



# دانشگاه حکیم سبزواری

دانشکده ادبیات و علوم انسانی

پایان نامه جهت دریافت درجه کارشناسی ارشد در رشته آموزش زبان انگلیسی

پرسی تاثیر آزمون ورودی دانشگاه بر شیوه ی تدریس معلمان زبان انگلیسی دوره ی دبیرستان در

ایران : تئوری زمینه ای

استاد راهنما :

دکتر سید علی استوار نامقی

استاد مشاور:

دکتر سعید غنی آبادی

نگارش :

فریده صمدی

خرداد ۹۱

**CHAPTER ONE**  
**INTRODUCTION**

## 1.1 Background

Traditionally, language testing researchers have focused their attention on inherent issues in tests. However, recently some researchers have turned to the empirical investigation of the washback phenomenon which is not limited to the test itself. It depends on other teacher and contextual factors which may be different from context to context (Cheng, Watanabe, & Curtis, 2004).

According to Ostovar Namaghi (2006), three forces control and steer teachers' work in the Iranian educational context. First, since teachers cannot choose a textbook which is in line with their students' needs, the input is controlled by the prescribed curriculum. Second, the output is controlled by the mandated national testing scheme so that teachers cannot develop tests which have a positive backwash on teaching and learning. Third, since high score is culturally equal to higher achievement, the process of teaching and learning is controlled by the grade pressure from students, parents and school principals. He argues that teachers are pure implementer of the prescribed initiatives and schemes surrounded by cultural constraints which prevent them from using their own professional knowledge and experience.

According to Ghorbani (2008), the centralized control of curriculum and assessment in Iran is assumed to have led to teaching towards the high-stakes UEE which affects the future career and lives of pre-university and senior high school students. Iranian senior high school English teachers, in general, and Pre-university English Teachers (PETs), in particular, feel they are expected to prepare their students for university entrance exams by having them translate English texts into Persian. They explain and put an emphasis on the grammatical structures explicitly. They also try to improve their students' reading skills at the expense of listening,

speaking and even writing skills. The dominant methodology is grammar-translation with a teacher-centered approach in which language usage and not actual language use is emphasized.

Many students at pre-university and academic senior high schools assume that the purpose of English teaching and learning is preparing students for the UEE. Students usually influence PETs' instruction through their expectations that they should prepare them for the UEE. Since the UEE influences students' future career and lives, PETs often teach to the test and students focus only on those activities and skills that are likely to appear on the test.

### **1.1.1 English Education in Iran**

In Iran, English is taught as a foreign language and is practiced within a context-restricted environment, in which the textbook and classroom teacher play the main role. Previously, English education in Iran formally started from second grade in junior high schools, but now it begins from the first grade. All schools at different levels follow the curriculum standards. The Ministry of Education (ME) compiles, develops and publishes textbooks and teaching materials for nationwide public and private senior high schools.

### **1.1.2 Iranian Nationwide University Entrance Examination**

According to Cheng (2005) examinations have been used as a means of control and as a way to counter nepotism and favoritism in the allocation of scarce opportunities. They have also been used to encourage the development of talent, upgrade the school performance, and select for education and employment for many years.

The UEE is a high-stakes test in which many students become disappointed. Students' failure to enter higher education institutes has led to undesirable outcomes such as dropping out of studies or brain drain and thus loss of a great portion of the county's potential.

As pointed out by Chapman and Snyder (2000, p. 458), high-stakes tests will have a greater influence on teaching and learning if their “primary use is to ration future opportunity as the basis for determining admission to the next layer of education or to employment opportunities”. Although the number of the country's higher education institutes has been increased and the capacity of many universities has been expanded, the fierce and tough competition among pre-university and high-school graduates is still a major concern. Less than one third of the students who sit for the UEE manage to enter universities.

Pre-university and senior high school graduates, who wish to enter the country's tuition-free public universities, participate in an annual and intense multiple-choice exam which usually lasts about 4.5 hours. The Education Evaluation Organization (EEO) which operates under the supervision of the Ministry of Science, Research, and Technology takes care of all aspects of the UEE.

One of the main criticisms regarding the UEE is that it is a very limited way of evaluating the caliber of pre-university and high-school graduates who wish to gain a place at university. In such a situation, doing anything out of interest which is unlikely to be tested in the exam is assumed to be a waste of time.

## **1.2 Context of the Study**

The study took place in Kashmar, a city located in Khorasan Razavi province. All participants were selected from urban areas. The study started with an open-ended interview with seven experienced senior high school teachers who were willing to share their perceptions and attitudes on the washback effect of the UEE toward their teaching activities with the researcher.

### 1.3 Statement of the Problem

The investigation of the washback effect of tests at the national level is the common point among lots of research studies (see, for example, Dore, 1976; Madaus, 1988; Buck, 1988; Wall and Alderson, 1993; Cheng, 1997, 2003, 2005; Cheng, Watanabe and Curtis, 2004; Alderson and Wall, 1996; Bailey, 1996; Wall, 1997, 2000; Chapman and Snyder, 2000; Watanabe, 2004). The importance of the UEE cannot be denied because it is a criterion for admission into higher education and it supports the processes of self-evaluation and improvement of educational institutions. There has been intensive work by the Ministry of Education, Ministry of Science, Research and Technology, Education Evaluation Organization (the organization in charge of administering of the UEE annually), as well as non-governmental organizations in large and smaller towns to inform the senior high schools, pre-university centers, teachers, students, parents, publishers, and the community in general about the importance of this test administration and its implications. Some Iranian students frequently complain that their English listening, speaking, and writing skills are poor. However they seldom complain about knowing grammatical points. Why? Generally, most English language teachers and researchers consider that the four skills, i.e., listening, speaking, reading and writing should be equally developed for a senior high school and pre-university student, but listening and speaking skills are not covered in the UEE at all, and writing subset only covers a small section in the UEE. Why do not listening, speaking and writing skills have any role in the UEE if we say that all of the skills are equally important to senior high school and pre-university students?

On the other hand, many Iranian language researchers and educators have assumed various levels of washback effect on English instruction. Despite numerous studies regarding the testing impact on English as a foreign language teaching and learning being of different contexts,

empirical research is still lacking on the Iranian senior high school English teachers' perceptions and students' attitudes towards the washback effect of the UEE on English learning and teaching. Thus, this study is designed to investigate the nature and scope of the UEE washback on English teaching in Iranian senior high schools based upon teachers' perceptions.

#### **1.4 Significance of the Study**

The purpose of this study is to investigate how English teachers in Kashmar senior high schools perceive the effect of the UEE on their teaching activities. The teachers' perceptions of the UEE are investigated with an aim to explain how their goals and actions are influenced by the effect of the UEE. More importantly, the findings should provide important information to lead the involved educational parties in Iran English education to an improvement of the UEE tests. This issue is worth scrutiny because more than 1.5 million Iranian students take this exam annually and it is considered as the sole criterion for admission into state universities.

#### **1.5 Objectives of the Study**

This study intends to use a qualitative research method, particularly GTM, to explain and predict various aspects of the UEE washback effect on teachers' activities of English teaching. Specifically, the objectives of this study are:

1. To examine the phenomenon of washback effect in the light of interpretive instruction.
2. To assess how the main participants of the study within Iran educational context react to the UEE as the most important nationwide exam.
3. To identify whether the UEE influences teaching activities performed in Iranian senior high school EFL classes, and whether such an impact brings about positive or negative consequences.

## 1.6 Research Question

In order to facilitate the investigation regarding the washback effect of the UEE as a large-scale exam on the Iranian senior high school English teachers' attitudes and perceptions of English teaching, the researcher formulated the following research question:

How does the UEE shape the teaching activities?

## 1.7 Definition of Terms in this study

**Washback:** Washback is described as “the extent to which the introduction and the use of a test influence language and teachers to do things they would not otherwise do that promote or inhibit language learning” (Messick, 1996, p. 4)

**High Stakes Tests:** high stakes tests—are used more widely at regional, national, and international level, and as the consequences of test use—especially the valid and ethical use of test results—come under greater scrutiny in the public domain.

**Grounded theory:** Grounded theory, “the discovery of theory from data” (Glaser & Strauss, 1967, p. 1), provides the opportunity for the researcher to theorize from evidence existing in the data.

**Open coding:** Open coding occurs at the beginning of a study. The primary goals of open coding are to conceptualize and categorize data, achieved through two basic analytic procedures: making comparisons and asking questions of the data.

**Axial Coding:** The second stage of data analysis is axial coding. Strauss and Corbin (1998, p. 123) described axial coding as the process of relating categories to their subcategories . . . linking a category at the level of properties and dimensions”.

**Selective Coding:** The final stage of data analysis in grounded theory is selective coding, which builds upon the foundation of the previous open and axial coding efforts.



Selective coding is “the process of selecting the central or core category, systematically relating it to other categories, validating those relationships, and filling in categories that need further refinement and development” (Strauss & Corbin, 1990, p. 116).

**Theoretical sampling:** Another aspect of the data analysis is theoretical sampling of concepts that are relevant to the emerging theory. Theoretical sampling is cumulative, increases the depth of focus, notes variation, and occurs in all three phases related to coding. Theoretical sampling is terminated once theoretical saturation is reached.

**Theoretical saturation:** is achieved when (a) no new data emerges regarding a category, (b) the category is dense enough to cover variations and process, and (c) relationships between categories are delineated satisfactorily as well aids for Data Analysis.

## **1.8 Limitations**

This study concentrated on investigating how the senior high school teachers in Kashmar perceived the impact of the UEE on their teaching activities. Since the population which involved in the investigation was confined to English teachers in senior high schools in Kashmar, this study had no attempt to investigate washback impact experimentally caused by a different type of examination or in a different context. In addition, this study focused on the explanation on how the UEE influenced English teaching in Kashmar senior high schools. Data collected in this study were only adequate for describing perceptions of washback impact of the UEE on Kashmar senior high school English teaching and how it could be explained and predicted by selected teachers and school characteristics. Thus, the results would be inappropriate to be generalized to other contexts or other examinations. Moreover, the findings in the discussion were based on teachers' opinions; further empirical data (e.g., classroom observations, questionnaires), especially from longitudinal studies, should be eventually collected and

analyzed to add up insight into the nature of this phenomenon. In this research, due to the lack of time, the participants were limited to seven teachers, which is a small sample.

## **CHAPTER TWO**

### **REVIEW OF THE RELATED LITERATURE**

## 2.1 Review of the Related Literature

Before the 1990s, washback was researched mostly in the field of general education and accounted for through informal observations and generalities (Alderson & Wall, 1993; Cheng et al., 2004; Wall, 2000). Educators discussed the negative effects of tests, whether intended or accidental, on teachers' methodology, the curriculum and students' learning, problematizing generally how to develop valid language tests which would have a positive influence (e.g. Davies, 1985). Popham (1987), a proponent of this 'measurement-driven instruction,' argued that if tests reflected 'useful' skills, then preparing for tests would be a beneficial activity for teachers and students. However, this idea was opposed by Madaus (1988) who insisted that, regardless of their validity and communicative constructs, tests inevitably inhibit students' learning and promote teachers' narrowing the curriculum, due to tests constructs' power to determine these activities (e.g. Cheng, 1998). Still, research demonstrating a link between valid tests and positive impact on teaching and learning merely consists of general experiences of stakeholders, with no systematic investigation of what is actually occurring in classrooms because of tests.

Since the 1990s, researchers began empirically accounting for language tests' 'washback' or 'impact.' Washback from high-stakes tests was given special attention since research confirmed these tests' deleterious effects on stakeholders (Shohamy, Donitsa-Schmidt & Ferman, 1996). Foundational to the emergence of washback research on high-stakes tests were Alderson and Wall's (1993) washback hypotheses which outlined the types of and conditions under which washback could occur. Additional models for how washback worked were developed, such as Hughes' (1994) trichotomy – washback to the participants (teachers, learners, test/materials developers), processes (course curriculum, classroom materials and teachers' methodology) and products (the quality of learners' learning). Furthermore, Bailey (1996)

combined Hughes's ideas to produce the notions of 'washback to the learners' – "the test-derived information provided to the test-takers and having a direct impact on them" – and 'washback to the program' – "the results of test-derived information provided to teachers, administrators, curriculum developers, counselors, etc." (pp. 263-264). Using these frameworks, subsequent washback research has explored the "nature of washback...and the conditions under which it operates" for teaching and learning in high-stakes, test-preparation classrooms (Alderson & Wall, 1993, p. 116).

Much research on the extent and types of washback from high-stakes tests have considered washback for the participants and processes concurrently because of the interdependence of teachers (participants) using certain preparation processes (changing course content; using certain materials) due to washback effects (Bailey, 1999; Spratt, 2005). These studies have identified language teachers as playing pivotal roles in effecting washback because of their pedagogical and ethical decisions about class content and methodology for high-stakes test-preparation courses as opposed to other courses (e.g. Read & Hayes, 2004; Watanabe, 1996). Teachers' pervasiveness as subjects of washback research arose because of their control in classrooms, coupled with the facts that they are easier to communicate with and have less imminent stress on their lives as a result of high-stakes tests (Shohamy et al., 1996; Spratt, 2005). Nevertheless, other washback causes were discovered alongside teachers' methods – for example, parents, textbooks, the class's focus, financial stability, the perceived 'highness' of a test's stakes and the timing of the research with the test's imminence (Alderson & Wall, 1993; Read & Hayes, 2004; Shohamy et al., 1996; Watanabe, 1996).

Not only the processes and participants, but also the products of washback have been investigated, especially in recent washback research concerned with ethical language testing.

Results distinguished the importance of valid criteria and score reporting, materials chosen for test preparation, specifications of the real-world tasks evaluated, as well as the diversity of high-stakes tests' impact on individual students and teachers as influencing these tests' washback on the products (Qi, 2005; Saif, 2006; Wall & Horák, 2006, 2008a).

Parallel to assessing washback's products; another frequently studied washback cause has been the intended effects of newly-developed or changed tests. The idea that designing or modifying a test to be authentic and direct will create positive washback has already been mentioned as popular among language testing experts (e.g. Popham, 1993). Over time, though, research has shown that the link between washback validity – the relationship between the test and teaching/learning practice in classrooms (Alderson & Wall, 1993) – and test innovation is less automatic than supposed and too simplistic an account of washback's instances (e.g. Alderson & Hamp-Lyons, 1996; Wall, 2000). Messick (1996) even argued against a causal relationship between washback and validity, because much of what happens in test-preparation classrooms is not caused by washback from the test but by other factors, outside of responsible test design. Any causal relationship between 'valid' test innovation and washback, then, must rest on a thorough, longitudinal investigation of the context, participants and intentions of the test developers (Bailey, 1999).

What researchers found about washback effects from innovations and new development for high-stakes tests is that these modifications did not always have their intended effects (e.g. Cheng, 1998). In fact, changes were predominantly "indirect," "unpredictable" and often mediated by other factors such as the teachers' education, individual learning styles of students, educational context, and the "selecting function" of high-stakes tests (Andrews et al., 2002, p. 221; Bailey, 1999; Cheng 1997; Qi, 2005, p. 163). Another common finding was that teachers

were crucial to effecting intended changes for positive washback in test-preparation classrooms, since they subjectively modified the content of their test-preparation classes, but rarely their methodology (Cheng, 1997, Wall, 2000). In other words, innovations for positive washback in high-stakes test-preparation classrooms could influence *what* teachers taught and learners learned but not *how* they taught and learned (Cheng, 1998; Messick, 1996).

Hence, while washback studies of high-stakes tests have identified various washback causes, the consensus is that washback is “malleable” and necessitates further investigation of its complexities (Spratt, 2005, p. 23). Despite this, students’ “point[s] of view on their washback-related behaviour before and after tests” (Bailey, 1999, p. 14) have barely been incorporated into recent examinations of washback (Wall, 2000).

## **2.2 Definitions of Washback**

The notion of “washback” is prevalent in language teaching and testing literature, but it is seldom found in dictionaries. Some writers use the term “washback” while others prefer “backwash” to describe the effects or influences brought by tests or examinations. Below, the definitions by various researchers are arranged under the groupings of (a) backwash or (b) washback.

### **a) Backwash**

- Hughes (1989): “The effect of testing on teaching and learning” is known as backwash.
- Spolsky (1994): The concept of backwash deals with the unforeseen side-effects of testing and not to the intended effects when the primary goal of the examination is the control of curricula.
- Biggs (1995): Backwash refers to the fact that testing controls not only the curriculum but also teaching methods and students’ learning strategies.

## **b) Washback**

- Alderson & Wall (1993): Washback compels “teachers and learners to do things they would not necessarily otherwise do because of the test”.

- Messick (1996): Washback is described as “the extent to which the introduction and the use of a test influence language and teachers to do things they would not otherwise do that promote or inhibit language learning”.

- Bailey (1996): Washback is the “influence of testing on teaching and learning.”

- Shohamy, et al. (1996): Washback is delineated as “the connections between testing and learning”.

- Cheng (2005): Washback indicates “an intended or unintended (accidental) direction and function of curriculum change on aspects of teaching and learning by means of a change of public examinations”.

## **2.3 Concepts similar to Washback**

In addition to “backwash” and “washback”, researchers used other similar terms stated below to investigate the phenomena of the influences or effects of tests on the educational field.

**a) Test impact:** (Andrews, 2004; McNamara, 2000; Wall, 1997; Bachman & Palmer, 1996) some researchers have argued that tests can have more far-reaching effects in the educational world than just in the language classroom. Bachman & Palmer (1996) used the term “test impact” to refer to the effects that tests have on individuals (teachers and students) or educational systems and on the society at large. Wall (Wall, 1997) held a similar view by stating that “Test impact refers to any of the effects that a test may have on individuals, policies or practices within the classroom, the school, the educational system, and society as a whole”. McNamara (2004) claimed that “Tests can also have effects beyond the classroom. The wider



effect of tests on the community as a whole, including the school, is referred to as test impact”. Andrews (2004) used “test impact” to describe “the effects of tests on teaching and learning, the educational system, and the various stake holders in the education process”.

**b) Systemic validity:** Systemic validity refers to the effects of instructional changes brought about by the introduction of the test into an educational system as stated tests induce “in the education system curricular and instructional changes that foster cognitive skills that the test is designed to measure”. (Fredericksen & Collins, 1989)

**c) Consequential validity:** Consequential validity encompasses concepts ranging from the uses of tests, the impacts of testing on test takers and teachers, the examination of results by decision makers, and the potential misuse, abuse, and unintended usage of tests. In other words, consequential validity implies that tests have various influences both within and beyond the classroom. In other words, consequential validity refers to the societal implications of testing that are only one facet of a broader, unified concept of test validity. (Messick, 1989, 1996)

**d) Curriculum-alignment:** (Shohamy et al, 1996) defined curriculum alignment as “the curriculum is modified according to test results”.

## **2.4 Definition of Washback in this Study**

After reviewing definitions of washback, the term can be defined according to two major perspectives: one at a narrower view within the classroom at a micro level, and the other at a wider and more holistic view beyond the classroom at a macro level. As suggested by Bachman & Palmer (1996), washback, at a macro level, refers to the extent to which a test influences within the society, ranging from government policymaking, school administration, publishing, and general opportunities, to parents’ expectations of their children. At a micro level, washback

refers to the extent to which a test influences within the classroom, mainly in the change or innovation of curricula and teachers' methodologies and the influence of students' learning. Bailey (1996, p. 5) used the phrase "washback to the learners" to indicate the effects of test on students, and "washback to the programme" to indicate effects of test on teachers, administrators, curriculum developers, counselors, etc.

To summarize, the narrower definition of washback focuses on the effects that a test has on teaching and learning. The wider or more holistic view of washback (also defined as test impact) looks beyond the classroom to the educational systems and society at large. All in all, tests can have "significant impact not only on individuals but also on practices and policies—in the classroom, the school, the educational system and in society as a whole" (Wall, 2005).

In this study, washback at a micro level will be adopted to investigate the test washback in the classroom, that is, the washback effect on teaching.

## **2.5 Positive and Negative Washback**

Generally, washback can be analyzed according to two major types: positive and negative, depending on whether it has a beneficial or harmful impact on educational practices (Hughes, 1989). This section explores positive and negative washback in terms of both the classroom setting and the educational/political system.

### **2.5.1 Positive Washback**

#### ***2.5.1.1 Classroom Setting***

Teachers and learners will be motivated to fulfill their teaching and learning goals (Anderson & Wall, 1993). Good tests can be utilized and designed as beneficial teaching-

learning activities so as to encourage a positive teaching-learning process (Pearson, 1988, p. 107).

A creative and innovative test can quite advantageously result in a syllabus alteration or a new syllabus (Davis, 1985).

### ***2.5.1.2 Educational/Societal System***

Decision makers use the authority power of high-stakes testing to achieve the goals of teaching and learning, such as the introduction of new textbooks and new curricula (Shohamy, 1992; Wall & Alderson 1993; Cheng; 2005).

Tests are encouraged to promote the idea of lifelong learning and encourage people to learn English (Language Testing and Training Centre, 2008).

## **2.5.2 Negative Washback**

### ***2.5.2.1 Classroom Setting***

The test will lead to the narrowing of content in the curriculum. What students have learned is test language, instead of total phases of understanding (Shohamy, 1992).

The tests may well fail to create a correspondence between the learning principles and/or the course objectives to which they should be related (Cheng, 2005). Many teachers detailed high anxiety, fear and pressure to cover the material, as they felt that their job performance was assessed by students' test scores (Shohamy, 1996). Educators experienced negative reactions to the stress brought about by public displays of classroom scores. Inexperienced teachers felt a greater degree of anxiety and pressure for accountability than did teachers with more experience.

“Testing programs substantially reduce the time available for instruction, narrow curricular offerings and modes of instruction, and potentially reduce the capacities of teachers to

teach content and to use methods and materials that are incompatible with standardized testing formats” (Smith, 1991a).

An increasing number of paid coaching classes are set up to prepare students for exams, but what students learn are test-taking skills rather than language learning activities (Wiseman, 1961).

Measurement-driven instruction will definitely result in cramming, narrowing the curriculum, focus of attention on those skills that are most relevant to testing, placement of constraints on teachers’ and students’ creativity and spontaneity, and disparage the professional judgment of educators (Madaus, 1988).

#### ***2.5.2.2 Educational/Societal System***

Decision makers overwhelmingly use tests to promote their political agendas and to seize influence and control of educational systems (Shohamy, 1996). Tests are used as a “lever” for change.

To summarize, in terms of the classroom setting at a micro level, the positive washback integrates meaningful and innovative learning activities in teachers’ educational methodologies, and thus educators will devote more attention to students’ intentions, interests, and choices.

Students at the same time will be encouraged and motivated to work harder. On the other hand, the negative washback is that teachers will usually teach to the test, narrow the curriculum and only focus on what will be tested. Moreover, cramming will be the washback brought by measurement-driven tests, even though there is an ongoing debate as to whether cramming is positive or negative washback. In terms of educational setting, the positive washback is that the authority can use the test to attain its goal of teaching and learning. However, the negative washback is that the authority uses that goal to control and obtain the power of the academic