cogent
social
sciences

*Corresponding author: Mohammad Salehi, Sharif University of Technology, Iran
E-mail: m_salehi@sharif.ir

SOCIOLOGY | RESEARCH ARTICLE

An investigation of a nationwide exam from a critical language testing perspective

Mohammad Salehi[1]* and M. Tarjoman[2]

Abstract: The present study aims at investigating two main test parties' viewpoints—those of university teachers and TEFL MA students—regarding a nationwide Exam to find out their opinions about using the current method of the TEFL MA University Entrance Exam, those aspects of the test which can be improved and the amount of power or control that these two main test parties own at different phases of the Exam.The survey also examined the washback effect of this exam on university teachers'methodologies and instructions. To this end seven university professors and sixty TEFL students who had passed the exam were selected using a convenience random sampling. Subsequently, a validated researcher-made questionnaire was administered to the students. In order to collect more reliable data, some students were randomly selected to be interviewed so as to cross-check the data collected through the questionnaire. The results of this study indicated that all the university professors and the majority of the students demanded to have control over the content, the time of the administration and other issues related to the TEFL MA Entrance Exam. They claimed that they had no power in the test development and administration processes. They also believed that the test should serve as an indicator of language ability or knowledge, rather than test-taking skills. In addition, the results indicated that the exam had a negligible effect on professors' academic

ABOUT THE AUTHORS

Mohammad Salehi is a language testing specialist living and working in Iran. He has written papers on topics like alternative assessment, validity, *C*-Test and cloze test, reading assessment, grammaticality judgment tests, differential item functioning, needs analysis, receptive intelligibility, motivation, self-regulated learning, learner autonomy and acquisition. His main interest is language testing and second language acquisition research. He is interested in any type of research that brings language testing and second language acquisition together. Recently he has taken a keen interest in critical language testing.

M. Tarjoman received her BA degree in English Language from Gilan University, Iran in 1999 and her MA degree in TEFL from Islamic Azad University, South Tehran Branch, Tehran, Iran in 2016. Now she is working at high schools as an English teacher.

PUBLIC INTEREST STATEMENT

Critical language testing is an important concept in language testing. Decisions about individuals may be made without their having any control over the process and content of language tests. A nationwide was examined in terms of language testing issues in this paper. A questionnaire was constructed and distributed to some students who had already taken the test. The results showed that test takers wished to have control over the testing process. But they did not have any control over anything. In other words, asymmetricality prevailed throughout the testing process. Policy makers must have imposed their ideologies over the test takers who were powerless. We also examined washback or curriculum alignment in this paper. Washback is defined as the effect of testing on pedagogy. In other words, we wanted to see if teachers and professors tailored their teaching according to the demands of the test. The results revealed that few teachers or professors did curriculum alignment. Their teaching was not informed by test results.

cogent·oa

behavior and teaching techniques, hence the low degree of curriculum alignment.

Subjects: Teaching Learning - Education; Educational Research; Education Studies; Bilingualism / ESL; Theories of Learning; Classroom Practice; Language & Linguistics; Language Teaching & Learning

Keywords: Questionnaire; interview; TEFL MA examination; critical language testing; washback

1. Introduction

Assessment is generally understood as a sourse of gathering information regarding what people have learned to make decisions about them. However, in the past decade, the traditional notion of tests has undergone a major shift after the advent of critical pedagogy. In this new approach which is called democratic assessement the test takers' views, experiences and expectations taken into the account and validated in the educational system. Furtheremore, the learners are at the same time participants in the educational systems and have their own rights that are respected.

Philosophers and educationalists have been considering tests as powerful devices in society for some time. For Madaus (1991), tests represent a social technology deeply embedded in education, government and business; as such, they provide the mechanism for enforcing power and control. In fact, tests are most powerful as they are often the single indicators for determining the future of individuals. Being considered as the main criteria for acceptance and rejection, tests dominate other educational aspects including the curriculum, textbook and instruction. Pennycook (2001) suggested a highly skeptical perspective for analyzing human freedom, a view different from what Marxists, structuralists, and poststructuralists propose for the analysis of social and political power. This critical view identifies the missing link in social chains, i.e. the individuals. In fact, Pennycook questions the Marxist view that results in power relations through class relations instead of individulas' freedom to decide about what they want. This approach is similar to the one followed by Foucault (1975). Taking tests into consideration, Foucault refers to a new modality of power where each test taker obtains his/her individuality as the scores demonstrate some features that identify the test taker as a particular "case". Auerbach (1995) asserts that decisions related to curriculum development, materials and language use processes are to be shaped politically and based on professional measures, although they may impose specific ideologies and socioeconomic roles on learners. In the same vein, Rea-Dickins (1997) argues that a lot of stakeholders are influenced by test scores including students, teachers, parents, administrators, and even the public. Farhady (2006) classifies these stakeholders into educational, social and political groups, and asserts that each group has a different type of interest and intention for utilizing tests as a source of power.

Similarly, Noam (1996) viewes tests as a tool to impose implicit ideas about success, knowledge and ability. He notes that, "How we assess can support careers and make people successful, but it can also destroy people's careers, place unfair burden on individuals self perception and unnecessary hurdles in the path of their achievement" (p. 9). Foucault (1980) stated that viewing tests in reference to social, educational and political contexts situates the field of testing in the domain of critical testing. As for language testing, this approach is referred to as critical language testing.

Shohamy (2001), one of the most well-known figures in the comprehensive model counting Critical Language Testing principles, proposed her theory of critical language testing advocating that test takers should be given the right to question the test itself, its value and its methods. She also believes that test takers' views regarding test are sometimes unique and test makers can learn from test takers. According to her, in the past few years, language test developers have begun to show a growing interest in the role that language tests play in society, specifically addressing the issues including the extent to which tests define linguistic knowledge, determine membership, classify people, and stipulate criteria for success and failure of individuals and groups. Shohamy (1992) also describes the utilization of external language tests to affect and drive foreign language learning in

the school context. Although there is general agreement in the field of language assessment on the existence and importance of the washback phenomenon, there still is considerable variety in opinions about how washback functions (Bailey, 1996). The way it is perceived depends on the theoretical standpoint researchers take as well as the educational context that they are associated with. Washback could be either positive or negative, or both. Its influence on teaching and learning is a complex phenomenon. Undertaking the review of World Bank research carried out in 14 African countries on the effects of examinations over the curriculum and on the quality of teaching and learning, Kellaghan and Greaney (1992) concluded that most of the focus is on the negative effects that high-stakes examinations can have on the classroom (e.g. the neglect of subjects and skills which are not examined, the excessive use of past exam papers as teaching material, frequent practice in examination-taking techniques) and on individuals (e.g. the promotion of a passive concept of learning). There are other researchers who believe in positive impact of the tests. Alderson (1986), for example, argued for innovations in the language curriculum through innovations in language testing. Noble and Smith (1994) proposed that "the goal of current measurement-driven reforms in assessment is to build better tests that will drive schools toward more ambitious goals and reform them toward a curriculum and pedagogy geared more toward thinking and away from rote memory and isolated skills" (as cited in Cheng & Curtis, 2004, p. 5). Crooks (1988) emphasized that tests can have a positive effect on learning if teachers stress the need for deep learning rather than surface learning. Meanwhile, many researchers have claimed that it is difficult to determine the positive or negative effects of the tests. For instance, Alderson and Wall (1993) stated that there is no simple relationship between tests, good or bad, and their effects on the classroom. Messick (1996) highlighted that, "a poor test may be associated with positive effects and a good test with negative effects because of the other things that are done or not done in the education system" (p. 242). Brown and Hudson (1998) claimed that when the assessment procedures correspond to the course goals and objectives, a positive washback effect occurs. In contrast, if the assessment procedures do not reflect the curriculum objectives, they are likely to generate a negative washback. There is a centralized educational system in Iran in which undergraduates should take a high stakes test to pursue their studies. TEFL MA UEE is as high stakes test administered annually to select qualified applicants for graduate studies. The process of this exam includes a wide range of stake holders like undergraduate students, university professors and those who develop, administer and make decisions related to the test. While they have been studies regarding other high stakes tests in an Iranian context (e.g. Rezaee & Salehi, 2008; Salehi, 2011), so study has ever investigated a high stakes test in this country in terms of critical language testing. To this end, the study used Shohamy's (2006) theory of critical language testing as the theoretical framework, this study investigated test takers' views regarding the TEFL MA exam and its uses for high-stake purposes as well as the amount of power that different groups of stakeholders assume in the exam. It also aims at studying the washback effect of MA Entrance Examination on university teachers' methodologies and attitudes. Therefore, the purpose of this study can be summarized as the following:

(1) Providing the stakeholders with a chance to view the test critically.

(2) Empowering the stakeholders, so that they can voice their concerns and opinions about the TEFL MA exam.

(3) Transmitting the stakeholders' comments on the TEFL MA exam to TEFL MAEE testees, and the anonymous test constructors of this exam, so that they can gain insights and improve the quality of the tests accordingly.

(4) Being hopefully useful for the university teachers teaching the English major at the BA level.

(5) Informing the supreme council of education authorities of the stakeholders' opinions about the current method of the TEFL MA exam for getting higher degrees.

Having these points in mind, the researchers tried to find answers to the following questions:

(1) Are the stakeholders and test developers, as important parties in the social context of the test, equally powered?

(2) What are the stakeholders' views regarding the adoption of the TEFL MA exam as one of higher degree requirements?

(3) In the stakeholders' views, are there any aspects of the TEFL MA exam which can be improved?

(4) Does the TEFL MA exam have any washback effect on the professors' performance in undergraduate cources?

2. Methodology

2.1. Participants

To consider the critical language testing issues regarding the TEFL MA Entrance Exam, the researchers invited representatives of two social groups who took the exam or were affected by the test results, including TEFL MA students and TEFL university professors, to take part in the present study. To have a sample representing the population, the students and the professors were selected from different universities. The participants, therefore, included seven university professors (four males and three females) and 60 students of TEFL (17 males and 43 females) from four different Iranian universities. All students taking part in the survey had had the experience of taking the TEFL MA Entrance Exam before participating in the study. Thus, they were cognizant of the content and format and were therefore more likely to view it critically. Five out of the seven instructors were Ph.D holders in TEFL and two had an MA in TEFL. All the teacher participants, ranging from 32 to 45, had over 6 years of teaching experience at the university level. At the time of the study, some of the student participants were still studying TEFL at an MA level and some of them had graduated. All of the participants were selected randomly from four different universities.

3. Instruments

The instruments used in this study were a researcher-made questionnaire—which enjoyed expert judgments, as well as semi-structure interviews to draw out insightful information about participants' opinions, experiences, beliefs, feelings, and thoughts about a topic (Anderson & Arsenault, 1998; Ary, Jacobs, & Razavieh, 2002; Best & Kahn, 2003; Holstein & Gubrium, 1995; Rubin & Rubin, 1995). The questionnaire also consisted of open-ended questions to permit greater freedom of expression. Out of 75 TEFL MA students receiving a copy of the questionnaire, 60 turned it back answered. The interview questions were developed and asked in English. All of the interviewees preferred to respond to the questions in English.

4. Students' questionnaire

4.1. Rationale and aim

The purpose of this questionnaire was to explore the amount of power and control of students in the administration of TEFL MAEE. It also intended to investigate the students' viewpoints on the exam (Tables 1 and 2).

4.2. Structure, content and sources

The students' questionnaire consisted of three types of questions. In the first type, the students were asked to provide some information on the test makers, the time that the test was made, the amount of control of different test parties in test development and their awareness of the main sources of the exam questions. In the second type, there were 13 items probing participants' viewpoints of TEFL MA Entrance Exam issues. The items in these two parts were designed on a semantic differential scale where a characteristic statement preceded five boxes (1 = have no idea; 2 = never/strongly disagree; 3 = seldom/slightly agree; 4 = frequently/often agree; 5. always/strongly agree). The third part of the questionnaire consisted of three questions with an open-ended format to permit greater

cogent · social sciences

Table 1. Structure and themes of students' questionnaire		
Structure	**Contents**	**Items**
Theme 1	*Demographic features:* • Age, sex, major, place of university • Participation in the exam	5
Theme 2	*The amount of information about the:* • test makers, the time that the test was made, the main sources of the exam questions, the degree of their involvement and control in test development, provision of information about the circumstances of different universities by the ministry of higher education	6
Theme 3	*Students' viewpoints:* • Whether they prefer to be tested by their own university as opposed to a centralized testing method • Whether or not the exam questions should match their learning styles and individual differences • Whether they prefer the present testing method (multiple –choice) as opposed to open-ended questions • Their overall satisfaction with the quality of study materials in BA courses as the main sources for the exam, the time required to complete the test, the feedback they receive on their test results, the quality of knowledge they gain from the test • The extent to which they thought the test provided a true and correct reflection and indication of their actual knowledge • Aspects of the TEFL MAEE that can be improved	13

Table 2. Major categories, themes, and coding schemes	
TEFL MA students' viewpoints about:	**Theme code**
Test makers, the time that the test was made,	T1
The main sources of the exam questions	T2
The involvement and control of different test parties	T3
The present testing method of the exam	T4
(multiple-choice) items as an appropriate method	T5
Improvement of any aspects of the TEFLMAEE	T6
Being tested by their own university as opposed to a centralized testing method	T7

freedom of expression on the part of the participants. A clear outline of the students' questionnaire is displayed in Table 3.

4.3. Design and validation procedure

The students' questionnaire was first distributed to 12 students at Khorram Abad University, in the presence of their professors and one of the researchers. Following the questionnaire administration, some defective items were revised and modified. Furthermore, the students' questionnaire enjoys expert validity, two of university professors reviewed the questionnaire as the result of which some revisions were made.

cogent ·· social sciences

Table 3. Test statistics: awareness of the main sources of the exam questions			
	Observed value	**Expected value**	**Percent**
Have no idea	2	12	3.3
Never/strongly disagree	30	12	50.0
Seldom/slightly agree	11	12	18.3
Frequently/often agree	8	12	13.3
Always/strongly agree	9	12	15.0
Total	60		100
χ^2	37.50		

Notes: d.f = 4; p = 0.05; χ^2 = 9.49.

5. Data analysis

5.1. Students' questionnaire analysis

The data collected via the questionnaire was analyzed by means of χ^2 test using SPSS 18 and bar-graphs. The transcribed interviews were coded through a comparative analysis as suggested by Glasser and Strauss (1967). The analysis involved a process of repeated shifting throughout the data to distinguish similarities and patterns of reference in the transcripts of the interviews. Analyses of these similarities and patterns gradually led to an evolving coding system for the categories. More specifically, the units of analysis and coding schemes were defined and developed; then the codes were transformed into categorical labels or themes that were repeated or appeared as patterns in the interviews. This procedure, according to Patton (2002), tends to help a researcher with "developing some manageable classification or coding scheme" as "the first step to analysis" (p. 463). Data analysis proceeded incrementally and once the coherence of the data was accomplished, conclusions were drawn based on the analyzed data. Tables 4 and 5 illustrate the major categories and coding schemes along with the thematic categorization which emerged from the data analysis.

5.2. The procedure

To assess the amount of power that different groups own at the levels of developing and adminis-trating the MA TEFL exam in Iran, and to investigated the test-takers' viewpoints, first the question-naire was piloted with the target group. Based on the pilot study, the questionnaire was revised, and the finalized form of the questionnaire was administered. Four groups of students from different universities were asked to fill out the questionnaire. To achieve highly reliable and valid results and to delve into the critical views of the two groups of participants about the TEFL MAEE, semi-struc-tured interviews were conducted. For the students' group, semi-structured interview was applied to cross-check the data collected through the questionnaire. Thus, after a time interval of three weeks, five of the respondents to the questionnaires were required to be interviewed. The interviews took place at the mentioned universities where the study was conducted and approximately 10 min was allocated for each interview. All the interviews were audio-recorded and transcribed verbatim. The transcriptions were then reviewed to come up with the desired information. For the students group,

Table 4. Test statistics: percentage, appropriateness of multiple choice format			
	Observed value	**Expected value**	**Percent**
Have no idea	12	12	20.0
Never/strongly disagree	10	12	16.7
Seldom/slightly agree	22	12	36.7
Frequently/often agree	9	12	15.0
Always/strongly agree	7	12	11.7
Total	60		100
χ^2	11.50		

Notes: d.f = 4; p = 0.05; χ^2 = 9.49.

cogent · social sciences

Table 5. Test statistics: the match between exam questions and learning styles and individual differences

	Observed value	Expected value	Percent
Have no idea	7	12	11.7
Never/strongly disagree	21	12	35.0
Seldom/slightly agree	11	12	18.3
Frequently/often agree	12	12	20.0
Always/strongly agree	9	12	15.0
Total	60		100
χ^2	9.67		

Notes: d.f = 4; $p = 0.05$; $\chi^2 = 9.49$.

the questions mostly mirrored the statements of the questionnaire. However, for the university teachers group, distinctive items were developed.

6. Results

6.1. Students' questionnaire
Research Question 1: Are stakeholders and test developers, who are important parties in the social context of the test, equally powered?

The questionnaire included several items to check the amount of involvement and power of students as one of the main test parties who receive the impact of MA TEFL Entrance Exam. The results suggest that more than 70% of the participants had no information about test makers, more than 63% had no information about the time that the test is made and 50% of the test takers were not aware of the main sources of the exam (Figures 1 and 2, Table 3). Such answers to the questions related to test developers, test sources and the time that the test was made provide nothing but evidence for the participants' unawareness which in turn suggests their lack of power in test development and decisions about the sources.

Based on the results it can be concluded that 63.4% of the participants believed the present testing method of the TEFL MAEE (multiple-choice) was an appropriate one. The results of the chi-square test ($\chi^2 = 11.50$, $p = 0.05$) indicated that the differences in Table 4 were statistically significant. The respondents believed that the present testing method of the exam (i.e. multiple-choice) was an appropriate one. On the other hand, 68.3% agreed with an open-ended format.

As can be viewed from Table 5, over 53% of the participants also agreed that questions should not match the learning styles and individual differences. The results of the chi-square test ($\chi^2 = 9.67$, $p = 0.05$) indicated that the differences were statistically significant. In other words, they implied that if all participants received one test, their rights were respected. Through CLT, it is possible to raise doubts about such contradictions among test parties' views. These results (views regarding the format and equity of the test) show that the method has been practiced so intricately that participants rarely question the policies behind it. CLT also provides a chance to critically question the dominant testing methods, and consider the control that each test party exerts over the high-stakes test of MA TEFL EE.

As illustrated in Table 6 more than 80% of the participants agreed with involving and controlling different parties (students and university teachers) in test development. The results of the chi-square test ($\chi^2 = 22.50$, $p = 0.05$) indicated that the differences in Table 6 were statistically significant. These views contradict what participants have suggested about the testing method (63.4% approved of the present testing method of the TEFL MAEE). It is evident that the percentage of the participants'

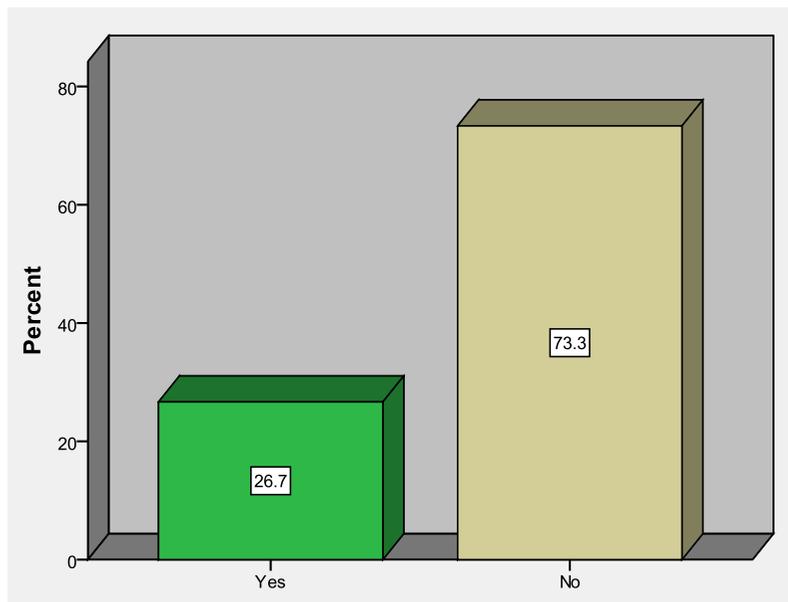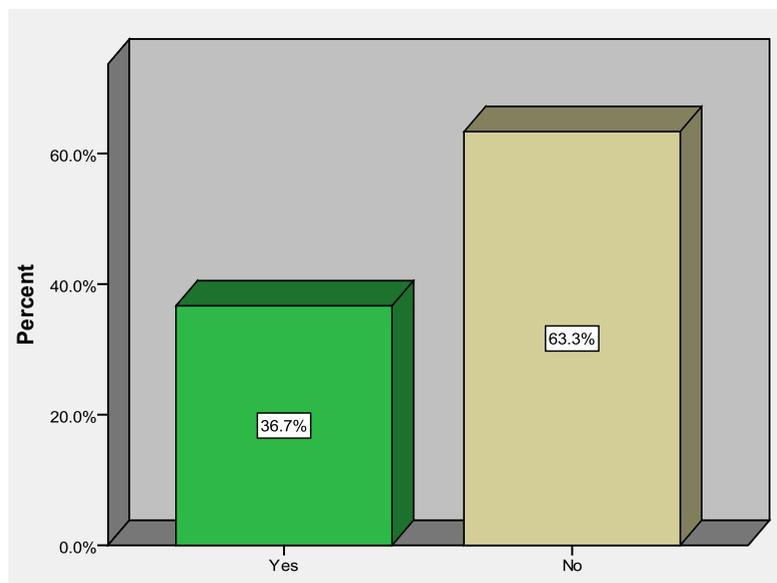**Figure 1. Percentages, information about test makers.**

Figure 2. Percentages, information about the time that the test was made.

approval of democratic practices in testing is well above that of the mainstream method of the TEFL MAEE. Here, other arguments are also possible. It can, for instance, be concluded that they are willing to have hands in test preparation and moderating of the consequences of tests. Democratic assessment makes it possible not to get succumbed to the ruling mainstream system, and protected individuals against unethical practices in language testing (Shohamy, 2001b).

In addition to the above observations, the findings of the study revealed that more than 60% of the participants preferred to be tested by their own university rather than a centralized testing method. The results of the chi-square test ($\chi^2 = 24.17$, $p = 0.05$) indicated that the differences in Table 7 were statistically significant. It could be inferred from this finding that the respondents of this study preferred to be selected by their own professors in their BA courses, as they have access to the main and real sources and materials based on which the nation-wide tests are designed.

Table 6. Test statistics: involvement and control of different test parties (students and university teachers) in test development			
	Observed value	**Expected value**	**Percent**
Have no idea	8	12	13.3
Never/strongly disagree	2	12	3.3
Seldom/slightly agree	9	12	15.0
Frequently/often agree	21	12	35.0
Always/strongly agree	20	12	33.3
Total	60		100
χ^2	22.50		

Notes: d.f = 4, p = 0.05; χ^2 = 9.49.

The fact that 70% of the participants have an overall satisfaction with the test (45% strongly and 25% often agree) may also reveal the position of different groups of test takers. The results of the chi-square test (χ^2 = 30.0, p = 0.05) indicated that the differences in Table 8 were statistically significant. Seemingly, the respondents were convinced that the present method of testing is the best to differentiate among them.

6.2. Open-ended part of the questionnaire
Research Question 2: What are the students' views on adopting the TEFL MA EE as the main requirement for a higher degree?

Questions number 13, 14, and 15 of the questionnaire had an open-ended format to provide greater freedom of expression for the participants to express their viewpoints. These 3 items are successively dealt with below:

Q13: Did the test provide a true and correct reflection and indication of your actual knowledge?

The investigation into the students' responses to this question revealed that the majority of the students admitted that the test did not provide a true and correct reflection and indication of their actual knowledge. The reason underpinning their rejections was that they believed it was impossible to check every person's knowledge in such exams, arguing that the domain of materials is limited to a number of questions. In fact, the participants believed that exams of this kind could only cover a limited number of the required sources for the exam. They believed that a two-hour multiple-choice test could never measure their actual knowledge, given the fact that all participants might suffer from debilitative and detrimental stress, which hinders straight thinking.

Q14: Do you think any aspects of the TEFL MAEE can be improved?

Table 7. Test statistics: preference to be tested by their own university as opposed to a centralized test			
	Observed value	**Expected value**	**Percent**
Have no idea	2	12	3.3
Never/strongly disagree	5	12	8.3
Seldom/slightly agree	16	12	26.7
Frequently/often agree	23	12	38.3
Always/strongly agree	14	12	23.3
Total	60		100
χ^2	24.17		

Notes: d.f = 4; p = 0.05; χ^2 = 9.49.

cogent •• social sciences

Table 8. Test Statistics: overall satisfaction with the test			
	Observed value	**Expected value**	**Percent**
Have no idea	3	12	5.0
Never/strongly disagree	6	12	10.0
Seldom/slightly agree	9	12	15.0
Frequently/often agree	15	12	25.0
Always/strongly agree	27	12	45.0
Total	60		100
χ^2	30.00		

First of all, the majority of the respondents referred to the sources of the test, and believed that through a careful and unidirectional survey and agreement on the main exam materials, many aspects of the exam could be improved. They complained that, especially for the general part of the test, there were not any specific sources for the participants. Other respondents suggested that it would be better if the exam evaluated listening and writing skills as well. They also stipulated that the students' whole performance during their four-year BA program should be considered in the final interpretation of the exam results. The rest of the students declared that the present method could sufficiently serve the purpose of the exam due to the large number of the participants.

Q15: Do you think that the ministry of higher education provided you with thorough information about the circumstances of different universities such as their qualified professors, educational facilities, etc.?

In response to this question, the respondents claimed that they did not receive any information regarding the quality of the universities when they were about to choose among them. Of course, they stated that some universities' websites might have helped them but access to many of the necessary information was impossible. Furthermore, they mentioned that shortcomings made them (i.e. the ministry of higher education) hide realities, they prefer to be mute; otherwise, they would boast about the quality of universities.

6.3. Students' interviews
The questions of the interviews with the student participants are also discussed below:

Q1: Do you have any information about test makers, about the time that the test is made?

Q2: Were you aware of the main sources of the questions?

With regard to the questions concerning the respondents' information about test makers, the time that the test was made and the main sources of the exam, the findings of the data analysis revealed that the study plans they had mapped out in the light of the TEFL MA exam were defectively dubious. Four out of five interviewees pointed to confusion in deciding on the technical and general contents to be covered in the exams as the chief reason why they built their study plans on insecure foundations. For instance, one of the students, Maryam, in this regard contended that "when I was about to take the exam, I had no standard and official sources for the exam, I found them on the internet or collected some sources from my friends who had already passed the exam." Regarding the test makers, the students claimed they had no information, yet they suggested that they might be some of the professors in the state universities. None of the respondents had any information about the time that the test was made.

Q3: Do you agree with the involvement and control of different test parties (students and instructors) in test development?

The results show that the majority of the students agree with the involvement of different parties in the TEFL MA exam administration. As an example, Mohsen maintained that "although involvement of all of students in such a large scale test seems not to be possible I think there must be some kind of match between what students studied during their BA courses and the MA examination." He continued

> For example, during our BA courses, we studied different books on teaching or testing, but for MA, we had to study other sources. At least, university teachers can start teaching the courses they are going to be important for the MA exam. Most of the time, teachers know that a specific book is needed for MA exam, but they offer other sources in BA courses. So, I think, at least, there should be a connection between the students, teachers and test makers. It would save time, energy and …

In line with Mohsen, Razieh asserted that "ministry of higher education should hold a kind of workshop every year and invite students and university professors to come up with their expectations and enjoy their suggestions for further examinations."

Q4: Do you believe that the present testing method of the exam (i.e. multiple-choice) is an appropriate one? Or you think that the exam should also consist of open ended questions?

Four of the interviewees were of the opinion that the current testing method, i.e. multiple-choice, is an appropriate one. They believed that multiple-choice exams can be corrected in a very easy way and can test a large number of applicants.

Q5: Do you think any aspect of the TEFL MA exam can be improved?

Some of the interviewees put forward their ideas on some aspects of the MA TEFL EE that needed to be improved. One of the aspects referred to by the students concerned the necessity of assessing all skills in the exam. "In fact, this test has to test all skills, but which one of them is involved? Just reading and the components of the language such as grammar and vocabulary. Listening and writing are left out. They did not exactly test the students' proficiency, it is based on memorization and reading skill" (stated by Leila).

Q6: Did you prefer to be tested by your own university as opposed to a centralized testing method?

In response to this question, Javad answered: "no, I didn't prefer to be tested by my own university because I wanted to change the place of my study for higher education in order to gain new insights from new professors with new methods and new styles." In his argument on the same issue, Mohsen stated that:

> In a large scale test like TEFL MAEE, the best method to assess the applicants is a centralized method, because university professors may be biased, they may prefer to choose their own students. So, I myself preferred to be selected through the centralized testing method.

Another respondent, Leila, asserted that: "My own university will increase my chance to get to higher levels of success in my field, because when I continue my education in my own university, I know the test developers who are my own teachers. In fact, I know my teachers' favorite areas and try to study those areas more carefully."

6.4. Professors' interviews

Apart from the remarks of the students, the viewpoints of the university professors were also scrutinized. The questions are dealt with below:

Q1: In the curriculum development by MST (ministry of science and technology), are university teachers' opinions and ideas considered?

The investigation into the professors' responses to this question revealed that even if authorities ask for the university teachers' opinions and ideas, they do not include or apply them. Therefore, it can be concluded that some scholars and university professors have certainly been consulted. However, information such as who they are, what their interests and affiliations are, where they work, and whether they are qualified or is not revealed and maybe open to question. The authorities prefer not to incorporate different professors' opinions unless a catastrophe happens and necessitates reforms!

Q2: What aspects of the test need amelioration?

Professors claimed that the nuclear essence of testing humans by humans, the unitary nature of the judgments and lack of interaction throughout the assessment process need modification. The current MA exam, in the instructors' view, has made the applicants develop exam-oriented strategies of memorization and comprehension rather than analysis, evaluation, and production strategies. "Most of the students believe that they only need to be test-wise to pass the exam and it seems that possessing abilities, such as logical analysis of subject matters, is not necessary to take the TEFL MAEE successfully"(stated by one of the instructors). The other professor mentioned that "rethinking in developing the curriculum and textbooks, which are considered as references for preparation for the MA entrance exam, is necessary as most of the references of the MA EE are not valid sources and are useless. When you are admitting a student according to that invalid source, it is not a valid judgment."

Q3: If you want to develop a model of the exam, what would be your suggestion?

The professors agreed that as a test designer, they should do a kind of survey and ask teachers what they teach in BA courses and then the dominant syllabus should be picked up and the test designer should construct the test items based on the most frequently used topics, syllabi and documents in different TEFL courses. They mentioned that their favorite model is an interactive, research-based one based on the needs of the test-takers.

Q4: To what extent do you think the MA entrance exam influences your instruction?

The professors' answers to this question revealed that they had no tendency toward adapting their selected course book to the materials used by test developers in constructing the TEFL MAEE test. In fact, they stated that, as they almost have no control over the sources of the test and the materials are chosen by a specific group of the administrators, they do not take into the account the content of the TEFLE MA EE in designing their syllabi. Likewise, they do not have enough time to dedicate their time to cover the exam-related points during the term in the classroom. Therefore, it can be concluded that the format and content of the TEFL MAEE questions do not influence their instruction in BA courses.

7. Discussion

Through analysis of the data, it was found that the stakeholders and test developers, as important parties in the social context of the test have no power in the different phases of administering of the exam. In line with Khany and Tarlani-Aliabadi (2016), students are basically viewed as the passive and powerless recipients of predetermined policies in Iranian educational system. They are not

given enough autonomy to express their reactions and voices against the key educational aspects of curriculum development and assessment design (Khany & Tarlani-Aliabadi, 2016).

The findings contradict with what (Shohamy, 2001a, 2007) recommend about the democratic assessments. According to democratic assessment, the professional tester serves mostly as a facilitator who assists other parties in the strategies of collecting the information and its interpretation (Shohamy, 2007).

The second research question tried to investigate students' views regarding the adoption of the TEFL MA exam as one of their higher degree requirements. As can be viewed from the results over 53% of the participants agreed that questions should not match the learning styles and individual differences and 63.4% of them agreed that the current method of examination (multiple-choice) is an appropriate one but on the other hand, 68.3% agreed with an open-ended format. In other words, they have implied that if all the participants received one test, their rights would be respected. But they demanded another form of the examination which is the open-ended format. Their agreements with the current multiple- choice method of the exam go well with what Shohamy (2001b, 2007) suggested that some practices may evolve into ideologies as test receivers do not challenge the ruling method of evaluation, or they are manipulated to do so. So they are persuaded that the current method of the exam can differentiate among them. Also their demands for adding the open-ended section to the current format of the exam raise doubt about contradictions among test parties' views. These results revealed that the method has been accepted to some extent that test takers rarely question the policies behind it.

Based on the results, more than 80% of participants agreed with involving and controlling different parties (students and university teachers) in test development. These views also contradict with what participants have suggested about the testing method. It can be concluded that they were willing to participate in test preparation.

The findings of the study are also in accord with Shohamy (2001b) in that she encouraged test takers to examine tests critically and stated that test takers should be permitted to question the test itself, its value and its methods.

Regarding the washback effect of the TEFL MAEE on the professors' performance in undergraduate courses the results of the professors' interviews revealed that they are not affected by Iranian MA Entrance Examination as the high stakes test. They use their own methods and never teach according to the issues related to the MA exam. The results of this study contradict with those of Mohammadi (2007) who conducted a survey on the effects of MA Entrance Examination on the teaching practices and reported signs of washback effect for the respondents. According to the finding of this study the majority of the university professors teach according to their own preferable sequence of importance. They do not spend their class time teaching tips and tricks of the exam. The point which is worthy of note is that the majority of the professors stated that the authorities prefer not to involve different professors' opinions in developing the test, so this can be the reason why the majority of the professors do not have any plan to prepare the students for such tests. These results do not coincide with those of Shohamy (2007b) who asserted that interaction between test constructors and teachers should be boosted and teachers should be trained in assessment measures. Professors' views and opinions in different phases of the test developing should be considered. As Al-Amri (2010) states, teachers who make a serious commitment to using criteria, training in the use of scales, and identifying and reflecting on their previous experiences, power, control and subjective decisions can plausibly anticipate more valid, reliable and practical assessments.

8. Conclusion

The results of the present study revealed that the agenda of the TEFL MA entrance examination are almost unknown to test takers; the test is mainly created by politically and educationally empowered parties. Considering the fact that the major group of stakeholders- i.e. the students—have no control over the content and the time of the administration of the test, it can be inferred that power is not equally shared among all test parties. The results generated from the data analyses indicated that, from the students' viewpoint, the testing method should consider all participants' expectations and invite their views via inclusion of different test formats. They believed that along with the multiple-choice format, the test should consist of open-ended questions as well.

According to the respondents, the multiple-choice format cannot be an appropriate method to test different students with different styles and individual differences. They also maintained that, if all participants receive one similar test, their rights are respected. Additionally, a large number of participants agreed that parties (including the students and university teachers) should be involved and empowered in the process of test development. This view contradicts what participants have suggested about the testing method. In other words, while they are willing to take part in a democratic test, they do not believe in the possibility of this event.

The observations made in the current study also evidenced that most of the participants prefer to be tested by their own university, rather than a centralized testing method. They assumed that being admitted by their own university gives them the chance to be aware of test generators and their favorite areas and thus they try to study those areas more carefully. They also maintained that the exam requires the applicants to limit themselves to comprehending and memorizing details instead of gaining deeper insights to produce extended answers. Regarding the improvement of the exam, the participants stated that this test should test all skills, while in the current method of testing, just reading and the components of language such as grammar and vocabulary are included. As a result, evidently, this test cannot exactly test the students' proficiency as it is based on memorization and reading skills. With regard to the professors' viewpoints about the amount of their power in administrating the test, an interesting point made was that even if the authorities ask for their advice about the framework of the test, their opinion will not be included or applied. The point which attracts attention, with regard to the aspects of the test needing amelioration, the majority of the professors believed that the current MA exam encourage exam-oriented strategies of memorization and comprehension rather than analysis, evaluation and production strategies. They also mentioned that reconsidering the policies followed for developing valid sources for the exam is necessary. Being asked for their suggestive model for the exam, most of the professors agreed with an interactive, researcher-based and needs-based model. The findings of the survey also indicate that university professors are not affected by the TEFL MAEE as a high-stake test. They are using their own methods and techniques appropriate for their own final exam. They teach according to their own preferable sequence of importance. They do not spend their class time teaching tips and tricks related the MA exam. This survey found that the majority of the instructors are not examination-oriented. In fact, the TEFL MA UEE was found to have a negligible effect on the professors' academic behavior.

8.1. Implications

The results of this study may have useful implications for teachers, test developers and researchers. Implications of this study are manifold: first and foremost, this study is intended to reveal the extent and the nature of washback effect of the state TEFL MA examinations on the teaching and learning procedures of the state undergraduate level, hence one relevant upshot of the study would be that the state TEFL undergraduate level is critically investigated as to whether or not it is guided in the very direction of the TEFL MA university entrance examination.

This study, thus would weigh out the pros and cons of the present educational policy within state universities at TEFL BA level. Consequently it would bring into light the present undergraduate program deficiencies as well as providing some useful insights into the areas which are in urgent need of essential revision both in the realm of teaching and TEFL MA UEE.

cogent ·· social sciences

The study is also likely to convince the test developers to manipulate the content of the TEFL MA UEE by adopting some analytical or critical approaches to eschew the peril of students growing increasingly test wise. More specifically teachers and professors are pit into an acid test whether or not they are providing the students with the proper instruction if the undergraduate students' major objective is to sail through the their imminent TEFL MA examination.

Author details
Mohammad Salehi[1]
E-mail: m_salehi@sharif.ir
M. Tarjoman[2]
E-mail: tarjomanm@gmail.com
[1] Sharif University of Technology, Iran.
[2] Department of English Language, Islamic Azad University, South Tehran Branch, Tehran, Iran.

Citation information
Cite this article as: An investigation of a nationwide exam from a critical language testing perspective, Mohammad Salehi & M. Tarjoman, *Cogent Social Sciences* (2017), 3: 1396639.

References
Al-Amri, M. (2010). Direct spoken English testing: Is still a real challenge to be worth bothering about. *English Language Teaching, 3*(1). Retrieved May12, 2012 from www.ccsenet.org/journal/index.php/article/download/.../4339

Alderson, J. (1986) Innovations in language testing. In M. Portal (Ed.), *Innovations in language testing: Proceedings of the IUS/NFER conference* (pp. 93–105), Windsor: NFER-Nelson.

Alderson, J., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*, 115–129. https://doi.org/10.1093/applin/14.2.115

Anderson, G., & Arsenault, N. (1998). *Fundamentals of educational research* (2nd ed.). London: Falmer Press.

Ary, D., Jacobs, L. C., & Razavieh, A. (2002). *Introduction to research in education* (6th ed.). Belmont, CA: Wadsworth.

Auerbach, E. (1995). The politics of the ESL classroom: Issues of power in pedagogical choices. In J. Tollefson (Ed.), *Power and inequality in language education* (pp. 933–956). Cambridge: CUP.

Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing, 13*(3), 257–279. https://doi.org/10.1177/026553229601300303

Best, J. W., & Kahn, J. V. (2003). *Research in education* (9th ed.). Boston, MA: Allyn and Bacon.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4), 653–675. https://doi.org/10.2307/3587999

Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3–17). Mahwah, NJ: Lawrence Erlbaum Associates.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*, 438–481. https://doi.org/10.3102/00346543058004438

Farhady, H. (2006). *Twenty-five years of living with applied linguistics: Collection of articles*. Tehran: Rahnama Press.

Foucault, M. (1975). *Discipline and punish: The birth of the prison*. New York, NY: Vintage Books.

Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings*. New York, NY: Pantheon Books.

Glasser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. New York, NY: Aldine.

Holstein, J. A., & Gubrium, J. F. (1995). *The active interview*. Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781412986120

Kellaghan, T., & Greaney, V. (1992). *Using examinations to improve education: A study in fourteen African countries*. Washington, DC: World Bank.

Khany, R., & Tarlani-Aliabadi, H. (2016). Studying power relations in an academic setting: Teachers' and students' perceptions of EAP classes in Iran. *Journal of English for Academic Purposes, 21*, 72–85. https://doi.org/10.1016/j.jeap.2015.12.002

Madaus, G. (1991). *Current trends in testing in the USA*. Paper presented at the conference on Testing and Evaluation, Feedback Strategies for Improvement of Foreign Language Learning, Washington DC.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*(3), 241–256. https://doi.org/10.1177/026553229601300302

Mohammadi, M. (2007). High-stakes testing washback: A survey on the effect of Iranian MA entrance examination on teaching. Retrieved October 10, 2011, from http://faculty.ksu.edu.sa/aljarf/documents/english%20language%20teaching%20conference%20-%20iran%202008/mojtaba%20mohammadi.pdf

Noam, G. (1996). Assessment at a crossroads: Conversation. *Harvard Educational Review, 66*(3), 631–657.

Noble, A. J., & Smith, M. L. (1994). *Measurement-driven reform: Research on policy, practice, repercussion* (Tech. Rep. 381). Tempe, AZ: Arizona State University, Center for the Study of Evaluation.

Patton, M. Q. (2002). *Designing qualitative studies. Qualitative research and evaluation methods* (3rd ed.). Thousands Oaks, CA: Sage.

Pennycook, A. (2001). *Critical applied linguistics*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rea-Dickins, P. (1997). So why we need relationship with stakeholders in language testing? A view from UK. *Language Testing, 14*(3), 304–314.

Rezaee, A. A., & Salehi, M. (2008). The construct validity of a language proficiency test: A multitrait multimethod approach. *TELL, 2*(8), 93–110.

Rubin, H. J., & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage.

Salehi, M. (2011). Test taking strategies: Implications for validation. *Journal of Language Teaching and Research, 2*(4), 850–858.

Shohamy, E. (1992). Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal, 76*(4), 514–521.

Shohamy, E. (2001a). Democratic assessment as an alternative. *Language Testing, 18*(4), 373–391. https://doi.org/10.1177/026553220101800404

Shohamy, E. (2001b). *The power of tests: A critical perspective on the uses of language tests*. Essex: Longman.

Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. London: Routledge. https://doi.org/10.4324/9780203387962

Shohamy, E. (2007). The power of language tests: The power of the English language and the role of ELT. In J. Cummins & C. Davison (Eds.), *International handbook of english language teaching* (pp. 521–531). New York, NY: Springer. https://doi.org/10.1007/978-0-387-46301-8

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com