

«نشریه علمی-پژوهشی آموزش و ارزشیابی»

سال نهم - شماره ۳۳ - بهار ۱۳۹۵

ص.ص. ۱۵۹-۱۸۰

## Washback Effects of the New Iranian TEFL Ph.D. Program Entrance Exam on EFL Instructors' Teaching Methodology, Class Assessment, and Syllabus Design: A Qualitative Scrutiny

Reza Rezvani<sup>1\*</sup>  
Ali Sayyadi<sup>2</sup>

---

تاریخ دریافت مقاله: ۱۳۹۴/۰۹/۱۸

تاریخ پذیرش نهایی مقاله: ۱۳۹۴/۱۲/۲۵

---

### Abstract

Given the significance and sensitivity of the Iranian Ph.D. program entrance exams, such high-stakes exams are assumed to exert considerable influence on Iranian public higher educational curriculums. The present study was motivated to scrutinize the EFL instructors' insights to explore the potential washback effects of the new Iranian TEFL Ph.D. program entrance exam on their teaching methodology, class assessment, and syllabus design. To this end, the researchers conducted in-depth interviews with ten experienced instructors teaching M.A. courses at four Iranian universities. The results emerging from the close qualitative content analysis of the respondents' introspection indicated that the new exam, despite its considerable significance, fails to generate substantial effects on the instructors' teaching methodology and class assessment due to the inconsistency between the objectives supposed to be met by the new exam administration and those expected to be achieved in TEFL M.A. and Ph.D. programs. Further, it was revealed that the new exam, on the contrary, substantially influences the instructors' syllabus design through conditioning their material and content selection. This study further discusses the implications of the findings in the context of Iranian TEFL higher education.

**Keywords:** washback, university entrance exams, TEFL postgraduate programs, teaching methodology, class assessment, syllabus design

---

1- Assistant Professor of TEFL, Yasouj University. rezvanireza@gmail.com

2- M. A. Graduate of TEFL, Yasouj University. ali.sayadi1989@gmail.com

## مطالعه کیفی تأثیر آزمون جدید ورودی دوره دکتری رشته آموزش زبان انگلیسی بر روش تدریس، ارزشیابی کلاسی و برنامه درسی استادان زبان دانشگاه های ایران

### چکیده

با توجه به اهمیت و حساسیت آزمون ورودی دوره دکتری در ایران، انتظار می رود که این آزمون تاثیرات قابل توجهی بر برنامه های آموزشی تحصیلات تکمیلی در دانشگاه های ایران بگذارد. مطالعه حاضر تلاشی برای بررسی دیدگاه های استادان دانشگاه های ایران در رابطه با اثرات احتمالی آزمون جدید ورودی دکتری رشته آموزش زبان انگلیسی بر روش تدریس، ارزشیابی کلاسی و برنامه درسی آنها بود. برای نیل به این هدف، پژوهشگران این مطالعه، مصاحبه هایی دقیقی با ده استاد مجرب در تدریس دروس رشته آموزش زبان انگلیسی در مقطع کارشناسی ارشد انجام دادند. نتایج به دست آمده از تجزیه و تحلیل اطلاعات جمع آوری شده نشان داد که آزمون جدید، علی رغم اهمیت بسیار بالای آن، تاثیر چندانی بر روی روش تدریس و ارزشیابی کلاسی اساتید نگذاشته است. از دید شرکت کنندگان در این مطالعه، دلیل این امر عدم همسویی مهارت های مورد نیاز جهت موفقیت در آزمون ورودی دکتری و اهداف و انتظارات در نظر گرفته شده در دوره های کارشناسی ارشد و دکتری رشته آموزش زبان انگلیسی است. از طرفی دیگر، نتایج بیان گر آن بود که این آزمون، با اثرگذاری بر انتخاب منابع و موضوعات درسی استادان، موجب تغییر چگونگی طراحی برنامه درسی آنها شده است. این مطالعه در ادامه به بحث پیرامون تاثیرات نتایج به دست آمده در بافت تحصیلات تکمیلی رشته آموزش زبان انگلیسی در دانشگاه های ایران میپردازد.

### 1. Introduction

Various aspects of human performance have been assessed and decided upon mainly through the application of high-stakes standardized exams. The results of such exams often give rise to serious consequences which immediately and directly impact the educational and occupational decisions to be made for the applicants (Madaus, 1988; Stobart & Eggen, 2012). The primary use of such exams is "to ration future opportunity as the basis for determining admission to the next layer of education or to employment opportunities" (Chapman & Snyder, 2000, p. 458).

The potential influence of standardized high-stakes exams is more profound in countries such as Iran, where such exams are considered as the sole avenue to higher education and consequently to procuring prestigious jobs. More to the point, the prevalent imbalance between the number of higher education applicants and matriculates arguably has boosted the criticality of the high-stakes university entrance exams in Iranian context. A striking case in point is the Iranian Ph.D. Program Entrance Exams (PPEE),

where a considerably great number of applicants strive to qualify for the highest educational degree through sitting the Iranian PPEE. As facts and dates speak, more than 216000 and 240000 applicants sat the Iranian PPEE for state universities in 2013 and 2014, respectively (Sanjesh, 2014). However, despite such mounting tsunami of the applicants, a small proportion of them were qualified to enter such determining programs.

The Iranian PPEEs have been planned and administered in their new forms since 2012 when the monopoly on the development and administration of PPEEs was taken from universities and instead granted to Sanjesh Organization, a subsidiary of Iranian Ministry of Science, Research, and Technology (IMSRT). The Iranian PPEEs subsequently involved substantial modifications in their formats along with their administration, scoring, and content determination procedures. More elaborately put, as suggested by the authors (2014), the new exam, unlike the old one, has been more standardized and administered under uniform situational conditions and scoring procedures. Moreover, unlike the old exams which were planned in essay-type questions, the newly designed exams have been employing more objective items.

Widespread application of newly designed exams is assumed to give rise to disparate behaviors in classrooms, test-development institutions, and education ministries (Loumbourdi, 2013) due to the fact that exams are generally perceived as the substantial determiners, both positively and negatively, of what goes on in the classroom, the educational system and society as a whole (Hughes, 2003; Wall & Alderson, 1993). The effect of a test on teaching and learning practices is broadly known in the literature as washback effect (Alderson & Wall, 1993; Green, 2007; Richards & Schmidt, 2002). As Hamp-Lyons (1997, p. 295) concisely argues, washback, in general, “is one of a set of terms that have been used in general education, language education and language testing to refer to a set of beliefs about the relationship between testing and teaching and learning”. Messick also elaborates on the notion of washback and its tracing by maintaining that washback concerns “the extent to which the test [positively or negatively] impacts language teachers and learners to do things they would not otherwise necessarily do” (1996, p. 241). The test impacts are found to be stronger when language instructors and learners are to be in touch with high-stakes tests (Loumbourdi, 2013). Given their sensitivity, scope, and decisiveness, high-stakes tests exert direct and profound effects on the educational curriculum along with teaching materials and activities (Alderson & Wall, 1993; Ramezany, 2014; Shohamy, 2001). More

specifically, language instructors are often deemed to teach to decisive tests (Cheng, 2005; Davies, 1997; Green, 2007; Hayes & Read, 2004). In other words, they display a great tendency to tailor their teaching methodology and assessments to prepare their students for the tests and ignore the activities that have little contribution to the test takers' success and revolve around those with the maximum amount of contribution (Chapman & Snyder, 2000; Hughes, 2003). They are also constrained by a fear of disadvantaging their students on the exams even when they want to introduce new content or instructional practices (Chapman & Snyder, 2000). The present study accordingly is an attempt to scrutinize to what extent and how the new TEFL PPEE has influenced the M.A. instructors' teaching methodology, syllabus design, and class assessments from the view of university instructors.

## **2. Review of related literature**

### **2.1. Washback**

The concept of washback is prevalently highlighted in language teaching and testing literature, as well as general education. It has been referred to as both 'washback' and 'backwash' interchangeably and is basically defined as the influence that tests exert on teaching, learning, classroom practices, and syllabus and curriculum planning (Buck, 1988; Wall & Anderson, 1993). Such test effects are assumed to involve positive or intended and negative or unintended aspects, influencing educational practices with varying degrees (Bachman & Palmer, 1996; Hughes, 1989).

Bachman and Palmer (1996) theorize that washback is a feature of a wider process known as "test impact". However, close review of the related literature suggests that there has been considerable dispute over the distinction and clear definitions of the terms impact and washback. A great number of researchers, as argued by Loumbourdi (2013), claim that washback is part of something more general, called *the impact of testing*. As Hughes (2003, p. 2) asserts, "the term impact, as it is used in educational measurement, is not limited to the effects of assessment on learning and teaching but extends to the way in which assessment affects society as a whole". On the other hand, for some researchers (e.g., Cheng, 1998, 1999, 2003; Choi, 2008; Hamp-Lyons; Shohamy et al., 1996), washback and test impact might be used interchangeably.

### **2.2. Impacts of high-stakes exams**

As Stobart and Eggen (2012) argue, high-stakes tests, which have been utilized for over two thousand years now, may give rise to serious

consequences. Decisiveness and criticality of a given exam's results are assumed to be treated as two priming factors conditioning the intensity of the impacts arising from the administration of the exam (Loumbourdi, 2013; Qi, 2008). High-stakes exams would probably come to dominate most of the teaching time and materials and, more specifically, condition instructors' and learners' tendencies in their classroom behaviors (Chapman & Snyder, 2000; Hughes, 2003; Loumbourdi, 2013; Madaus, 1988; Stobart & Eggen, 2012 Wall & Alderson, 1993).

Many researchers (e.g., Davies, 1990; Green, 2007; Wall & Horak, 2006) have addressed the impacts of high-stakes exams on language instructors' pedagogical practices, and their findings, in general, indicate that teaching to the high-stakes exams is a common practice. Cheng (2005) and Green (2007), for instance, found test-related activities (e.g., offering test-taking tips, doing question analysis) and test-taking strategies instruction more prevalent in IELTS/TOEFL preparation classes than in regular classes. Instructors are assumed to benefit from teaching to high-stakes exams to help their students boost their scores (Noble & Smith, 1994) and to overcome their anxiety to sit the exams (Hayes & Read, 2004).

However, the related literature substantially underlines the negative impacts of high-stakes exams in different aspects as well. Chapelle and Brindley (2002), for instance, contended that the influence of a decisive test often tends to be harmful due to the temptation of language instructors to spend time on activities that will help their pupils succeed in the exam rather than to develop insightful skills. Noble and Smith (1994) also found in their study that high-stakes tests might affect teachers "directly and negatively" (p. 3), and that "teaching test-taking skills and drilling on multiple-choice worksheets, for example, are likely to boost the scores but unlikely to promote general understanding" (p. 6). Based on an extensive qualitative study of the role of high-stakes testing in elementary schools in the United States, Smith (1991b) similarly stated that the "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" (p. 8). Further, as it has been observed and pointed out by Prodromou (1995), many teachers who in general have not been nervous about teaching to other previous tests have developed anxiety during teaching to high-stakes standardized exams.

### 2.3. Washback effects of university entrance exams (UEE)

The potential influence UEEs exert on educational behaviors has been considerably taken into consideration in the literature of high-stakes exams washback effect. Researchers from China, Turkey, Hong Kong, and Japan, among others, have dealt with the English section of the university entrance exams in their own countries. Cheng (2004), for instance, examined the washback effect of Hong Kong Certificate of Education Examination on secondary school instructors' teaching behaviors and concluded that certain washback effects on the instructors' attitudes towards the modified exam could be detected, although the instructors' teaching activities were not influenced by the exam reform. Qi (2004) similarly scrutinized the washback effect of the English section of the National Matriculation Test in China. The results of her study revealed that there was a noticeable discrepancy between what the test constructors intended and what the teachers and students practiced in schools, indicating the inefficiency of the test for bringing about pedagogical changes in Chinese schools. Likewise, Watanabe (2004) investigated the washback effects of the Japanese university entrance exams on classroom instructions and concluded that the entrance exams gave rise to the teachers' use of the exam preparation as an opportunity to improve the English learners' proficiency.

### 2.4. Washback effects of Iranian university entrance exams

As the sole gatekeeper to higher education and potentially to procure highly qualified jobs, Iranian UEEs, also known as *konkour*, have stimulated substantial sensitivity. Therefore, exploring the washback effects ensued from the Iranian UEEs has been subject to discussion and research in Iranian context.

The washback effects of the undergraduate program entrance exams (UPEE) on Iranian EFL instructors' and learners' classroom behaviors have been addressed by many researchers (e.g., Ghorbani, 2008; Kazemi & Sayyadi, 2014; Mahmoudi & Abu Bakr, 2013; Ramezaney, 2014; Salehi & Yunus, 2012; Tabatabaei & Safikhani, 2011). As regards the learners, diverse studies have been carried out to explore how their motivations and quality of English learning were affected by the Iranian UPEE. Kazemi and Sayyadi (2014), for instance, examined the washback effect of the Iranian undergraduate program entrance exams on the motivations of high school students to learn English. The results of their study indicated that "the increased degree of concern about the Iranian UEE could serve as a booster for high school students to gradually leave their integrative motivation

behind as they approach the day of taking the UEE due to the fact that no place has ever been set for assessment of applicants' communicative abilities in the English section of Iranian UEE" (p.163). Considerable attempts have also been made to explore how teachers' pedagogical strategies, test development, and curricular planning, among others, are affected by Iranian UPEE. Ghorbani (2008), for instance, investigated the washback effect of the Iranian UPEE on high school language instructors' curriculum planning and instruction. The findings of his study suggested that UEE strongly affects the "what of teaching" but not the "how of teaching" of Iranian high school teachers, and that almost all high school language teachers, regardless of their teaching experience, educational background, gender, the school type, and the school location, perceived the negative effects of the UEE. Likewise, Ramezaney (2014) sought to investigate the Iranian UEE impacts on high school EFL teachers' curricular planning and instruction techniques and concluded that the teachers often adjust their teaching methodology and class assessment to the requirements of the UPEE, drawing on the concrete content and format of such a high-stakes exam. Further, in their study on the washback effect of Iranian undergraduate program entrance exam on high school instructors' classroom behaviors, Salehi and Yunus (2012) suggested that the UPEE negatively and implicitly influenced English teachers to teach to the content and format of the exam. They also asserted that little attention was given to three language skills of speaking, writing, and listening in the classroom as these skills are not tested in the UPEE.

Care was also taken by researchers to investigate the washback effects of the Iranian TEFL mastering program entrance exams (MPEE) on university instructors' and students' educational practices (e.g., Mohammadi, 2008; Mousavi & Amiri, 2011). For instance, Mohammadi (2008) sought to explore the washback of the Iranian TEFL MPEE on university language instructors' attitudes and teaching methodologies. He noted that the university instructors' attitudes and pedagogical practices were set to the demands of the exam. Moreover, Mousavi and Amiri (2011) made an attempt to study the impacts of the TEFL MPEE on university instructors' and students' teaching and learning behaviors. Not concurring with the general trend, the results emerging in their study revealed that the TEFL MPEE "had a negligible effect on the students' and professors' academic behaviors" (p. 103) and failed to motivate them to plan their learning and teaching practices towards this test.

Whereas substantial concern has been raised to study the washback effects generated by UPEE and MPEE, the washback effects of PPEE, as the review of the related literature unveiled, has garnered notoriously little attention, despite its sensitivity and significance. A single study exclusively concerned with the washback effects of the Iranian TEFL PPEE was undertaken by the authors (2014), where they were motivated to examine qualitatively the impacts of the new TEFL PPEE on the applicants' study plans and strategies. The study findings indicated that the introduction and dominance of multiple choice PPEEs has obviated the need to develop and maintain analytic, synthetic and evaluative qualities on the part of the applicants and induced them to tailor their study plans and strategies towards developing lower cognitive abilities of comprehension and recall of the information crammed.

It seems that the potential washback effects of the new Iranian TEFL PPEE on language instructors' classroom behaviors are under-researched. Accordingly, the incentive behind the present study was to undertake this strand of research by examining the washback effects of the new Iranian TEFL PPEE on university instructors' teaching methodology, class assessment, and syllabus design.

### **3. Research questions**

In this study the following research questions are addressed:

- 1. What washback effects, if any, does the new TEFL PPEE exert on university instructors' teaching methodology?**
- 2. What washback effects, if any, does the new TEFL PPEE exert on university instructors' class assessment?**
- 3. What washback effects, if any, does the new TEFL PPEE exert on university instructors' syllabus design?**

### **4. Method**

The washback effects of exams are recommended to be examined through undertaking qualitative inquiries (Cheng, 2004). The current study, accordingly, is motivated to qualitatively scrutinize the washback effects of the new Iranian TEFL PPEE on M.A. instructors' teaching methodologies, class assessment, and syllabus design.

#### **4.1. Participants of the study**

The participants of the study were ten experienced Iranian Ph.D. instructors teaching M.A. courses at four different Iranian universities including Shiraz,

Esfahan, Sheikh-Bahaie, and Shahre-kord Universities. The instructors had at least 4 years of teaching experience at universities and aged between 43 and 56. From among the instructors, three were females and seven were males. Table 1 shows the demographic information of the instructors.

Table 1: Demographic information of the interviewed instructors

No	Name*	Age	Gender	M.A. teaching experience	Current university
1	Mona	48	Female	6 years	<b>Shiraz</b>
2	Hesam	54	Male	13 years	Shiraz
3	<b>Maziyar</b>	46	<b>Male</b>	<b>7 years</b>	<b>Shiraz</b>
4	<b>Puya</b>	45	<b>Male</b>	<b>7 years</b>	<b>Shiraz</b>
5	<b>Nadim</b>	44	<b>Male</b>	<b>4 years</b>	<b>Shahre-kord</b>
6	<b>Akbar</b>	43	<b>Male</b>	<b>4 years</b>	<b>Shahre-kord</b>
7	<b>Behnam</b>	50	<b>Male</b>	<b>9 years</b>	<b>Esfahan</b>
8	<b>Atefeh</b>	48	<b>Female</b>	<b>7 years</b>	<b>Esfahan</b>
9	<b>Mehrdad</b>	44	<b>Male</b>	<b>5 years</b>	<b>Sheikh-Bahaie</b>
10	<b>Parisa</b>	45	<b>Female</b>	<b>5 years</b>	<b>Sheikh-Bahaie</b>

Note. The names are fictitious.

#### 4.2. Data accumulation procedures

In accord with the motivations behind the conduction of the present study, researchers sought to conduct profound interviews with the subjects of the study to elicit their views on the washback effect of the new TEFL PPEE on M.A. instructors' teaching methods and assessment procedures. To this end, they took the advantage of open-ended interview questions to closely examine the issues in detail and to provide the chance for the interviewees to adequately reflect on their insights. The interview questions were developed and asked in English. The participants, however, exercised a measure of absolute freedom in their choice of the language to respond to the questions. Interviews took ten to twenty minutes and were recorded with the permission of the interviewees. Once the data were accumulated, they were transcribed into written texts and then analyzed. The trustworthiness of the findings was also a critical consideration. Intensive care was taken to avoid bias through employing a prolonged and persistent field-work and accounting for participants' language verbatim accounts meticulously documented as recommended by McMillan and Schumacher (2006). When the responses were in Persian, the statements were carefully

rendered into English. In addition, the researchers frequently used member checking to check the data informally with the participants for accuracy during the interviews, and were sensitive to discrepant data that did not conform to the emerging patterns.

#### **4.4. Data Analysis management**

Qualitative investigations on the data collected through interviews are firmly tied up with interpretation of the data thorough content analysis (Elo & Kyngas, 2008; Glaser and Strauss; 1967). Accordingly, constant comparative content analysis, as suggested by Glaser and Strauss (1967), was utilized in the process of data analysis to code the transcribed interviews involving an inductive reasoning process of frequent sifting through the data to identify similarities and patterns of reference in the interview transcripts. Close analyses of the patterns and similarities subsequently led to the emergence of an evolving coding system for the categories. More specifically, the units of analysis and coding schemes were defined and developed during the process of the content analysis; then, the codes were transformed into categorical labels or themes that were repeated or appeared as patterns in the interviews. This iterative procedure, according to Patton (2002), is intended to help the researchers in “developing some manageable classification or coding scheme” as “the first step of analysis” (p. 463). Data analysis proceeded incrementally and once the coherence and saturation of the data were accomplished, conclusions were drawn based on the analyzed data.

### **5. Findings and results**

#### **5.1. Familiarity with the old and new PPEEs**

The instructors were initially asked to reflect on their degree of familiarity with the old and new TEFL PPEEs and the affected modifications in the development of the new one. The close analysis of their responses indicated that they, as experienced instructors, were quite conversant with both exams. Hesam, for instance, argued that:

The previous one was an essay type... and you had to write at least a paragraph for each question. The new one is multiple choice, as you know, and they [the test developers] have included other stuffs such as intelligence parts... The previous one was decentralized. It was decided by professors in different universities. This one is centralized..., (respondents' wording).

Puya also was of the opinion that he was well-versed in the modifications recently involved in the development of PPEEs and stated that:

About ten year ago when I took the [Ph.D. program entrance] exam, exams differed from one university to another. Some students had to take essay type exams, and some combination of essay type and multiple choice, depending on the universities. However, the new one has been standardized and developed in multiple choice format only, (respondents' wording).

In line with Hesam and Puya, Behnam also discussed his familiarity with the new PPEE and maintained that:

You know, it's not a long time the new exam has been used. I'm not sure but I think for about 3 years. In the past, every university had its own criteria to admit new Ph.D. students. The Ph.D. entrance exams resembled teacher-made exams, I suppose, because for example only the teacher who was in charge with teaching testing in a university developed the testing questions of the entrance exam. As I stated earlier, every university held its own exam to admit new students. The new ones have been nationalized and more standardized. Students from different parts of Iran are supposed to take the same exam, which was not the case in past. By the new exam, you have more options. In the past, you could only take your chance to be admitted in one university. Of course there were some ways through which you could find your way on a new university, but generally speaking you were more limited. Old ones were more subjective and essay type, (respondents' wording).

## **5.2. Exam impacts on the instructors' teaching methodology**

As regards the questions concerning the effect of the new TEFL PPEE on the instructors' pedagogical methodology, in-depth analysis of their insights revealed that they indeed did not teach to the new PPEE. More specifically put, nine among ten instructors did not perceive preparing their M.A. students for the TEFL PPEEs as a concern while planning the dynamics of their teaching procedures. Mehrdad, Parisa, Maziyar and Behnam, for instance, argued that they committed themselves to the application of the

teaching principles and procedures which ensured accomplishing their classroom objectives. Mehrdad accordingly cited that:

The classroom objectives are what actually count for me. Before I start instructing a certain course to my students, I assume the objectives of the course based the students' capacities and their required degree of insightfulness on certain technical topics. Subsequently, I decide upon the degree of significance in different technical areas and contents to be discussed in the class based on my own intuition and ponder the most fruitful techniques to be utilized in order to make my students insightful in the covered contents, especially the most significant ones. PPEE is not a consideration in my decisions, (translated by the researchers).

Likewise, when Parisa was asked whether the new exam had exerted any influence on her pedagogical practices, she succinctly held "no, *my* objectives in *my* classes are quite different from the ones to achieve in Ph.D. entrance exams".

Besides, Atefeh reckoned the potential negative impacts of the new PPEE on their students' cognitive behaviors and learning strategies as the major grounds for eschewing the teaching activities which called the students' attentions to the new PPEE. She decried the new PPEE and voiced that:

As a matter of the fact, sensitive standardized tests are supposed to critically influence applicants' learning strategies and teachers' teaching methods and assessments. As far as my teaching and assessment techniques are the considerations, I should confess the new Ph.D. exam has not been that influential. Let me tell you why. You know, the new exam makes the applicants focus on many trivial details and avoid shaping insightful perceptions of the contents they study. Therefore, if I teach to the new exam, I will directly or indirectly encourage my students to limit themselves to learning at low cognitive levels, (respondent's wording).

Hesam also declared that he had never placed any import on the new PPEE. He was subsequently asked whether he did not regard preparing M.A. students for the Ph.D. programs as one of the priming practices of the instructors. Voicing critical concerns about the new TEFL PPEE, he responded that:

Actually, we *do* prepare the students for the Ph.D. programs. But the point is that the new exam does not indeed filter the Ph.D. applicants based on the capacities Ph.D. students ought to possess. We are indeed preparing our students for the Ph.D. programs and not the entrance exam. This is because the capacities the students need to have for the entrance exam are not the same as the ones required for the Ph.D. programs. How much do you think the superficial knowledge memorized for the sake of passing the new entrance exam can help the Ph.D. students make insightful experts in their field? What we need in Ph.D. programs is students with highly evaluative capacities. Nevertheless, the Ph.D. students who come and follow the studies are really weak. It shows that the [new] exam is not actually filtering out good students. So I see no logic in placing any import on such an exam. Therefore I prefer to stick to my old classroom practices yet keep my students insightful (respondent's wording).

However, in Puya's view, the new exam has induced him to place more emphasis on the bits and pieces in the technical contents. More elaborately put, he cited that:

As you know, for multiple choice items students go through the materials more detailed in order to get familiar with dedicate aspects which might be important for exam purposes... so when I come across an important point which might be raised as Ph.D. exam item, I draw my students attention to the point, ( respondent's wording).

### **5.3. Exam impacts on the instructors' class assessment**

The instructors' reflections on the washback effect of the new TEFL PPEE on their class assessments revealed that, not unlike their teaching methodologies, the instructors' class assessments were not substantially touched by the new exam. In other words, to the instructors, the policies behind the new PPEE were not in line with their course objectives. Mehrdad, Hesam, Maziyar, and Parisa correspondingly declared that they designed and developed class exams in accordance with their course objectives which basically departed from the ones to be accomplished by the new TEFL PPEE. Hesam, for example, concisely stated that he had regularly expected his students to shape in-depth judgmental and evaluative

views on the contents by the end of each course and the new PPEE was potential of overshadowing that fundamental view. Further, Maziyar precisely asserted that:

The exam is not my purpose. I don't want to have a preparatory course. My purpose is just to make the students familiar with the basic issues in testing. And because of this, it's not important to me whether they are in line with the [new] test or not, (respondent's wording).

Mehrdad also disparaged the new exam's format and cited "assessing M.A. students through multiple choice exams is a critical mistake, and filtering Ph.D. applicants through multiple choice exams is a disaster". He further pointed to the reasons why he gave the TEFL PPEE short shrift while planning exams to assess his M.A. students and held that:

I assess my students on the basis of my course objectives. The way Ph.D. entrance exam is planned and developed gives rise to neglecting comprehensibility of the technical areas. The exam questions in M.A. course need to be essay type due to the fact that M.A. students' capacity to discuss technical topics within one or two paragraphs is an undeniable necessity, which is not taken into consideration in Ph.D. exams, (translated by the researchers).

Mehrdad's critical comment is also represented in Parisa's view when she cited that:

Well, I think what actually the new exam asks the applicants to do is not that logical and acceptable. Why should my class assessment be affected by something illogical and unacceptable? I follow my objectives. For example, when I teach test reliability and validity to my pupils, I expect them to develop full-scope or comprehensive knowledge of such critical issues in their minds rather than to memorize the superficial aspects of each. I prefer to develop exams which assess my pupils' analytical abilities rather than their memorization capacities, (respondents' wording).

The other instructors also simply confessed that the new TEFL PPEE did not have noticeable effects on their class assessments. Puya, however, pointed at some minor effects and stated that:

I cannot say its [PPEE's] effect has been that noticeable. But a couple of years ago when most universities followed the essay type version, I actually designed test items based on

that single format, I mean all of my questions appeared in the form of essay type items. But since the change in the policy, I occasionally add multiple choice questions to my exams as well, (respondent's wording).

#### **5.4. Exam impacts on the instructors' syllabus design**

As the instructors' syllabus design is the concern, careful examination of their views indicated that the PPEE is nevertheless a determining account when it comes to selecting and justifying the contents to be covered in the TEFL M.A. programs. The instructors were of the opinion that the questions posed in new PPEE are suggestive of the technical areas of interest. It induces the instructors to parallel the covered technical contents with those raised in the PPEEs and consequently involve occasional modifications in the contents to be covered in the programs. Atefeh accordingly discussed about the effects of the new exam on her content selection. More specifically, she argued that:

Although I do not frequently and directly follow the Ph.D. exam questions each year, I have been informed that the areas from which the questions are selected involve changes time to time. It shows that the important areas and topics are subject to change. I try to get aware of these changes and select materials which can cover the new and important areas more efficiently. But I try to teach these new contents through my old teaching methods.

Maziyar also discussed the modifications in his testing course syllabus arisen by the new exam and stated that:

Last year I added a topic like dynamic assessment to my syllabus, but at the end of the term I decided to omit it from the syllabus. But later on, when I saw that the same area was presented in the Ph.D. Exam, this gave me kind of support that I was going on the right track.... So probably, I had very minor modifications in my syllabus in terms of standing more sessions on some topics and adding some topics to my syllabus. I used to stand on dynamic assessment for one session, but now I spend two or three sessions on it, (respondent's wording).

Nadim and Akbar also asserted that although they did not change their previously utilized materials, they frequently introduced the books from

which the recent Ph.D. exam questions have been selected. Akbar, for instance, asserted that:

I have been using the conventional textbooks since my first experience in teaching M.A. courses. I think they are informative enough, and it has not been necessary to change them. But as some new technical areas and topics are occasionally added to Ph.D. exam, I actually try to introduce newly published textbooks to my students, (respondent's wording).

## 6. Discussion

The potential decisiveness of high-stakes public exams has mounted their sensitivity. Recent growth in high-stakes examination systems has substantially gave rise to renewed concerns about the influence of highly critical exams on the behaviors in classrooms and more broadly in various educational programs (Alderson & Wall, 1993; Rezvani & Sayyadi, in press; Stecher et al., 2000). It might be arguably suggested that the plans pursued by teachers, learners and other in-touch parties are perceived to be markedly impacted by high-stakes exams.

The intensity of the consequences arising from the introduction of a critical exam ties up with the decisiveness and significance of the exam (Loumbourdi, 2013; Qi, 2008). Loumbourdi (2013), accordingly, considers the decisiveness and significance of a high-stakes exam as two elements which are strongly indicative of the potential prominent washback effect created. He succinctly argues, in other words, that "the more high-stakes a test is, the more washback it creates" (p. 85). As far as the exam impacts on instructors' teaching methodologies and class assessments are of concern, the high-stakes exams are assumed to be the influential determiners of instructors' tendencies in their pedagogical and judgmental classroom practices (Cheng, 2005; Davies, 1990; Green, 2007; Hayes & Read, 2003; Hughes, 2003; Loumbourdi, 2013; Wall & Alderson, 1993). Paving the way to procuring highest educational degree and consequently prestigious jobs, the high-stakes Iranian TEFL PPEE is also perceived to be a highly decisive and significant exam (Rezvani & Sayyadi, in press). However, in-depth qualitative inquiry on the washback effect of the new Iranian TEFL PPEE on the basis of university instructors' insights indicated that the new exam despite highly considerable sensitivity and import does not create noticeable washback effect on M.A. instructors' pedagogical methodologies and class assessments. Such findings concur with the results of Tsagari's (2006) study

on high-stakes exams' washback in Greek context where he asserts that "high-stakes tests, powerful as they are, might not be efficient agents for profound changes in an educational context..." (p. 310).

In the light of the new exam's incapacity to create considerable washback effects on the instructors' teaching methodology and class assessment procedures, it would appear worthwhile to possibly inspect the reasons in the critical comments of the interviewees taking part in the present study. Hesam, for instance, underlined the defective validity of the new exam as the ground for questioning the new exam and avoiding test-centered instruction. This is in accord with the findings of the writers (in press) in their study on the reliability and validity of the new TEFL PPEE where they came to the point that the new exam is perceived to demonstrate defective face, content, predictive, and construct validities. Such views may potentially reinforce Messick's (1996) perspective on the relationship between the validity and washback where he maintained that washback is a sub-product of construct validity and positive washback can be promoted by heavily relying on construct validity and avoiding construct underrepresentation and construct irrelevant variance. Likewise, Morrow (1986) and Frederiksen and Collin (1989) also argued on the close association between the degree of validity in a given test and the washback it builds. Therefore, it seems that the defective validity ascribed to the new TEFL PPEE might be of the grounds highlighting the instructors' reluctance to teach to the new exam. In addition, the new TEFL PPEE's tendency to tap its applicants' low-level cognitive competence (see Rezvani & Sayyadi, in press) was also deemed as an underlying factor inducing the instructors to abstain from tailoring their teaching methods and class assessment practices to preparing their students for the PPEE. In other words, in the respondents' views, the new exam is more bits-and-pieces-oriented, obviating the need to make evaluative judgments, and thus, provoking the applicants to confine themselves to comprehending and memorizing details to tackle the questions. Seeking to make university instructors and experts, M.A. students and especially Ph.D. candidates, as the respondents maintained, are nevertheless required to develop analytic and judgmental competences in order to gain profound insights on their field of study. This supports Noble and Smith's (1994a) findings which indicated that high-stakes examination is unlikely to promote general understanding of applicants. Hence, it appears that the prevalent inconsistency between the capacities expected to be developed in M.A. and Ph.D. programs and those required for sitting the TEFL PPEE has induced the M.A. instructors to avoid outweighing

concerns about preparing their students for the new exam. More simply put, as Hesam asserted, the instructors are potentially preparing their students for what they are to possibly encounter in Ph.D. programs rather than in Ph.D. entrance exams. This conforms closely to the findings emerging in the study recently submitted by the writers where they voiced concerns about the predictive and construct validities of the new TEFL PPEE and highlighted its incapacity to predict effectively the applicants' future performance through taking the advantage of questions developed based on clear construct characterization. Furthermore, close analysis of the respondents' insights suggested that the instructors sought to set their pedagogical classroom practices on the basis of their own course objectives and did not confine themselves to test-driven instruction. This stance supports Madaus's (1988) as well as Wall's (2000) criticisms on high-stakes exams where they held that treating high-stakes exams as a consideration in pedagogical practices substantially limits teachers' creativity.

Turning to the instructors' syllabus design, the new exam, on the contrary, creates a critical washback effect on their material and content selection. As emerged from the participants' views, the technical areas and contents covered in the new exam make the instructors insightful about the most recent areas of interest. It might consequently inspire the instructors to resort to various materials to cover the most critical technical realms and accordingly assign varying periods of time to efficiently teach the areas. However, as Atefeh argued, they employed their old teaching methodologies to teach the new contents and areas.

## **7. Conclusion and implications**

The present study was purposed to qualitatively scrutinize the potential impacts of the new PPEE on teaching methodology, class assessment, and syllabus design of Iranian university instructors teaching M.A. courses. The results emerging from the close analysis of the respondents' insights indicated that the new exam, despite its considerable criticality, fails to exert remarkable influence on the instructors' teaching methodology and class assessment owing to its defective validity and its tendency to tap the applicants' low-level cognitive capacities. In regard to the new exam's impact on the instructors' syllabus design, the research findings suggested that the new exam nevertheless creates considerable washback effect on the instructors' material and content selection.

The results of the current study might raise university instructors' awareness of the potential impacts of the new TEFL PPEE on their syllabus

design. They may also lead to enhancing the university instructors' knowledge of the potential criteria applied by instructors to assess TEFL M.A. students in classroom context. Further, the results may have important implications for policy makers and test developers to take into consideration how the new exam is viewed and treated by the university instructors. The perceived inconsistency between the objectives supposed to be met by the new exam administration and those expected to be achieved in TEFL M.A. and Ph.D. programs calls for taking the service of the in-touch experienced experts' insights suggested by the current study in order to develop and administer TEFL PPEEs to be more closely aligned with the assumed capacities to be fostered in Iranian postgraduate education. Last but not least, the study findings concerning the new exam's incapacity to tap into its applicants' more complex cognitive behaviors might immensely give rise to an inclination towards a more scientifically profound approach involving better construct characterization along with more accountable formats and content determinations when developing PPEE.

### References

- Ary, D., Jacobs, L. C., & Sorensen, C. (2006). *Introduction to Research in Education*. B Belmont: Wadsworth, Cengage Learning.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Buck, G. (1988). Testing listening comprehension in Japanese university entrance examination. *JALT Journal*, 10, 15-42.
- Chappelle, C. A., & Brindley, G. (2002). Assessment. In N. Schmitt (Ed.), *an Introduction to Applied Linguistics* (pp. 267-287). New York: Oxford University Press Inc.
- Chapman, D. W., & Snyder, C. W. (2000). Can High-Stakes National Testing Improve Instruction: Re-Examining Conventional Wisdom? *International Journal of Educational Development*, 20 (6), 457-474.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 147-170). Mahwah, New Jersey: Lawrence Erlbaum.
- Cheng, L. (1998). Impact of a public English examination change on students' Perceptions and attitudes toward their English learning. *Studies in Educational Evaluation*, 24, 279-309.
- Cheng, L. (1999). Changing assessment: washback on teacher perceptions and actions. *Teaching and Teacher Education*, 15 (2), 253-271.

- Cheng, L. (2003). Looking at the Impact of a Public Examination Change on Secondary Classroom Teaching: A Hong Kong Case Study. *Journal of Classroom Interaction*, 38 (1), 1-10.
- Cheng, L. (2005). *Changing language teaching through language testing*. Cambridge: CUP.
- Davies, A. (1997). Demands of being professional in language testing. *Language Testing*, 14(1), 328-339.
- Choi, I. C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, 25 (1), 39-62.
- Frederiksen, J. R., & Collins, A. (1989). A Systems Approach to Educational Testing. *Educational Researcher*, 18 (9), 27-32.
- Ghorbani, M. R. (2008). *Washback effect of the university entrance examination on Iranian pre-university English language teachers' curriculum planning and instruction* (Doctoral dissertation), University of Putra, Malaysia.
- Green, A. (2007). Washback to learning outcomes: a comparative study of IELTS preparation and university pre-essential language courses. *Assessment in Education: Principles, Policy & Practice*, 14(1), 75-97.
- Hamp-Lyons, L. (1997). Washback, impact and validity: Ethical concerns. *Language Testing*, 14, 295-303.
- Hawkey, R. (2006). *Impact Theory and Practice: Studies of the IELTS test and Progetto Lingue 2000*. Cambridge: CUP.
- Hawkey, R. (2009). *Examining FCE and CAE*. Cambridge: CUP.
- Hayes, B., & Read, J. (2004). IELTS Test Preparation in New Zealand: Preparing Students for the IELTS Academic Module. In Y. Cheng, Y. Watanabe, & A. Curti (Eds.), *Washback in Language Testing: Research, Context, and Methods* (pp. 97-111). Mahwah, New Jersey: Lawrence Erlbaum.
- Hughes, A. (2003). *Testing for Language Teachers* (2nd edition). Cambridge: CUP.
- Kazemi, A., & Sayyadi, A. (2014). Effect of University Entrance Exam on Gifted High School Students' Motivation Scrutinized: An Iranian Perspective. *International Journal of Applied Linguistics & English Literature*, 3 (5), 150-165.
- Loumbourdi, L. (2013). *The Power and Impact of Standardised Tests: Investigating the Washback of Language Exams in Greece*. Oxford: Peter Lang Edition.
- Madaus, G. F., (1988). The Influence of Testing on the Curriculum. In L. N. Tanner (Ed.), *Critical Issues in Curriculum: 87th Yearbook for the National Society for the Study of Education* (pp. 83-121). Chicago: University of Chicago Press.
- Mahmoudi, L., & Abu Bakr, K. (2013). Iranian Pre-university English Teachers' Perceptions and Attitudes towards the Iranian National University Entrance

- Exam: A Washback Study. *International Journal of Education & Literacy*, 1(2), 57-83.
- Messick, S. (1996). Validity and Washback in Language Testing. *Language Testing*, 13(1), 241-256.
- Mohammadi, M. (2008). *High-Stakes Testing Washback: A Survey on the Effect of Iranian M.A. Entrance Examination on Teaching and Testing Methods*. Paper presented at the First Conference on English Language Teaching & Literature, Islamic Azad University, Roudehen Branch.
- Mousavi, M., & Amiri, M. (2011). The Washback Effect of TEFL University Entrance Exam on Academic Behaviors of Students and Professors. *Journal of English Studies*, 1(2), 103-144.
- Morrow, K. (1986). The Evaluation of Tests of Communicative Performance. In M. Portal (Ed.), *Innovations in Language Testing: Proceedings of the IUS/NFER Conference* (pp. 1-13). London: NFER/Nelson.
- Noble, A. J., & Smith, M. L. (1994). *Measurement-driven reform: Research on policy, practice, repercussion*. Tempe, AZ: Arizona State University, Center for the Study of Evaluation.
- Prodromou, L. (1995). The Backwash Effect: From Testing to Teaching. *ELT Journal*, 49 (1), 13-25.
- 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, New Jersey: Lawrence Erlbaum.
- Ramezany, M. (2014). The washback effects of University Entrance Exam on Iranian EFL Teachers' Curricular Planning and Instruction Techniques. *Social and Behavioral Sciences*, 98, 1508-1517.
- Richards, J.C., & Schmidt, R. (2002). *Longman Dictionary of Language Teaching and Applied Linguistics, 3rd edition*. London: Longman.
- Rezvani, R. & Sayyadi, A. (2014). A qualitative study on the Washback effects of the new Iranian TEFL Ph.D. program entrance exam on applicants' study plans and strategies: Instructors' and applicants' insights. *Journal of Research in Applied Linguistic Studies*. 5(2), 113-127.
- Salehi, H., & Yunus, M. M. (2012). The Washback Effect of the Iranian Universities Entrance Exam: Teachers' Insights. *GEMA Online Journal of Language Studies*, 12 (2), 609-628.
- Shohamy, E. (2001). *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. London: Longman.
- Shohamy, E., Donitsa-Schmidt, S. & Ferman, I. (1996). Test Impact Revisited: Washback Effect over Time. *Language Testing*, 13(3), 298-317.
- Smith, M. L. (1991b). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11.

- Stobart, G., & Eggen, T. (2012). High-stakes testing: Value, fairness and consequences. *Assessment in Education: Principles, Policy and Practice*, 19(1), 1-6.
- Stecher, B., Barron, S., Chun, T., Krop, C., & Ross, K. (2000). *The effects of Washington education reform on schools and classrooms*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Tabatabaei, O., & Safikhani, A. (2011). The Washback Effect of University Entrance Exam on EFL Teachers' Methodology and Test Development. *Journal of English Studies*, 1(2), 145-172.
- Tsagari, K. (2006). *Investigating the Washback Effect of a High-Stakes EFL Exam in the Greek context: Participants' Perceptions, Material Design and Classroom Applications* (Unpublished doctoral dissertation), Lancaster University, UK.
- Wall, D., & Alderson, J. C. (1993). Examining Washback: The Sri Lankan Impact Study. *Language Testing*, 10 (1), 41-69.
- Wall, D., & Horak, T. (2006). *The impact of changes in the TOEFL examination on teaching and learning in central and eastern Europe: Phase 1, the baseline study*. Princeton, New Jersey: Educational Testing Service.