wong, 2002; Wall, 2000).

Turkey is known with its test-oriented education system which exerts strong negative washback on teaching and learning (Akpınar & Cakildere, 2013; Karabulut, 2007; Ozmen, 2011; Sevimli, 2007). Within this test-dominant context utilizing tests as a lever for change to engineer positive washback in education and examining test effects on teaching gain importance. The research context of this study is a proficiency testing situation at the tertiary level in a Turkish university. The Test of Readiness of Academic English (TRACE) aims at assessing whether language learners have the sufficient ability to use English for academic purposes in an English-medium instruction university. In an attempt to depict real-world communicative acts displaying language proficiency the main premise of TRACE is inclusion of integrated tasks of listening-to-writing and reading-to-writing in a proficiency test, since it was assumed that integrated writing assessment tasks can better meet the demands of authentic communication. Academic writing in many academic contexts entails source material being read or listened to (Gebriel, 2009; Plakans, 2009; Weigle, 2004). Thus the construct of integrated assessment mirrors academic literacy activities since in many academic contexts, writing requires the integration of reading and listening (Cumming, 2013). Integration of skills also encourages authenticity since research studies confirmed that academic writing tasks in English for Academic Purposes courses and university content courses commonly resort to use of external sources to guide and assist learners in building arguments and developing content (Leki & Carson, 1994, 1997).

In order to align instruction with the test in the research context, teaching materials are organized around a single theme that entails an integrated skills approach and replicating exam tasks. It is often accentuated that when there is a curricular alignment in a language program between what is taught and what is tested washback is apt to be strong (Madaus, 1988; Smith, 1991) for “what is assessed becomes what is valued, which becomes what is taught” (McEwen, 1995 in Cheng & Curtis, 2004, p. 3). Given the unique role of integrated assessment in engineering positive washback, examining test effects on teachers gain utmost importance since they play a significant role in mediating washback.

2. Literature review

2.1. Definition of washback

The terms ‘backwash’ (used in general education field) and ‘washback’ (more commonly used in applied linguistics, language education, and language testing) have been used to refer to the effect of testing on teaching and learning (Hughes, 1989). Washback has also been formulated as the influence of tests which lead “teachers and learners to do things they would not necessarily otherwise do” (Alderson & Wall, 1993, p.17). Washback is conceptualized as being either negative (harmful) or positive (beneficial). A test may exert negative washback when a test’s content or format is based on a narrow definition of language ability, and consequently cause constraints in the teaching and learning context. Positive washback occurs when a testing procedure brings about good teaching practice (Taylor, 2005).

Hughes's washback trichotomy model (1994) categorizes test effects in three areas: participants, process, and product. First, participants refer to stakeholders such as students, teachers, administrators, materials writers, and publishers, whose perceptions related to the teaching and learning process may be influenced by examinations. Second, process refers to endeavours in teaching and learning such as materials development, syllabus design, modifications in instruction and methodology, and use of learning and/or test taking strategies. Finally, product encompasses learners' intake, skills, and quality of learning (Bailey 1996, p. 262). Tests have a number of effects: they influence participants (the teachers, learners, and materials writers involved in test preparation, and the perceptions and attitudes they bring to the task), and they trigger modification of their processes (teaching and learning behaviours). Consequently, these impact learning outcomes (Green, 2007a, p.78).

2.2. Washback on teaching

A number of researchers (e.g., Cheng, 2004, 2005; Green, 2007b; Wall & Alderson, 1993; Wall, 2005) proposed that while “what” teachers teach (content) is affected by test washback, there are no changes on “how” teachers teach (methodology). Wall and Alderson (1993) studied the impact of a new English examination in Sri Lanka which was, as a curricular innovation towards more learner centered instructional approaches, expected to bring about positive washback effect on language teaching. Researchers concluded that although the test had washback effect on the content of teaching, no evidence was found for any influence of the test on how teachers taught. In a follow up study Wall (2005) discovered that teachers focused on the tested skills involving reading and writing on the New O level English examination in Sri Lanka. As a result, an uneven instruction in skills was particularly eminent during the examination preparation period. In contrast to the student-centered approach that the examination was expected to induce, teacher-centered approach was generally employed. In a similar vein, Cheng (2005) disclosed that the new Hong King Certificate of Education Exam in English encouraged teachers to prioritize speaking and integrated skills aspects of the course. However, despite the inclusion of learner centered communicative activities, Teacher Talking Time (TTT) still remained a substantial part of classroom teaching. Furthermore, after an analysis of teacher perceptions collected through a survey and focus group interviews, Chen (2002) construes that the influence of public examinations on teachers' curricular planning and instruction as ‘superficial’ and indicated that the washback may influence content but not methodology. It is often asserted that washback research should “relate teachers' attitudes to an understanding of exams to observations of classrooms in order to understand why teachers teach the way they do, and why tests might not have the impact that is frequently asserted” (Wall & Alderson, 1993, p.41). Therefore, the main objective of the current study is to explore teachers' perceptions towards the test and explore how test effect is manifested in instruction.

On the other hand, a number of researchers (e.g. Alderson & Hamp-Lyons, 1996; Shohamy, 1993; Watanabe, 1996) confirmed tests affect not only ‘what’ but also ‘how’ teachers teach. Stecher, Chun and Barron (2004) examined the effects of assessment-driven reform endeavours on the teaching of writing in 70 Washington elementary and middle schools, surveying 277 teachers and analysing test scores as well as classroom practices. The researchers found that education reform to be brought through exams and

engineering positive washback in Washington affected not only the content, but the teaching methodology.

Some studies also noted that tests exerted effects and changed the methodology of the teachers in varying degrees. Lam's (1994, p.91) findings revealed that more experienced teachers were significantly more "examination-oriented" in comparison to the less experienced. Similarly, research by Shohamy (1993) and Shohamy, Donisa-Schmidt, and Ferman (1996) revealed differences between experienced and novice teachers, arguing that while experienced teachers taught towards the test, prioritising materials to be included on the test; novice teachers focused on a wider repertoire of oral language activities. These findings support prior research which reported that washback generates a narrowing of curriculum, pointing out an overall negative washback effect of the English language tests on materials and narrowing of the curriculum to testable skills (e.g. Alderson & Wall, 1993; Li, 1990; Read & Hayes, 2004).

2.3. Teacher's role in washback

While a number of research studies have reported contradictory findings regarding washback effect upon what (content) and how (methodology), researchers concluded that most teachers had the inclination to 'teach to the test' to increase success rates in test, although individual teachers may be affected by the tests to different degrees, and it seems that the washback may variously affect content, methodology, or both. (Cheng, 2004; Ferman, 2004; Gu, 2007; Shohamy et al., 1996; Wall, 2005).

Although scholars could not reach an agreement as to the areas of teaching that washback affects, it is commonly acknowledged that teachers play an essential role in determining the extent to which washback operates, the areas it should operate in, and how it operates (Spratt, 2005). Chen (2002) elucidated a variety of factors which affect teachers' perceptions of the impact of public exams on their teaching. These factors are classified into teacher and context characteristics. The former involves teaching experience, education, in-service training education, perceived professionalism in teaching, perceived importance of the exam, gender and perceived awareness of the exam. Context characteristics include school type, school location, the age of students, students' perceived learning attitudes, perceived attention from external forces, and class size). Similarly, some studies have indicated that the washback effect of new tests, which were intended to be a lever for educational change, exerted varying effects on teaching based on the attitude of individual teachers (Alderson & Hamp-Lyons, 1996; Burrows, 2004; Watanabe, 1996).

2.4. Research methods in washback studies

Due to variety of mediating factors and the intricate nature of the phenomenon, researchers stress the need for a range of data collection methods to explore washback in context (e.g. Burrows, 2004; Cheng, 2001; Cheng, 2004; Watanabe 2004). For instance, Scott (2007) suggests that interviews can be beneficial for exploring the perceptions of different stakeholders, and capturing rich, multi-layered accounts to provide in-depth insights into attitudes, and provide descriptions of reported practices. However, it is also commonly noted that participant perceptions may not always correspond to actual teaching and learning practices. Therefore, classroom observations have been suggested

as a means of probing complex interrelationships between variables and processes (Alderson & Wall 1993; Hughes, 1994), and have become a frequently used instrument in empirical washback studies (e.g. Alderson & Hamp-Lyons, 1996; Cheng, 1996; Huang, 2009; Qi, 2005; Tsagari, 2007; Watanabe, 1996). Direct observation allows researchers a more accurate view of instruction by providing them with the opportunity of collecting live data (Cohen et al., 2000 in Huang, 2009, p. 97). Watanabe noted that the type of observation instrument varies according to need, based on the contextual factors, the examination under inquiry, and the purpose of the research. In many cases, researchers had to devise a suitable observation tool (Cheng, 2004; Ferman, 2004; Saville & Hawkey, 2004; Stecher, Chun, & Barron, 2004; Qi, 2004) or modify an existing one. Communicative Orientation of Language Teaching Observation Scheme, COLT, (Spada & Frohlich, 1995) is a widely used and modelled observation tool among washback researchers (i.e. Burrows, 2004; Hayes & Read, 2004; Huang, 2009). COLT has been used to examine the extent to which different language classrooms display the features of the communicative approach to language teaching since it was constructed with the aim of differentiating communicative language teaching from the more teacher-centered and form-focused teaching (Huang, 2009).

The limited research based regarding washback on teaching in the local context and theoretical controversy as to how exams affect teaching in ELT literature prompt this study to examine the following research question: Is there a potential washback effect of the TRACE exam on teaching with respect to methods, materials, and tasks?

3. Method

Washback research indicates the necessity of mixed-method approach to explore complex washback phenomenon. As Creswell and Plano Clark (2006) advocate combining qualitative and quantitative approaches into mixed methods is used when a single approach does not provide a full picture. In this way, the mixed approach allows results of different phases of study to compensate for one another and enrich findings. Using both qualitative and quantitative research methodology, the present research investigates the research question in order to fill a gap in the literature regarding the washback effect of integrated assessment on teaching.

3.1. Research context and participants

This study was conducted at the Preparatory English Language Program (PEP) at an English medium foundation university in Istanbul, Turkey. The aim of the PEP is to improve students' general English ability and academic skills to meet the language requirements of their major fields of study. There are five levels within PEP which are aligned with the Common European Language Framework, namely elementary (A1), pre-intermediate (A2), intermediate (B1), upper-intermediate (B2) and advanced level (B2+). Learners are placed into intensive general English courses (elementary, pre-intermediate, and intermediate levels) based on the scores of the placement test. Others take the TRACE which is an institutionalised English language proficiency which determines whether the test-taker's skills and language level are sufficient for academic coursework. Thus, those who score at or above the cut off score of 65 out of 100 on TRACE are allowed to take their mainstream courses in their departments. Students

with scores between 50 and 65 are required to take advanced level course and scores lowers than 50 are directed to upper-intermediate course in the PEP.

TRACE entails four sections; introduction, reading, listening, and writing. All sections focus on one general topic such as psychology, sociology, environment, or business. In the introduction section, test-takers are given visuals regarding the topic of the exam and brainstorming worksheets to take notes. Then, they are provided with four reading texts with multiple choice comprehension questions. In the third section, test-takers listen to a lecture and take notes. The final section entails discursive essay writing using a variety of sources (ideas from readings, lecture notes of the listening, and notes from introduction section). Pictures and visuals in the introduction, reading passages and listening texts provide a substantial context for the test-takers. Having taken into account the real life academic needs of university students and the belief that the student needs to use the language to communicate actively and effectively, test developers integrated three language skills (reading, listening, and writing) in TRACE exam so as to reflect the authentic context. Replicating common assignment in academic coursework and test content, instruction in the PEP emphasizes an integrated skills approach in order to bring about positive washback upon teaching and learning. Consequently, the limited research base in the washback effect of theme-based integrated assessment prompted the present study to explore the test effects on teaching.

Participants included non-native and native language teachers employed full-time at the PEP. In order to improve students' general English ability and prepare them in terms of the academic skills needed to meet the requirements of their fields, these language instructors are assigned to teach 20 contact hours per week in classes of 15-18 students, each class hour lasting 50-minutes. Additionally, instructors are required to conduct 4 office hours for tutorials and focused feedback. Participants taught at upper-intermediate and advanced level English language courses in the PEP and each course lasted for 7 weeks. In both courses teachers made use of a variety of teaching materials including (1) commercialized books which focused on integrated skills (2) in-house prepared supplementary materials that were closely aligned with the requirements of the exam regarding content, text types, task types and format and (3) online materials with exam type tasks and questions given to the students through course management system as remedial work. There were 15 teachers teaching these levels in the PEP and all of them participated in the classroom observations. 14 took part in one-to-one interviews. They had varying teaching experience, ranging from two to more than twenty-five years, and Master's Degree in English Language Teaching (ELT).

Gender	Major	Years of experience	Highest Qualification	(Non)Native
--------	-------	---------------------	-----------------------	-------------

M	F	B.A	M.A	1-5	6-10	11-15	16+	M.A 14	NNSE	N
6	9	Literature 4 EFL 8 Applied Linguistics 2 Education 1	EFL 4 Education 2 Literature 1	2	4	8	1	PhD: 1	13	2

Table 1. Background information about language teachers who participated in observations

3.2. Data collection instruments

3.2.1. Interviews

One-on-one interviews were carried out with the participants teaching Upper-Intermediate Level ($n=7$) and Advanced Level ($n=7$) in the PEP. Interview questions mainly focused on teachers' perceptions regarding TRACE washback in relation to process (teaching materials and classroom activities) and product (how much learning happened). Questions aimed to examine teacher perceptions into the amount of learning in terms of reading, listening and writing, and reveal teachers' attitudes towards materials and tasks. Also, interviews aimed at surveying opinions regarding the correspondence between PEP objectives and the effect of TRACE on teaching and learning. Interviews were in English and lasted between 25-38 minutes.

3.2.2. Classroom observation

Washback researchers suggest conducting classroom observations as a means of data collection to investigate the webs of interrelationships between variables and processes (Alderson & Wall 1993; Hughes, 1994) for triangulation purposes. Thus, classroom observations were conducted in order to validate the teachers' self-reported data. In other words, observations were predicated on the underlying hypothesis that the respondents' statements about their teaching practices and the students' learning would be observable. Classroom observations were conducted through Communicative Orientation of Language Teaching Observation Scheme, COLT, (Spada & Frohlich, 1995). The class observation tool in this study was modelled after COLT Part A, in which "the observer makes a detailed note in real time on the activities and episodes that occur during the lesson, including the time taken for each one" (Hayes & Read, 2004, p.102). Qualitative input and a pilot observation were the preliminary steps in the development of the observation note-taking and analysis tool. Qualitative input consisted of theoretical resources from related research studies (Burrows, 2004; Hayes & Read, 2004; Huang, 2010) in which data was gathered through observations, as well as an interview with a colleague who had trialed the observation analysis form with a videotaped lesson. Similarly, the observation instrument and interview questions were shaped by feedback from a pilot session and suggestions resulting from piloting procedures.

The observation instrument consisted of six categories; time allotment, teaching materials, skill focus, activity, student work mode, and comments. Under 'activity', classroom activities and time allocation were noted. The 'activity' category was open-ended, excluding any pre-determined descriptors. The 'teaching materials' category pertains to commercial course books, supplementary materials, self-edited materials, internet materials, and others. The 'skill focus' category consisted of listening, speaking, reading, writing, integrated, vocabulary, and grammar. Finally, 'student work mode' was

described in terms of individual, pair, group, and choral. Classes consisted of 13-17 students. The duration of each activity was calculated as a percentage of total class time in minutes. One 50-minutes lesson was observed for 15 language instructors teaching at upper- intermediate ($n=9$) and advanced levels ($n= 6$) in PEP; a total of 800 minutes, or 13.3 hours.

Before the observations, brief meetings were conducted to obtain information regarding teachers' educational background and work experience. All classroom observations were video recorded, and data gathered from recordings and field notes were analyzed through quantitative and qualitative methods. COLT is used to examine the extent to which language classrooms display the features of the communicative approach, in line with its aim of differentiating communicative language teaching from more teacher-centered and form-focused teaching (Huang, 2009, p. 98). Upon completion of each observation, post-observation meetings were held to discuss the rationale behind the teaching activities employed, as a measure of confirmation.

3.3. Data analysis

The data gathered via semi-structured interviews were analysed using Bogdan and Biklen's (1998) framework. Transcriptions were e-mailed to instructors to check whether any details had been lost during the transcription process. Then, as a result of intensive and repeated reading, conceptual themes pointed out by the recurring words and ideas were identified. The emerging conceptual categories leading to the major themes, were classified under specific headings according to their relevance to the research question. Additionally, the results were quantified where possible to get a preliminary overview of data. Finally, the researcher used the data to produce detailed explanations addressing the research question.

Qualitative and quantitative analysis of the classroom observations was conducted following a two-step process. First of all, data was retrieved from the coding scheme in the observation instrument used in real-time observation. The data gathered provided a general description of the lessons, and served as a background for further analysis of the video-taped lessons. In the second step, the researcher watched the videos and specified the recorded length of time allocated to each teaching activity, and identified the number of minutes allotted to teacher and student talk. The focus was on what and how teachers taught, by analyzing elements of classroom activity. The unit of analysis was the particular classroom activity, since these represent the teacher's methodology and content. Elements of classroom communication shed light on the methodology employed in the classroom (the 'how'), and classroom activities, on content of instruction (the 'what'). When examining classroom activity quantitatively, the focus was on frequency of the language skills or knowledge focus of the activity, the student work mode during the activity and the time allocated to teacher and student talk. English medium teaching and learning is expected as an institutional policy; therefore, occasional student talk in Turkish was ignored. In qualitative terms, field notes were analyzed to identify differences among the 'what' and 'how' of different teachers.

4. Findings and discussion

4.1. Interviews

Findings revealed that teachers employed a variety of class tasks and activities. Teachers indicated that their repertoire involved a very wide variety of activities ranging from technology-integrated tasks (i.e. using Padlet; online posters to recycle vocab, summarizing content of units, and synthesize information across texts), preparing presentations, essay writing, listening to web-based lectures and note-taking to exam practice. Although these reported activities displayed a great variety of methodological approaches a few teachers focused on integrating skills deliberately or employing activities that fostered synthesizing information from reading and listening into writing if these tasks were not already enclosed in their materials.

In addition, when asked whether their selection of activities and tasks was influenced by TRACE, teachers provided a positive response. The majority (64%) stated that TRACE affected their selection of tasks, and a further 29%, agreed to a certain extent. Only one (7%) stated that TRACE had no effect on selection of tasks. Most (64%) responded that helping students perform better in TRACE was their aim in all tasks. All teachers in the advanced level indicated that exam practice in the form of previous TRACE exams and timed reading, listening, and writings were believed to result in boosting scores. Furthermore, majority indicated that they believed classroom activities would improve students' scores because teaching was geared towards the exam. One participant stated:

I think the writing sessions can be given as an example. Because in the writing part, you know, they write an essay related to the topic they read in reading and listening. So, in our lessons they write an essay related to the topics that we covered in the lesson. Again in listening they took notes and then answered the questions, lectures. Towards the end of the course, I give the supplementary readings to students as a test. I tell them to imagine they are in the exam.

Additionally, activities that were assumed to be beneficial were strategy training on answering specific exam questions (such as cross textual reading), and using the published exam preparation booklets from other tertiary level institutions which had similar exam tasks.

In response to a set of questions which aimed at surveying correspondence between teaching, learning and success on TRACE, teachers concurred that the content of the course and TRACE were similar; both the exam and the course content were theme-based, involved discourse synthesis and had identical objectives that are tested in the exam. It was pointed out that even the question types used in course materials and those in the exam were the same.

However, when teachers were asked what changes they would initiate in their teaching if there were no requirement to take TRACE at the end of the year, nearly all suggested that they would use more creative and communicative activities, and more allocation of time to speaking tasks. It can be inferred that these responses may indicate a potential negative washback of TRACE, which was characterized as narrowing the curriculum purely to tested skills. Another frequently mentioned potential change included focusing on tasks such as creative writing, role-plays, and drama activities,

which would foster students' creativity. Many teachers stated that if it weren't for the exam, they would abandon multiple choice questions and ask more open-ended questions in order to engage students in more critical thinking, responding to texts, and interacting with the texts as a reader. Similarly, it was suggested that there would be more tasks which require students to link concepts to their own experiences, as well as extensive reading and summary writing. Furthermore, some teachers stated that they would shift the focus on vocabulary away from academic wordlists, and spend time on more productive activities. Many stated that they would make use of more communicative, productive student centred activities, taking into account student preferences. In other words, students would be provided with more involvement in decisions on tasks. It was also claimed that there would be more variety in terms of tasks, and materials would be related to daily life, and to students' academic departments, and "not just [to] this test". All respondents conceded that they utilized exam oriented teaching methodology, such as practice exams and exam strategy training.

In an attempt to unveil attitudes to teaching materials used in class, respondents were asked whether the course materials contributed to their students' learning English. An overwhelming majority (79%) voiced negative perceptions about contribution of the course books used at both levels to learning, stating that the course books were not effective in improving students' English. A wide array of reasons was put forth, ranging from the inclusion of unfamiliar, uninteresting, conceptually difficult topics, texts that are challenging for the students' level, and short listening texts in comparison to the ones used in the test, to a lack of variety in reading texts and explicit skills training. However, the most commonly stated reason was that commercial course books were not closely aligned with the test content. Many teachers expressed their dissatisfaction for not being able to exploit some tasks in the books, such as responding to open-ended questions, because they were not perceived by the students as replicating exam-type tasks and questions. One of the teachers suggested: "We don't use most of the book, and the reading and listening texts that are most commonly used from the books are supplemented it with multiple choice questions". It was claimed that the mismatch between the course book tasks, exercises and questions types, and those of the TRACE increased teacher load, by requiring them to develop materials that replicate exam type tasks and questions. Respondents often regarded writing and speaking exercises of the course book as 'irrelevant' to content of the exam, which indicated negative washback in the form of a narrow focus on curriculum. On the other hand, with respect to in-house produced supplementary materials, all respondents expressed positive attitude. One teacher claimed that the reason for this positive attitude was that supplementary materials had "face value".

It looks like something which they would come across in the exam. It also has the same, similar length to the listening or reading they come across in the exam. Therefore, it provides a high motivational value to the students and I find that they're more receptive to supplementary materials. Whereas in our book the texts are generally short, and listenings are generally shorter and they are of a high level.

Teachers often stated that their supplementary materials for each level were tailor-made made for students' exam-oriented expectations, their needs and the course objectives. It

was also suggested that supplementary materials reflected the specifications of the exam so well that “supplementary materials give an idea about TRACE if you are a new teacher”. One of the respondents stated: “If we didn’t have supplementary materials, I would be lost”. It was concluded that TRACE led to a positive test effect on materials, which were designed in close alignment with its specifications and underlying principles.

Finally, when teachers were asked about how well they thought the course supported their students’ learning and success in TRACE, all claimed that the preparatory program (especially for the advanced level students) gave support in the language and skills necessary to pass TRACE. Remarks about close alignment of course content and the exam content included the following: “I think the course helps students greatly because we do not expect them to do something really different from what we have done in the classroom. Note-taking and answering the questions, I think yes, everything is parallel with what we did in class and TRACE.” However, some teachers voiced their belief that the program was beneficial in passing the test but not learning ‘real English’ as in the following comment.

When I think about our system, our aim is to teach students English, right? But, during the modules, it changes, because students are concentrated on the exams, they want to pass the module. So, in a way we start to give exam preparation. When we think about our supplementary materials and our course-books- they are also supplemented- with multiple choice questions... But, if you are teaching someone English and if your aim is to really teach English, you should not have lots of multiple choice questions, because [it] is not real teaching. They need to understand the content; they shouldn’t choose the answers from the options. Because in real life, it is not like this. Nobody gives you the options after asking something. So, I don’t know...we need to change something in the system I think.

4.2. Classroom observations

In most of the observed classes, the skill focused on was listening followed by reading. All, except one, made use of supplementary materials aligned to TRACE content in terms of text length, genre, and exam type multiple choice questions. The majority (73%) seemed to deal with skills in isolation if the teaching material lacked a task which required the synthesized process of writing with other modalities such as listening and reading. In many lessons teachers covered materials without integrating them with other activities or sources to practice four skills in a more balanced approach. In contrast, the remaining 27% tended to make use of a variety of other sources (e.g. short videos, texts from the internet and other books and photographs) and taught with a deliberate focus on integrated skills oriented approach by incorporating tasks that demanded reading and listening-to-writing and speaking.

Also, data revealed variety across classrooms in terms of students’ work during instruction. Although some teachers were inclined to ask students to work in pairs and groups, the majority opted for individual working mode, requesting students to work alone. There were few instances of group work. However, working alone and responding to teachers’ questions in a choral mode were the most frequently observed student activity in class. The interactional pattern in most of the classrooms followed the

traditional Teacher Initiation-Learner Response-Teacher Follow Up (IRF) communication pattern (Sinclair & Coulthard, 1975 in Huang, 2009). Also, the classroom interaction had a great tendency to be exclusively between the teacher and the students, whereas little communication took place among students themselves.

As for time allotment to teacher and students talk, it was observed that there was a substantial variation, with the former considerably higher. The very short responses in student turns in most classes, seen in the timed analysis, can be regarded as indicators of limited time allocation to communicative activities. As it can be seen in the table below, on average, teacher talk took up 28', 21" in upper intermediate level classrooms, and 28', 10" in advanced level of the 50-minutes class, whereas student talk time was considerably less in these levels, 6',67" and 5',30" for upper and advanced level respectively. Time allotment to teacher and student talk in observed upper-intermediate and advanced level classrooms is also summarized in Figure 1 below. This variation can be attributed to activities carried out in the lessons as speaking was never the main skill focus of the lessons observed. The classes observed were mainly devoted to skills that appear in the exam. This also accounts for the relative lack of student-to-student interaction pattern in classroom communication. Therefore, corroborating interview findings, the elements of classroom communication corresponded to a negative washback effect of TRACE in narrowing of the curriculum, since the data reiterated that what was tested was taught.

Table 2. Summary of class observations

	No of students	Skill/ Knowledge Focus	Student Modality	Teacher talk time	Student talk time
Upper Teachers					
T1	15	Listening, Vocabulary	Group, Individual, Choral	19',11"	13',54"
T2	15	Listening	Individual, Choral	31',30"	3',25"
T3	14	Listening	Individual, Choral	29',40"	5',22"
T4	15	Integrated	Pair, Individual	22',38"	9',15"
T5	13	Listening	Individual, Choral	27',06"	5',40"
T6	14	Reading	Individual, Choral	26',15"	6',32"
T7	11	Reading	Individual, Choral	32',56"	6',38"
T8	11	Listening, Grammar	Group, Individual	34',55"	4',10"
Advanced Teachers					
T9	14	Integrated	Individual, Pair, Choral	22',58"	4'
T10	12	Integrated	Individual, Pair	28',18"	8',08"
T11	15	Writing	Individual, Pair	3'3"	3',35"
T12	14	Integrated	Individual, Pair	31',2"	4',32"
T13	11	Reading	Individual, Group, Choral	24',22"	6',15"
T14	14	Reading	Individual, Group	25',15"	9',05"
T15	12	Listening	Choral	32',26"	2',15"

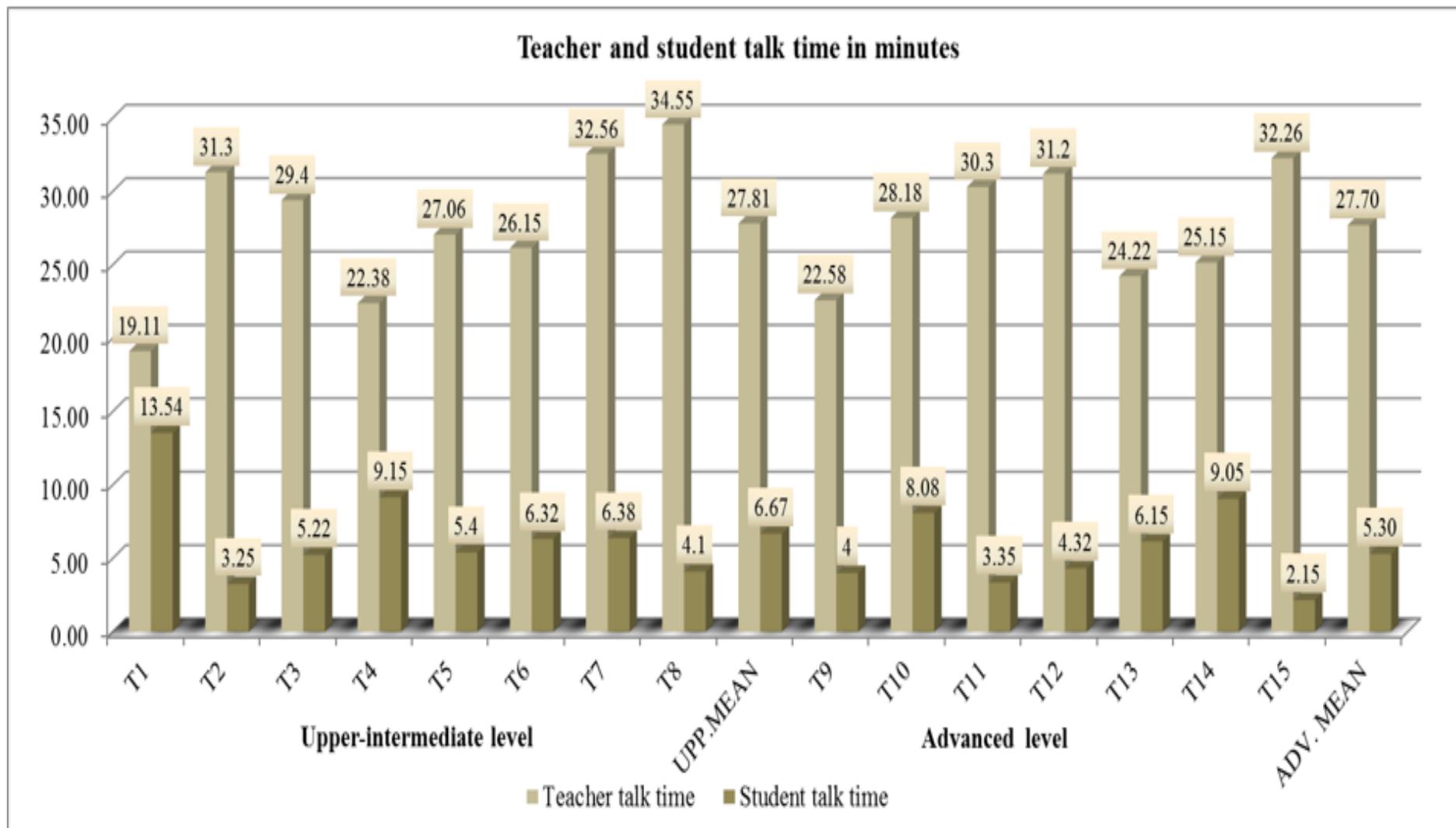


Figure 1: Bar graph showing time allotment to teacher and student talk in observed upper-intermediate and advanced level classrooms

Observations revealed some marked differences among teachers in terms of instruction in four areas: the exploitation of materials, focus on integrated skills and discourse synthesis, explicit strategy training, and exam coaching.

4.2.1. Exploitation of materials

Confirming interview findings, data analysis revealed strong positive washback effect of TRACE on in house-prepared supplementary materials. It was evident in the observations that nearly all teachers employed supplementary materials. However, some teachers displayed great variety in terms of exploitation of teaching materials in comparison to others, who moved through the materials in a predictable and linear fashion. To illustrate, some teachers were observed to raise awareness on critical thinking and reading strategies by replacing the exam type multiple choice questions with open-ended ones projected on power point slides. Teachers expressed that they wanted to raise awareness on the importance of critical reading skills through responding to open-ended questions, though some students resisted such questions, as these did not mirror exam type items. Some teachers scaffolded the given supplementary materials/ worksheets. As an illustration, one of the teachers chose to first ask students to read, find the main idea, and discuss it in groups. Afterwards, they were provided with the supplementary materials consisting of exam type multiple choice items. The teacher required them to match the paragraphs with given main ideas as a follow up of their discussion. Similarly, although summarizing and paraphrasing were not directly tested by TRACE, a few required students to summarize paragraphs in one sentence as an alternative to multiple choice.

Observations also unveiled disparities with regard to the use of teaching materials. Although the majority of the teachers relied on utilizing the supplementary materials without modifications, some teachers resorted to the integration of extra oral and written sources. Students were required to work across different resources and practice different language skills, for example, by answering specific questions on an authentic video as a lead-in to a reading supplementary material. It was observed that some teachers displayed a deliberate effort towards integrating skills through use of authentic materials such as videos and texts from newspapers and blogs. Some teachers pointed out that students would be required to read and listen to different sources and make use of these in the writing section of the TRACE exam, showing a clear understanding of TRACE specifications, and an ability to adapt their methodology to suit the theme-based integrated nature of the exam. Based on classroom observations, it was concluded that some drew on a more restricted range of teaching materials, whereas others supplemented with a variety of other resources. This reiterates the interview finding that teachers play an important role in mediating washback.

4.2.2. Focus on integrated skills and discourse synthesis

Data from questionnaires on washback revealed that teachers claimed to have a solid understanding of the requirements of TRACE. However, later, during observations, it was observed that a deliberate focus on integrated skills approach and

discourse synthesis was completely lacking, except for one instance. Very few teachers seemed to have made an effort to adopt an integrated skills approach. Actually, it seemed that most instruction directly depended on supplementary materials because the majority strictly followed them without making modifications. One rare example was making use of a short listening as a lead-in to the reading text which was the focus of the lesson. Afterwards, the teacher required the students to refer to the ideas in listening and reading in a follow up speaking activity. This illustrates that some teachers were able to adapt teaching materials to a certain degree and have students practice integrated skills. Even if the major teaching objective was to focus on a certain skill, some teaching included a broader range of skills, and displayed an even coverage of four skills through lead-in and follow up tasks. Such classes were observed to be less-teacher centred, since more classroom time was allocated to activities that required students to engage in speaking and writing.

4.2.3. Explicit strategy training

In some classes, covering the assigned teaching material was prioritised, but in others, the focal points of the lesson were addressed through explicit skills training. To illustrate, in one of the classes, after the lead-in task, the teacher asked about the number of reading texts in TRACE, length of these texts, questions types and time allotment. After a brief discussion, the teacher handed out the reading text and asked the students about the time needed to read and understand the text. Then she asked whether it would take less time if she gave out the multiple-choice questions. Students stated that it would take considerably shorter time because they would only read the parts related to the questions. Then the teacher distributed a small rectangular printed paper to each student and asked them to punch a small hole in the middle. She said that it was a new reading device which would enable them to read better and asked them to read the text by placing the paper on the text and read through the hole. After experiencing the difficulty of focusing on isolated words that were visible through the hole, the students stipulated that reading without understanding the meaning-relations between words and sentences was quite difficult. They discussed the similarities of this experience with reading a text for specific answers without grasping the main ideas and suggested ways to read more efficiently. In addition, it was observed that some teachers raised awareness on skills training by engaging students in activities such as examining text titles, headings and visual clues in order to guess the content, exploring main ideas (examining key words, analyzing examples given and asking questions), looking for specific information, and deciding where it is usually located in the text, and having students discuss the strategies they employed to read more efficiently. On the other hand, in some classes, teachers read with students the texts paragraph by paragraph, directing students' attention to context clues to understand the main ideas when they encountered difficulty in decoding the main ideas.

It was also observed that some teachers explicitly taught students how to recycle and recall vocabulary through modelling, working on word forms and vocabulary

games. Similarly, one teacher focused on note-taking skills strategy training, by asking students to compare their notes, and elicited techniques for efficient note-taking.

Consequently, observations indicated that while some teachers focused on explicit note-taking, vocabulary, and reading strategy training, others simply set the task and then gave the answers. In the former, students' awareness on the learning process were more likely to be raised, whereas in the latter, the focus seemed to have shifted directly to product.

4.2.4. Exam coaching

Many teachers, especially in the advanced level, were observed to utilize test-related activities, include drawing a link between class activities and the exam, raising students' awareness of the tasks by referring to the exam, giving the tasks under exam conditions, evaluating performance in comparison to exam related task, and pointing out exam skills and strategies. This finding is also corroborated with outcomes of teacher interviews.

5. Conclusion

In relation to the potential washback of the TRACE interviews and classroom observations indicated that there were both positive and negative test effect exerted on the teaching methodology, materials, and classroom activities. In addition to positive washback on materials, this study also found negative washback in the form of narrowing of curriculum. Observation and interview data concluded that the test narrowed curriculum offerings, since teachers were employing instructional approaches they would not have used without the exam. This conclusion is supported by the limited variety of classroom organizational patterns, time allotment to tested skills at the expense of communicatively oriented activities, and practice of speaking (which is not tested on the exam). It can also be inferred that for some teachers, TRACE exerted a negative washback and changed not only what they taught but also how they taught, but others were less affected, since even though they were prepping students for a high-stakes proficiency exam, their teaching was largely unchanged. Consequently, the exam exerted negative washback for some teachers who were more focused on tested skills and were coaching their students to increase scores instead of deeper learning. Findings imply that the teacher's methodology and choice of classroom activities were adapted to the tested skills and contents of the test. Teachers stressed that their methodology was affected by requirements of the exam, which, for example, led them to omit or reduce the time spent on tasks that required speaking which received little attention since it did not directly contribute to the requirements of the exam. Teachers explained that they had the tendency to employ materials and rely on activities compatible with the principles underlying the exam, and placed more value on tested skills. These findings are in line with previous research studies which concluded that, when teachers perceive that teaching and learning are 'circumscribed and controlled' by the examinations, and consider that

students are exam-oriented, the focus of teaching and learning shifts towards the tested skills (Calder, 1990 in Cheng Cheng & Curtis, 2004, Hayes & Read, 2004)

The present study also implied that the exam affects both content (what gets taught) and methodology (how teachers teach), but amounts and types of washback varied depending on the teacher involved. These findings are in line with previous studies on washback in language education (Alderson & Hamp-Lyons, 1996; Andrews, 1994; Stecher, Chun & Barron, 2004; Watanabe, 1996b), rather than those that found no direct connection between the test and teaching (e.g. Chen, 1997; Wall & Alderson, 1993).

As the research findings are based on analysis of data regarding a proficiency test in a specific educational context, it could be argued that it is not possible to generalize the findings to the broader English language teaching and testing populations in other contexts. However, researchers (e.g. Perrin, 2000; Tsagari, 2006) argued that any washback research is innately context-based, and therefore that investigating those forces in a specific educational context may shed light on similar forces in a broader context. The focus of the study was teasing out the washback of a context-specific theme-based integrated tertiary level proficiency test which mirrors authentic language use in academic settings. Findings can be related to other contexts, and may have implications for EFL students, teachers, and test designers in similar contexts in which it is aimed to engineer positive washback.

Since “teachers are the final arbiters of policy implementation” (Menken, 2008, p.401), teacher perceptions play a significant role upon the motivation and effort made to improve teaching methodology and influence learning (Clark & Peterson, 1986; Lane, Parke & Stone, 1998 in Pan & Newfields, 2012). Wall and Horak (2007) indicated that there may be times when nature of the tests and intended washback effect might be ambiguous to the teachers, and they highlighted the importance of communicating intentions of the testers and underlying principles of tests to teachers. Teachers should familiarize themselves with the exam specifications, and be aware of the underlying constructs, as well as the purpose and the value of the exam. Consequently, they will be able to inform students of exam requirements and explain the underlying constructs (e.g. discourse synthesis and how this would contribute to their academic studies). Acknowledging requirements would also help teachers modify and adapt classroom materials and exam-preparation materials, and integrate learning and testing materials. Also, teachers should balance teaching and testing, avoiding excessive focus on exam-coaching or and narrowing the curriculum to the tested skills. Teachers should devise efficient methods for resisting pressure from students to allocate a disproportionate amount of class time to exam preparation activities (mock exams, exam-strategy training, revision, multiple choice items) and focusing on tested skills at the expense of communicative activities. To this end, teachers could give information and feedback to diagnose learning needs, rather than equating language proficiency purely with test scores. Teachers could make use of a variety of channels to assess their students’ performance and share their observations with the students.

Future studies can focus on the perception of other stakeholders, especially students' and examine how tests affect their learning process as well as outcomes. In addition, due to multi-faceted and complex nature of the washback phenomenon, washback studies should adopt multiphase and longitudinal research designs which utilize a variety of data collection methods. More empirically grounded research is warranted to explore how teachers' attitudes to and understanding of exams affect the way they teach, and why tests, as curricular innovations, might not bring about positive washback effect that is intended.

References

- Akpınar, K.D. & Cakildere, B. (2013). Washback effects of high-stakes language tests of Turkey (KPDS and UDS) on productive and receptive skills of academic personnel. *Journal of Language and Linguistic Studies*, 9(2), 81-94.
- Alderson, C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing*, 13(3), 280-297.
- Alderson, C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129.
- Azadi, G. & Gholami, R. (2013). Feedback on washback of EFL tests on ELT in L2 classroom. *Theory and Practice in Language Studies*, 3(8), 1335-1341.
- Bailey, K.M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- Bogdan, R. C. & Biklen, S. K. (1998). *Qualitative research in education: An introduction to theory and methods*. Needham Heights, MA: Allyn & Bacon.
- Burrows, C. (2004). Wash-back in classroom-based assessment: A study of the washback effect in the Australian adult migrant English program. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and methods* (pp. 147-170). Manwah, New Jersey: Lawrence Erlbaum Associates.
- Chen, L.M. (2002). Washback of a public exam on English teaching. (ERIC Document Reproduction Service No. ED472167).
- Cheng, L. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54.
- Cheng, L. (2004). The washback effect of a public examination changes of 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). Manwah, New Jersey: Lawrence Erlbaum Associates.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge, UK: Cambridge University Press.
- Cheng, L., & Curtis, A. (2004). 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. 3-18). Manwah, New Jersey: Lawrence Erlbaum Associates.
- Clark, C.M. and Peterson, P.L. (1986), Teachers' Thought Processes. In M.C. Wittrock (Eds.), *Third handbook of research on teaching* (pp. 255-296). New York: Macmillan.
- 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. 3-18). Manwah, New Jersey: Lawrence Erlbaum Associates.

- Fullilove, J. (1992). The tail that wags. *Institute of Language in Education*, 9, 131-147.
- Green, A. (2007a). *IELTS Washback in Context Preparation for Academic Writing in Context*. Cambridge, UK: Cambridge University Press.
- Green, A. (2007b). Washback to learning outcomes: a comparative study of IELTS preparation and university professional language courses. *Assessment in Education*, 14(1), 75-97.
- Gu, Y. P. (2014). The unbearable lightness of the curriculum: what drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school? *Assessment in Education: Principles, Policy and Practice*, 21(3), 286-305.
- Hayes, B., & Read, J. (2004). IELTS test preparation in New Zealand: preparing students for the IELTS academic module. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and methods* (pp. 129-148). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Huang, L. (2009). Washback on teacher beliefs and behaviours: Investigating the process from a social psychology perspective. (Unpublished PhD thesis), Lancaster University, Lancaster, UK.
- Hughes, A. (1994). Backwash and TOEFL 2000. Unpublished manuscript, commissioned by Educational Testing Service (ETS). University of Reading.
- Karabulut, A. (2007). Micro level impacts of foreign language test (university entrance examination) in Turkey: a washback study (Master's thesis), Iowa State University, Ames, Iowa. <http://lib.dr.iastate.edu/rtd/14884/>
- Lam, H.P. (1994). Methodology washback and insider's view. In D. Nunan, R. Berry, & V. Berry (Eds.), *Bringing about change in education* (pp. 83-99). Hong Kong: University of Hong Kong.
- Leki, I., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28, 81-101.
- Leki, I., & Carson, J. (1997). Completely different worlds: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly*, 31, 39-69.
- Li, X. (1990). How powerful can a language test be? The MET in China. *Journal of Multilingual and Multicultural Development*, 11, 393-404.
- Madaus, G. F. (1988). The Influence of Testing on the Curriculum. In L. N. Tanner (Eds.), *Critical Issues in Curriculum* (pp. 83-121). Chicago, IL: The National Society for the Study of Education.
- Manjarres, N.B. (2009). Washback of the foreign language test of the state examinations in Colombia: A case study. *Arizona working papers in SLAT*, 12, 1-12.
- Özmen, K. (2011). Washback effects of the inter-university foreign language examination on foreign language competences of candidate academics. *Novitas-ROYAL Research on Youth and Language*, 5 (2), 215-228.
- Perrin, G. C. (2000). The effect of multiple choice foreign language tests of listening and reading on teacher behaviour and student attitudes. (Unpublished PhD thesis) Lancaster, England, Department of Linguistics and Modern English Language, Lancaster University.
- Plakans, L. (2008). Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, 13, 111-129.
- Plakans, L. (2009a). The role of reading strategies in integrated L2 writing tasks. *Journal of English for Academic Purposes*, 8, 252-266.
- Plakans, L. (2009b). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26, 561-587.
- Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, 44, 185-194.

- Plakans, L., & Gebril, A. (2012). A close investigation into source use in integrated second language writing tasks. *Assessing Writing*, 17, 18–34.
- Qi, L. (2005). Stakeholders' conflicting aims undermine the washback function of a high-stakes Test. *Language Testing*, 22, 142–173.
- Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: research contexts and methods* (pp.147-170). Mahwah, NJ: Lawrence Erlbaum.
- Saville, N., & Hawkey, R. (2004). 'The IELTS Impact Study: Investigating washback on teaching materials', In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and methods* (pp. 19-36). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Sevimli, S. E. (2007). The Washback Effects of Foreign Language Component of the University Entrance Examination on the Teaching and Learning Context of English Language Groups in Secondary Education: A Case Study (Master's thesis), Gaziantep University, Gaziantep. <https://tez.yok.gov.tr/UlusalTezMerkezi/>.
- Shohamy, E. (1993). *The Power of Tests*. Washington DC: National Foreign Language Center.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*, 13(3), 298-317.
- Smith, M L (1991). Put to the Test: The Effects of External Testing on Teachers. *Educational Researcher*, 20(5), pp 8-11.
- Spada, N., & Fröhlich. M. (1995). *Communicative orientation of language teaching observation scheme (COLT)*. Australia: Macquarie University National Centre for English Language Teaching and Research
- 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.
- Stecher, B., Chun, T., & Barron, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington State. In L. Cheng & Y. Watanabe & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 53-71). Mahwah, NJ: Lawrence Erlbaum.
- Taylor, L. (2005). Washback and impact. *ELT Journal*, 59(2), 54–155.
- Tsagari, D. (2007). Review of Washback in Language Testing: What Has Been Done? What More Needs Doing? Washington, DC: Center for Applied Linguistics (ERIC Document Reproduction Services No. ED 497709).
- Wall, D. & Alderson, J.C. (1993). Examining washback: The Sri-Lankan impact study. *Language Testing*, 10, 41-69.
- Wall, D. & Horak, T. (2007). Using baseline studies in the investigation of test impact. *Assessment in Education: Principles, Policy and Practice*, 14 (1), 99-116.
- Watanabe, Y. (1996). Does grammar translation come from the entrance examination? Preliminary findings from class-room based research, *Language Testing*, 13(3), 318-333.
- Watanabe, Y. (2000). Washback effects of the English section of the Japanese university entrance examinations on instruction in-pre-college level EFL. *Language Testing Update*, 27, 42-47.
- Watanabe, Y. (2004a). Methodology in washback studies. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and methods* (pp. 19-36). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Watanabe, Y. (2004b). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research Contexts and methods* (pp. 129-148). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Wall, D. (2005). *The impact of high-stakes examinations on classroom teaching*. Cambridge: Cambridge University Press.

Wesdorp, H. (1982). Backwash effects of language testing in primary and secondary education. *Journal of Applied Language Study*, 1(1), 40-55.

Appendix A.

A.1. Semi-structured teacher interview topics and questions

1. Opening

Introduction

Key points of the study, purpose, confidentiality, media and timing

2. Instructional sensitivity of TRACE

How much learning takes place in terms of:

Reading

1. Do you think your students have improved their reading ability in the course? Why? Why not?
2. What do you think they have learned in terms of reading skills in the course? Can you give some examples?

Listening

3. Do you think your students have improved their listening ability in the course? Why? Why not?
4. What do you think they have learned in terms of listening skills in the course? Can you give some examples?

Writing

5. Do you think that your students have improved themselves in writing?
6. What do you think they have learned in terms of writing skills in the course? Can you give some examples?

WASHBACK: Attitudes Towards Teaching materials & Tasks

7. Think about the course materials (books, supplementary materials, web activities...etc.) Do you think that they have contributed to your students' learning English? Which ones were the most beneficial in your opinion? Why?

Coursebook:

Is the course book a B2 level book?

Does the book prepare your students for TRACE?

Supplementary materials:

Are the supplementary materials used for TRACE? Why?

Would you use a supplementary material if you didn't have to prepare students for TRACE?

Vocabulary & Grammar Booklet:

Do they prepare students for TRACE?

Do you think that they have contributed to your students' learning English?

Blended Learning Web Materials

Do they prepare students for TRACE?

Do you think that they have contributed to your students' learning English?

8. What kind(s) of reading, listening & writing activities and tasks have you done in the class?

9. Do you think that your selection of activities and tasks are affected by TRACE?

10. Do you remember any task that was directly related to the test and it may help your students improve their scores?

Correspondence between teaching-learning and being successful on TRACE (The relationship between objectives of the program, learning and TRACE)

11. Do you think that content of the course (what you teach in class) and TRACE are similar? How?

12. Do you do any special preparation for TRACE? If not then: How do you think what your students learned in the course may help them in TRACE?

13. What changes would you initiate in your teaching if your students didn't take TRACE at the end of the year?

14. In your opinion to what extent did the course support your students to learn English and be successful on TRACE? How well did the course prepare them to be successful on TRACE?

Round up and thanks

Copyrights

Copyright for this article is retained by the author(s), 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 (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).