



## WASHBACK EFFECTS OF THE INTER-UNIVERSITY FOREIGN LANGUAGE EXAMINATION ON FOREIGN LANGUAGE COMPETENCES OF CANDIDATE ACADEMICS

Kemal Sinan ÖZMEN\*

**Abstract:** This paper presents a study investigating the washback effect of Inter-university Foreign Language Examination (ILE) on candidate academics in Turkey. Over the last decades, washback studies have become even more critical in the field of applied linguistics due to unanimous remarks of the international research studies indicating the critical role of the high-stakes tests on students, teachers and societies. Also an effective understanding of washback is now much more important because of the increasing number of international testees taking these examinations. The present study follows twelve candidate academics during their preparation for ILE. In a qualitative design, the findings show that ILE has a negative washback effect on numerous micro and macro level variables.

**Keywords:** Washback effect, high-stakes tests, candidate academics

**Özet:** Bu makale Üniversiteler Arası Kurul Yabancı Dil Sınavı'nın (ÜDS) akademisyen adayları üzerindeki ket vurma etkisine (washback effect) dair bir araştırmayı sunmaktadır. Son dönemlerde, ket vurma etkisi uygulamalı dilbiliminde daha fazla önem kazanmaya başlamıştır çünkü uluslar arası araştırma çalışmalarının hepsi hayati sınavların öğrenci, öğretmen ve toplum üzerinde çok ciddi etkileri olduğunu dile getirmektedir. Ayrıca uluslar arası İngilizce sınavları ve bu sınavlara giren kişi sayısında artış sebebiyle ket vurma etkisini etkin bir biçimde anlamak büyük önem arz etmektedir. Bu çalışmada on iki akademisyen adayı ÜDS sınavına hazırlık dönemlerinde izlenmiştir. Nitel bir desen dahilinde araştırma sonuçları, ket vurma etkisinin çeşitli küçük ve büyük çaplı değişkenler üzerinde negatif etkisi olduğunu göstermektedir.

**Anahtar sözcükler:** Ket vurma etkisi, önemli sınavlar, akademisyen adayları

### Introduction

Washback or backwash refers to effects of language testing on teaching and learning (Aldersen & Wall, 1993). During the last few decades, high-stakes tests have been growing in number and gaining importance as student mobility has increased notably all around the world, and as institutions seek a official document to evaluate individuals' competences for a specific job. Also such tests are still the major tools utilized for employment and promotion in most countries. In this context, analyzing the washback effect of high-states tests is critically important for test-takers, institutions designing tests, and of course for employers and other institutions who use these tests to employ or educate individuals. To illustrate the effects of this test, this paper reports on a research study in which the possible washback effects of *the inter-university foreign language examination* (ILE), designed and applied by Higher Education Council Student Selection and Placement Center in Turkey, were investigated in a qualitative study.

---

\* PhD, Lecturer, Gazi University, Faculty of Education, ELT program, Ankara, Turkey, sinanozmen@hotmail.com.

The literature on washback reveals critical insights, mostly indicating that the effects of the tests are not easy to measure, analyze and accordingly infer regularities. Nevertheless, the studies may be grouped under certain categories. First group of studies would be the ones in which the high-stakes tests such as TOEFL or IELTS are investigated to find out whether these tests exert positive, negative or neutral effects on testees (Cheng, Watanebe, & Curtis, 2004; Madaus & Kelleghan, 1992; Nolen, Haladyna, & Haas, 1992; Shepard, 1990). Such studies mostly focus on areas in which students display certain difficulties in developing their communicative competences. As an example, Yıldırım (2010) carried out a study to investigate whether “The Foreign Language Examination<sup>1</sup>” in Turkey exerts a positive or negative impact on test-takers and found that the exam had certain negative effects on students’ foreign language competences. Then second group of studies would be the ones that adopt a remedial approach; that is, such studies aim at modifying examinations that exert a negative effect so that these exams may have a positive effect on the students, teachers and hopefully the institutions and the society. A down-to-earth example would be the modification that was made on TOEFL. Before 2005, TOEFL examination used to offer a grammar section, but then this section was replaced with a speaking part with the hope of getting test takers to focus on developing their academic speaking skills and strategies. Third group of studies may be the ones that specifically focus on the teacher's side. Washback studies may appear to concentrate on solely validity concerns and students, which are of course two of the major variables: The test itself and the testees. However, research into washback reveals another key actor of the phenomenon, which is the teacher and the teaching style s/he adopts (Watanabe, 2004).

In order for us to have a better understanding of the influence of examinations on our learning and teaching practices, we should recall the distinction made by Wall (1997) between “test impact” and “test washback” with regard to the scope of the effects. Wall (1997) claimed that *test impact* refers to any kind of effects that an examination may have on individuals, policies and practices in any given educational system or society, while *test washback* is the effect of tests on teaching and learning. It is then quite easy for a professional reader to infer that not only is washback effect easy to measure but also modifying exams to exert a positive effect is not a simple, recipe-based action. The reason why washback is a complicated phenomenon surely lies in its nature, which encompasses many influencing variables (Shepard, 1990). Therefore, the contemporary testing literature gives a special emphasis on the importance of washback studies, as the latest studies show clearly that it is such a critical variable that washback may also be considered as a part of test validity (Shepard, 1997).

ILE is an academic English test that is offered to thousands of candidate academics and to those who want to study in a graduate institution at master's or doctorate levels in Turkey. Recognized in all academic institutions and bodies in the country, the test also is taken by tenure-track professors or Ph.D. holders who aim to apply for the associate professorship examination that is offered by Inter-university Board under the auspices of Council of Higher Education. This examination is quite a critical one for the individuals aspiring to pursue further higher education at graduate level or become an academic. Therefore, investigating the possible washback effects of ILE on candidate academics may shed light on 1) how ILE shapes foreign language learning experiences of these group of learners, 2) whether ILE creates any positive, negative or neutral effects on the candidate academics' communicative competences. Considering the fact that ILE

---

<sup>1</sup> TR Acronym: YDS

and similar examinations such as "The Foreign Language Examination for Civil Servants<sup>2</sup>" (FLEC), "The Foreign Language Examination" (FLE) have been offered simply to millions of Turkish students over the past decades, it is quite surprising that little or no research was conducted to find out whether these examinations are simply beneficial or harmful for foreign language learners in Turkey. These and related information are mostly either unknown or incomplete for now. In this respect, the present study carries an important mission by analyzing the washback effect of ILE on candidate academics.

## Background

It is important to bear in mind that washback is a neutral term, which may refer to positive or intended (Bachman & Palmer, 1996) or beneficial effects (Buck, 1988; Hughes, 2003) or to harmful (Buck, 1988), negative or unintended effects (Bachman & Palmer, 1996; Hughes, 1989). Washback was differentiated from test impact by Bachman and Palmer (1996) who, with McNamara (1996), Hamp-Lyons (1997) and Shohamy (2001), believed that it should be placed within the scope of impact. While test impact may possibly happen at a 'macro' level such as in a social or an institutional context, washback can be observed solely at the 'micro' level of the individuals (mostly teachers and students). Yet another theoretical framework concerning washback was offered by Davies (1995), who mentions the backward direction of washback effect, which refers to the basic idea that the nature and the content of the test shape the choices of the students from the curricular content and of learning strategies. Similarly, Messick (1996) points out the bidirectional nature of the washback claiming that any teaching and learning actions taken in a learning context as a result of the nature and content of the test refer to this bidirectional nature.

The pursuits of exerting a positive effect by conducting certain modifications are quite challenging for institutions and test-designers. In order for those responsible to create a positive effect, the insights of the concerning literature should be taken into consideration: Messick (1996) believes that a positive washback effect can be created if the tasks and activities utilized in teaching language are included as test items in the examination. This means not only should the students be tested on what they were taught, which is the basic motto in testing, but also they should be tested *in the way* they were taught, which is a motto offered by scholars studying washback. Therefore, it is not that difficult to infer that, in addition to the methodological concerns in terms of materials utilized in teaching and testing, the methodology that the teacher adopts is also critical in exerting a positive effect. In spite of the effective materials and a good test, a negative washback effect can still be measured in a specific testing context. "A poor test may be associated with positive effects and a good test with negative effects..." (Messick, 1996, p. 242). Messick (1996) continues claiming that a good test can bring about a negative effect due to other factors, most of which may be found in the methodological perspective and beliefs of the teacher about learning and teaching.

Can we claim that a test which exerts a positive impact is a good test? If so, can we argue that the consequences of a test is a part of its development or design process? These and similar questions are the fundamental arguments that may lead us to place washback effect to a right place in testing literature or in educational studies. Applied linguistics has already maintained that acquiring a second language is quite complicated process, and a very difficult one if learned in a pure EFL setting like Turkey. The irony is that this complicated process has no more flexibility

---

<sup>2</sup>TR Acronym: KPDS

and energy to carry yet another burden, which may be created by the wrong and harmful testing procedures. Taking all these very well-known facts into account, the literature of applied linguistics currently presents many studies and discussions in which the consequences of a test are considered as an integral part of the test validity, which was named as *consequential validity* by Messick (1996). Consequential validity has been extensively discussed among language testers in the last decades (Kunnan, 2000; Messick, 1996; Shepard, 1997). Most testing experts now acknowledge that test washback and test impact are highly complex phenomena that should be considered as the fairness and ethical factors in designing tests (eg. Hamp-Lyons, 1997). Nevertheless, even the direction of the discussion, that washback effect is regarded as a part of validity in some or other way, is promising for the future studies.

### ***The Inter-university Foreign Language Examination***

ILE is a national high-stakes academic English test that was developed primarily for individuals holding a PhD degree and aiming to apply for associate professorship examination that is offered by Inter-university Board under the auspices of Council of Higher Education (ÖSYM, 2010). It is offered in three languages, English, French and German. ILE may also be taken by students who want to pursue a graduate degree in national universities and institutions. In practice, test takers are dominantly academic staff of universities who either want to start a Ph.D. program or aim to prepare a dossier for associate professorship. ILE has been offered biannually, namely in spring and in fall terms since 2000. In 2009, the number of test takers in spring term was 72.484 and in fall term 54.278, making a sum of 126.762 testees in a year. The test is conducted in 17 different city centers.

In terms of the content of ILE, the first remark to make is that this exam does not assess all communicative language competences; in other words, the exam is quite a traditional one with some grammar, vocabulary and reading questions, all of which are presented in multiple choice format. 80 multiple choice test items are expected to be answered in 180 minutes, which is also not an effective timing procedure. The exam simply does not offer any section or item for listening, speaking and writing. Obviously, these bring about serious validity problems. Also the test has three different versions that are designed for three major fields of study; namely, natural sciences, medical sciences and social sciences. In other words, a medical doctor needs to take ILE-Medical sciences English test to get a valid test score. Accordingly, a research assistant of a linguistics department should take ILE-social sciences test to get a score for a Ph.D. program of linguistics. This application can be considered as a sort of content-based testing, examples of which are quite rare, and possibly unnecessary, in academic testing literature.

As for the level of the exam, ÖSYM (2010) claims "Bachelor level has been taken as a criterion", which actually does not refer to any kind of academic English levels known nationally or internationally. However, we may say that exam includes multiple choice test items between B1 to C2 in terms of Common European Framework of Reference for Languages (CEFR). What may be crystal clear for an applied linguist is the fact that ILE has not been prepared in terms of the contemporary understanding of modern testing, and that the test may be criticized harshly in terms of its validity. That being the case, it is vitally important to investigate how ILE in its current format affects, and perhaps shapes, English language learning experiences of candidate academics. In this study, the major concern is to find out whether ILE has a positive, negative and neutral impact on testees. Then the findings may enlighten the possible updating process of ILE in the future.

### **3. Method**

#### ***Context***

The incentive behind the present study is to examine the effects of ILE on candidate academics who take the exam for various academic purposes. The study conducted in the capital city of Turkey, Ankara, with a group of academics, research assistants, graduate students of four universities located in Ankara in the academic year 2010.

Academics and graduate in Turkey need to take ILE to prove their knowledge of academic English. Different from the USA and the EU context, in which mostly internationally recognized tests such as TOEFL and IELTS are accepted by the universities, this national academic English test is the major tool for formally proving the level of English. Mostly, testees need to follow an ILE course which is abundantly available in English-schools market mostly located in big cities. There are also numerous internet sites and books available for the testees. While taking the exam is quite cheap (around 20\$ for 2010) in comparison with international high-stakes examinations, the process of getting prepared for this test can be expensive if the testee needs extra help from the institutions and experts.

#### ***Participants***

Fifteen testees were asked to participate in the study. However, twelve testees wanted to contribute to the study as a participant. Their age ranged between 24 and 37 years. All of the participants were either working or studying at different universities in Ankara. Also all of the participants were following a private ILE course. Contact was made with the participants mostly in these ILE courses. Eight of the participants are the candidate academics – three research assistants, four instructors and a lecturer who aspire to get prepared for associate professorship, and the rest four are either master's or Ph.D. candidates preparing for the next step – the comprehension exam or Ph.D. entrance interview – in their academic career. In addition to working or studying in an academic setting, another common ground of twelve different individuals was their limited English language learning experiences in their secondary and tertiary education.

#### ***Research Objectives***

The aims of this research study were to:

- investigate the possible negative, neutral and positive washback effects of ILE candidate academics and graduate students
- identify how ILE shapes foreign language learning experiences of candidate academics and graduate students

#### ***Data Collection and Procedure***

Washback researchers suggest qualitative inquiry to analyze washback effect of a given test (Cheng et al., 2004). Surely it is also important to refer to quantitative data where and when necessary. However, to understand the nature of the phenomenon via in-depth data, qualitative designs are mostly preferred in that the washback effect encompasses so many variables that the picture may be an incomplete one in a purely quantitative research (Cheng et al., 2004). In this study, a qualitative approach to data collection was adopted to answer questions which are mostly based on personal experience, personal feedback and reflection.

This is a qualitative study in which the research objectives were investigated by referring to purely qualitative data. All the data were collected merely from the interviews conducted with the participants. After the participants were informed about the content and objectives of the study, the interviews were conducted with one participant at a time, between 20 minutes and 30 minutes. All interviews were audio taped and transcribed verbatim. The interview sessions were completed in Turkish and in three weeks during September 2010. Then the analyzed data in a relatively categorized form were presented to participants one by one. Each participant was able to see the complete data, and was asked to make comments and reflect on them as much as possible. The procedure was completed after validating all the data with the participants. The final step was to translate the processed and categorized data into English. Three experts, one of whom was a native speaker of English, have worked on the translations and confirmed the final version.

### ***Data Analysis***

The data collected through semi-structured interviews were coded through constant comparative analysis (Glaser & Strauss, 1967, p. 101-116), which refers to a process of repeated sifting through the data to distinguish similarities and patterns of reference in transcripts of the interview. Analyses of these similarities and patterns gradually led to an evolving coding system for the categories. This repetitive process continued until coherence and a saturation of understanding the data was accomplished. Responses to interview questions were also coded by an experienced independent researcher who holds an Ed.D. on educational statistics one week later. The results of both codings indicated 93,6% reliability. The total analysis period took six weeks.

Apart from benefiting from the remarks of the participants on their ILE experience as a direct quotation in the *Results and Discussion* section, the coded and analyzed interview transcripts are also presented in a table format (see Table 2) in the same section to show a greater picture of the phenomenon. The data for Table 1, presented below, is completely based on the thematic coding. The codes and major categories identified through a cyclical data analysis process are presented in Table 1.

**Table 1.** Codes and categories identified in data analyses

Type of Washback Effect	Codes	Relation to Research Objectives (RO)
1. Negative and Unintended	NU	Washback Effect
2. Micro level	MiL	Washback Effect
3. Macro level	MaL	Washback Effect & Learning Experience
4. On Testee	OT	Washback Effect & Learning Experience
5. On Learning	OL	Washback Effect& Learning Experience

## Results and Discussion

This section is organized in terms of the research objectives, which means the following two headings are also the objectives that have led the study. The first objective is to find out the possible washback effects that ILE may exert, and the second one is the way ILE shapes the language learning experiences of the candidate academics.

### *Research objective 1: Washback effects of ILE*

The data that relate to washback effect of ILE is quite a clear one. All the participants have shown the same direction and talked about the same issue for the first question "Can you share your opinions about the effectiveness of ILE exam? For instance, does it have any negative/positive/neutral effect on you?", second question "Does ILE an effective exam that evaluate academic English skills and knowledge?" and for the fourth question " Do you believe that preparing for ILE helps you develop your English? " Some of the responses of the participants are as follows:

*-- "ILE is not a good exam because there is no real English in it. It also affects me rather negatively. All I have to do is to memorize some vocabulary lists and exam tips. This is annoying."*

*--"ILE cannot be an effective exam. I know many people who got high scores in ILE but cannot even write a complete sentence in English. I have completed my MA, and now I cannot even apply for a Ph.D. program because I have to take some ILE course."*

*--"I cannot say it is a modern exam. I used to speak English better, but now I completely forgot how to do it because ILE gets me to study grammar, grammar and grammar."*

*--"I got 48 from ILE last year and if I had got 50, I would be a Ph.D. student now. I lost a complete year just because I did not get 50. This is not fair. The scoring of ILE should be better. "*

*--"I need to get 65 to apply for associate professorship. Well, I am quite experienced about such exams. All I need to do is to memorize more vocabulary items and study more grammar. ILE therefore has a negative impact not only on me but also on many people: my students, my colleagues and even on the system because we spent our precious time here [referring to the ILE course]. I should be writing a research report or deliver a speech in a congress now. "*

Similar results were also indicated in the studies in which the negative influence of high-stakes tests were discussed (Alderson & Hamp-Lyons, 1996; Cheng et al., 2004). Also Yıldırım (2010) conducted a research study in Turkey on a similar high-stakes test, English Component of the Foreign Language University Entrance Exam (ECFLUEE), offered to secondary school graduates who want to enroll a English language program (Linguistics, English/American literature or English language teaching). He found that the exam exerts not only a negative impact but also results in an orientation problem in pre-service ELT programs due to lack of English communicative skills. ECFLUEE is just another version of ILE in that both offer multiple choice questions on syntax, lexis and some sub-skills of reading. The findings of Yıldırım (2010) and present study also indicate that preparing exams like ILE may harm the natural development process of L2 learning and possibly L2 acquisition. Negative washback effect is also associated

with the methodology the teacher adopts (Madaus & Kellaghan, 1992; Nolan, Haladyna, & Haas, 1992; Shepard, 1990). However, none of the participants in this study reported such an incident.

Another important finding also enables a discrimination between the participants. The academic staff that were interviewed in the study were found to be more negatively influenced by the ILE exam. They do believe that this exam is nothing but an unnecessary burden and a great obstacle in front of them. Younger participants who seek to be accepted to a master's program responded less angrily and even positively. However, they did report that ILE simply does not contribute to their communicative skills in English. In other words, younger participants were less influenced due to their high motivation to be accepted to a graduate school. However, older participants who are already teaching in higher education system were significantly discouraged and displayed notably a lower level of motivation. The data concerning the different levels of negative washback effect in relation to student motivation and age are quite a new one for the literature.

### *Research objective 2: ILE and language learning*

The third and fifth interview questions, "How does ILE affect your approach to learning a foreign language?" and "Has ILE influenced your feelings about learning English?" were utilized to collect data regarding language learning experiences of the participants. Of course, it is not possible to think such questions detached from the answers given to the washback-effect questions of the interview. Nevertheless, the insights provided by the participants were quite important, as follows:

*-- "ILE has actually turned language learning to a kind of a memory game. I don't even dream of writing or speaking English. But I can tell you which words are asked in ILE and which ones are not. So my experience of a learning a foreign language is mostly challenging, discouraging, being perfectly honest, a little bit boring."*

*--"I don't like learning a foreign language. Actually I find nothing interesting, attracting and fun in learning a language. Also I don't have years to learn it, otherwise I may be dismissed from my school and I may not be able to receive any salary. So I don't feel good things about learning English for now."*

*--"ILE affected my view and understanding of learning a foreign language badly. Therefore, I feel bad about learning a language."*

*--"Well, I feel disappointed. I might develop my English more effectively if ILE asked me to write a paragraph or utter something in English. In this format, the exam is nothing but a procedure for all of us. I regret it as we spent money and time, and so does the State! Generally, I like English but no way, not this way!"*

*--"I actually enjoy learning English. Sometimes ILE can be fun. But what am I going to do with ILE English? I can read abstracts, for example, or translate some passages. That is all and surely very limited."*

The findings on language learning experiences of candidate academics have shown that, during the preparation for ILE, the participants utilized mostly cognitive strategies which had led them to memorization and short term study habits. Various similar studies have previously focused on

the impact of a specific test on learning strategies or study skills of testees (Cheng et al., 2004; Green, 2007; Song, 2009). However, the findings of this study also maintain that a poor examination may also affect the motivation of the testees, which may also lead to a test anxiety or to avoidance behavior towards subject matter. These findings may be interpreted in terms of the theoretical discussion of Watanabe (2004) in which washback effect is associated with certain motivational theories. The types of the negative washback effect that ILE exerts on testees are illustrated below in Table 2.

**Table 2.** Thematic categorization of interview results

Type	Washback Effect	Learning Experiences
Micro level Test washback	<ul style="list-style-type: none"> <li>• Negative (unintended) effect on                             <ul style="list-style-type: none"> <li>○ L2 competences</li> <li>○ Academic career</li> <li>○ Motivation</li> <li>○ Cognitive learning</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Short term study habits</li> <li>• Low order thinking skills</li> <li>• Memorization and rote learning</li> </ul>
Macro level Test impact	<ul style="list-style-type: none"> <li>• Negative effect in terms of                             <ul style="list-style-type: none"> <li>○ Course and material expenses</li> <li>○ Time dedicated for ILE</li> <li>○ Teachers' anxiety and role</li> <li>○ A poor national exam</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Anxiety and inhibition</li> <li>• Feeling of failure</li> <li>• Avoidance behavior towards learning English</li> </ul>

In the light of the findings presented in Table 2, ILE was found to have a negative and unintended washback effect on testees and possible institutions with which these group of candidate academics interact for educative and professional purposes. The type of the effects can be addressed as 'Micro level' or the test washback (Wall, 1997) and 'Macro level' or the test impact (Wall, 1997). Micro level effects were categorized in terms of the effects of ILE on the individual's learning and life. In this respect, these effects were L2 competences that remain weak and undeveloped, the academic career that has to be somehow suspended for allocating time and energy for the exam preparation, a lower level of motivation for learning L2 and finally over-emphasis of cognitive learning due to the content of the exam. Learning a foreign language requires not only cognitive learning but also affective and social learning styles and strategies (Brown & Hudson, 1998), but getting prepared for ILE, the findings illustrate, requires little or no affective and social learning processes and contexts. The macro level effects, which refer to the social level effect of the exam, were the time and the money spent for the exam as well as the teachers' burden that was mentioned several times by the participants. Some participants believed that ILE was a heavy burden on their teachers' shoulder in that the teachers were considered unsuccessful if the students cannot, somehow, succeed in the exam. Many studies in the literature verify this critical point (Anderson, Muir, Bateson, Blackmore & Rogers, 1990; Noble & Smith, 1994; Watanabe, 2004). The final inference from the macro level effects would be that there was a general accord with the idea that ILE was a poor test, which should be updated and replaced with a more effective national test.

With regard to the learning experiences, the findings can also be discussed under micro and macro level categorizations. As for micro level learning experiences, the participants were found to display short-term study habits and low order thinking skills which were reported as the

memorization of the exam tips and the strategies that encourage a behaviorist or mechanical approach to foreign language learning, and finally as memorization and rote learning. This finding is supported by the study conducted by Anderson (Anderson et al., 1990), showing quite similar results. These micro level learning experiences were strongly related to the macro level ones, perhaps in a causal relationship. Participants reported to suffer from anxiety and inhibition, feeling of failure, both of which led them to an avoidance behavior towards learning English, which supports the previous literature such as Cheng et al. (2004) and Shohamy (2001).

### **Conclusion and Implications**

The number of studies on washback effect in language testing literature has been growing rapidly due to its critical impact on learning, learners and teachers and even the society or the complete educational system of a country. The high-stakes tests across the countries have also led the scholars focus on the consequences of such tests on students, teachers and interfering institutions (Anderson & Wall, 1993; Bachman & Palmer, 1996; Biggs, 1995; Cheng, 2005; Messick, 1996; Shohamy, Donitsa-Schmidt & Ferman, 1996; Taylor, 2005; Yildirim, 2010).

In the present study, the results revealed that ILE exerts a negative impact on candidate academics. The results also show the exam does not only have a micro level effect, which can be defined as the effect on an individual or a small group of individuals, but also a macro effect on a relatively populated group of individuals studying or working at university contexts. Participants also reported inhibition and avoidance problems, which can be regarded as a very critical problems for academics whose primary language is mostly English in both studying, publishing and presenting. In this study, the reason why ILE exerts the harmful and negative washback effects on the candidate academics is attributed to the content and style of the test itself, which is a multiple-choice test claiming to assess grammar, vocabulary and reading-skills knowledge in the L2 – English in the context and scope of this study. The unanimous reactions of all participating candidate academics of this study were that they need to develop more than what ILE assesses, and that ILE represents an obstacle in front of them. These findings are compatible with the related literature (Burrows, 2004; Song, 2006; Wall, 1999; Watanabe, 2000, 2001; Yildirim, 2010).

The study has many implications that may be put into practice in a possible modification of ILE and similar tests in Turkey. First and foremost, it is important to note that the findings of this study may be generalized with caution, as it presents a qualitative study on 12 participants. However, the findings may account for FLEC, which is quite a similar test to ILE in terms of content and testing style. Much older than ILE with a greater population of testees and offered in 24 languages, FLEC may be discussed under same implications necessary for ILE. In the light of the findings of the study, it is clear that ILE, FLEC and FLE exams offered to Turkish students and professionals should be modified immediately to assess four language skills in academic contexts. Fine examples are abundant in different countries (eg. Eurydice, 2009).

Another possible amendment in these tests will also enable Turkish education system to comply with the EU norms and standards. EU is currently trying hard to establish a well-functioning system of *European Higher Education Arena* (EHEA) among member and candidate countries (Eurydice, 2009). Relevant reports of the EU show that Turkey is among "the countries which have low rates of both incoming and outgoing mobility" (Eurydice, 2009, p. 52). In order for Turkey to conduct student mobility at all study cycles of EHEA and to comply with CEFR

standards, for both of which Turkey has already spent a lot of effort and fund, these exams must be updated immediately in parallel with the effective examples such as IETLS and TOEFL. Therefore, in addition to national needs and expectations, being part of a Bologna Process also necessitates such modifications (Eurydice, 2008).

Such implication suggestions naturally bring about the concerns of how to carry out probable modifications on ILE, FLEC or FLE. As a matter of fact, the response to such questions is not that complicated. The literature of applied linguistics is full of such examples (Cheng et al., 2004). However, the first step to take should be designing a test that is able to collect evidences of communicative competences of the testees in four skills. To this end, this great reform movement should be conducted with caution by carrying out well-monitored pilot studies on the progressing versions of the test. The result of such a pursuit will be beneficial not only for the candidate academics, students and professionals, but for all the country.

## References

- Anderson, J. O., Muir, W., Bateson, D. J., Blackmore, D., & Rogers, W. T. (1990). *The impact of provincial examinations on education in British Columbia: General report*. Victoria: British Columbia Ministry of Education.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115–129.
- Alderson, J.C. & Hamp (Lyons, L. (1996). TOEFL preparation courses, a study of washback. *Language Testing*, 13(3), 280-297.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, England: Oxford University Press.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, 10, 12-42.
- Biggs, J. B. (1995). Assumptions underlying new approaches to educational assessment. *Curriculum Forum*, 4(2), 1–22.
- Brown, J.D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Burrows, C. (2004). Washback in classroom-based assessment: A study of the washback in Australian adult migrant English program. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp.113-128). Manwah, New Jersey: Lawrance Erlbaum Associates.
- Cheng, L. (2004). The washback effect of a public examination change 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: Lawrance Erlbaum Associates.
- Cheng, L. (2005). *Changing language teaching through language testing: a washback study*. New York: Cambridge University Press.

- Davies, A. (1985). Follow my leader: Is that what language tests do? In Y. P. Lee, C. Y. Y. Fok, R. Lord, & G. Low (Eds.), *New directions in language testing* (pp. 1–12). Oxford: Pergamon Press.
- Eurydice (2008). *Key Data on Teaching Languages at School in Europe*. Brussels: Education, Audiovisual and Culture Executive Agency (EACEA P9 Eurydice)
- Eurydice (2009). *Higher Education in Europe 2009: Developments in the Bologna Process*. Brussels: Education, Audiovisual and Culture Executive Agency (EACEA P9 Eurydice)
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. New York: Aldine.
- Green, A. (2007). Washback to learning outcomes: a comparative study of IELTS preparation and university preessional language courses. *Assessment in Education*, 14(1), 75-97.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14, 295–303.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge, England: Cambridge University Press.
- Hughes, A. (2003). *Testing for language teachers (2<sup>nd</sup> edition)*. Cambridge: Cambridge University Press
- Kunnan, A. J. (ed.). 2000. Fairness and validation in language assessment: *Selected papers from the 19<sup>th</sup> Language Testing Research Colloquium*, Orlando, Florida. Studies in Language Testing, Vol. 9. Cambridge: UCLES/Cambridge University Press.
- Madaus, G. F., & Kellaghan, T. (1992). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 119–154). New York: Macmillan.
- McNamara, T. (1996). *Measuring second language performance*. London: Longman.
- Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13, 241–256.
- Noble, A. J., & Smith, M. L. (1994a). *Measurement-driven reform: Research on policy, practice, repercussion* (Tech. Rep. 381). Tempe, AZ: Arizona State University, Center for the Study of Evaluation.
- Nolen, S. B., Haladyna, T. M., & Haas, N. S. (1992). Uses and abuses of achievement test scores. *Educational Measurement: Issues and Practice*, 11(2), 9–15.
- Shepard, L. A. (1990). Inflated test score gains: is the problem old norms or teaching the test? *Educational Measurement: Issues and Practice*, 9, 15-22.
- Shepard, L.A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16(2), 5-13.
- Shohamy, E. (2000). Using language tests for upgrading knowledge. *Hong Kong Journal of Applied Linguistics*, 5(1), 1–18.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Harlow: Pearson Education.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test Impact revisited: Washback effect over time. *Language Testing*, 13, 298–317.

Novitas-ROYAL (Research on Youth and Language), 2011, 5(2), 215-228.

Song, D.Y. (2009). Applicability of an EBB Scale to English Speaking Performance Assessment in the Korean Secondary School. *English Language Teaching*, 21(1), 111-135.

T.C. Yüksek Öğretim Kurulu Öğrenci Seçme ve Yerleştirme Merkezi. (2010). Retrieved on December 08, 2010, from <http://www.osym.gov.tr>

Taylor, L. (2005). *ELT Journal*. 59/2. Oxford University Press. Retrieved on September 4, 2010 from <http://eltj.oxfordjournals.org>

Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (pp. 291–302). Dordrecht: Kluwer Academic.

Wall, D. (1999). *The impact of high-stakes examinations on classroom teaching: A case study using insights from testing and innovation theory*. Unpublished doctoral dissertation, Lancaster University, UK.

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. (2001). Does the university entrance examination motivate learners? A case study of learner interviews. In Akita Association of English Studies (Eds.), *Trans-equator exchanges: A collection of academic papers in honor of Professor David Ingram* (pp. 100–110). Akita, Japan: Author.

Watanabe, Y. (2004). 'Teacher factors mediating washback' in Cheng L, Y Watanabe & A Curtis (Eds.). *Washback in Language Testing: Research Contexts and Methods*. (pp. 129-146).NJ: Lawrence Erlbaum.

Yıldırım, Ö. (2010). Washback Effects of a High-Stakes University Entrance Exam: Effects of the English Section of the University Entrance Exam on Future English Language Teachers in Turkey. *Asian EFL Journal*. 12(2), 92-116. Retrieved from <http://www.asian-efl-journal.com> on December 21, 2010.

## **Appendix**

### **Interview Questions**

1. Can you share your opinions about the effectiveness of ILE exam? For instance does it have any negative/positive/neutral effect on you?
2. Does ILE an effective exam that evaluate academic English skills and knowledge?
3. How does ILE affect your approach to learning English?
4. Do you believe that preparing for ILE helps you develop your English?
5. Has ILE influenced your feelings about learning English?
6. Do you spent a lot of time and money for preparation?
7. What about the teachers? What is their role in this process?
8. Can you talk about the strategies you employ when studying English for ILE?