On the Authenticity of IELTS Academic Tests

Goodarz Shakibaei

Ahvaz Branch, Islamic Azad University

Student93ma@gmail.com

Abstract

The purpose of this paper was to investigate the authenticity of IELTS academic tests. Bachman and Palmer (1996, p. 23) define authenticity as "the degree of correspondence of the characteristics of a given language test task to the features of a target language task". Authenticity is then an important aspect of testing since it describes the relationship between the test and the real world. The participants of this study included 80 Iranian IELTS test takers who took IELTS academic test in 2014 and 2015, and 100 lecturers of university and IELTS teachers. Twenty IELTS test takers also were interviewed to collect the required qualitative part of data. The data were gathered through a semi-structured interview and authenticity checklist. Descriptive statistics used to analyze the collected data. The findings of the study showed that IELTS academic tests are not authentic in terms of tone, format, channel, form and scope of interaction. Moreover, mostly the interviewees stated that answering the items in IELTS academic test requires specific techniques and test takers can not succeed without mastering these techniques.

Keywords: Authenticity, IELT Academic Tests, Real Life, Test Takers

1. Introduction

One of major principles of language testing is authenticity, a concept that is a little slippery to define, especially within the art and science of evaluating and designing tests. Beckman and Palmer (1996, p. 23) define authenticity as "the degree of correspondence of the characteristics of a given language test task to the features of a target language task" and then suggest an agenda for identifying those target language tasks and for transforming then valid test items. Authenticity is defined in different ways. The authenticity of test tasks in recent years has increased noticeably. Two or three decades ago, unconnected, contrived items were accepted as a necessary component of testing. Things have changed. It was supposed that language testing scale could encompass the performance of the productive skills and writing components. Reading passages are selected from real-life sources that test-takers are likely to have encountered or will encounter.

The possible meaning of authenticity can be as follows:

a. Authenticity relates to the language produced by native speakers in a particular language community (Porter & Roberts, 1981; Little & Singleton, 1989).

b. Authenticity relates to the language produced by a real speaker or writer for a real audience, conveying a real message (Morrow, 1977; Porter & Roberts, 1981; Swaffar 1985, Nunan, 1988; Benson &Voller, 1997).

c. Authenticity relates to the interaction between students and teachers and is a "personal process of management" (Van Lier, 1996, p. 128).

d. The degree of correspondence of a given language test task to the features of a TLU task (Bachman & Palmer, 1996, p. 23-29).

e. A language test is said to be authentic when it mirrors as exactly as possible the content and skills under test (Davis, 1999, p. 13).

f. Authenticity relates to assessment (Bachman, 1991).

Authenticity is then an important aspect of testing since it describes the relationship between the test and the real world. The notion of authenticity was developed first within applied linguistics in the 1960. The debate mainly focused on the nature of authentic input until 1991, when Bachman suggested that two types of authenticity had to be realized: situational and international authenticity (Lewkowicz, 2000, p. 48). While situational authenticity was considered with the relationship between test task and the criterion behavior, interactional authenticity indicated the interaction between test task and test user (Authenticity in language assessment by: Eder, 2010).

Based on Blackman, classic fundamental considerations in language testing, an authentic language test is on which indicates how similar a test takers performance is to what he is likely to do in a real life situation .In other words, the more similar a test task is to what happens in real life, the more authentic it will be.

Despite the significance of authenticity, only a few studies have been carried out to investigate IELTS academic tests in terms of authenticity. To investigate this quality of IELTS academic tests, the following question was formulated and answered in this study.

-To what extent are the contents and tasks of the current IELTS academic tests practiced in Iran authentically representative of the tasks in target language situation"?

2. Review of Literature

According to Guariento and Morley (2001), one of the most important characteristics of tests is authenticity that has been vital important in language teaching and language teacher education for almost four decades.

The issue of authenticity has always been an important part of any discussion on language testing. There is an argument that only real language use should be counted when assessing language ability. Authenticity, in early communicative texts, meant the use of material in prompts that had not been written for non-native speakers of English, and a test could not be communicative unless it was authentic (Fulcher, 2000). Alderson (1981) terms authenticity a "sterile argument" (p. 48). Modern performance tests that attempt to mirror some criterion situation in the external world are no more than role-plays or simulations, in which the learner is asked to 'imagine' that they are actually taking a patient's details in a hospital, giving students a mini-lecture, or engaging is a business negotiation.

Language tests by their very nature are not mirrors of real life, but instruments constructed on the basis of a theory of the nature of language, of language use, and of language learning (Fulcher, 1999; Widdowson, 1978) making a distinction between 'genuineness' and 'authenticity' of language testing, argues that "genuineness is a characteristic of the passage itself and is an absolute quality. Authenticity is a characteristic of the relationship between the passage and the reader and has to do with appropriate response"(p. 80).

 Authenticity can also be defined as the degree of correspondence of the characteristics of a given language test task to the features of a target language use task. Bachman and Palmer (1996) show the relationship as following:

Characteristics of the TLU ↔ characteristics of the test task

They also consider the authenticity as a critical aspect of language test usefulness alongside validity, reliability, impact, interactiveness, and practicality. Accordingly, Douglas (2000) suggests employing the dual notion of authenticity in specific language testing. He believes that in EAP test development, what we must do is first to explain a target language use situation

in terms of features of content and task; we must then clarify how these characteristics will be recognized in the test in order to involve the test takers in the test tasks, performance on which can be interpreted as evidence of communicative language ability with reference to the target language situation.

Since authenticity is an important concept in general language purpose testing and in specific purpose language testing (Bachman& Palmer, 1996; Douglas, 2000), it is necessary to consider it as an important component of language testing development.

Birjandi, and Ahmadi (2013) reviewed different aspects of authenticity and mainly focused on a discussion of a social-constructivist approach to authenticity in second language assessment. They discussed that authenticity in second language assessment defined in terms of the interaction between the test taker and the input is a socially constructed product influenced by identified five factors namely test tasks, test constructors, test administrators, test takers, as well as contexts that   influence the nature of authenticity defined in terms of interaction.

Shomoossi and Tavakoli (2010) studied authenticity and authentication in language testing. They concluded that the authenticity as a test feature might be sidelined by a more important notion in test validity, i.e., the authentication process. In other words, the theoretical model of test authenticity (Bachman & Palmer, 1996) may need a dynamic and operational model to include the human elements and all steps of designing authentic test tasks, which may not be taken into account as the end of the testing process. They proposed a rough sketch for the authentication process as the operational side to Bachman and Palmer's (1996) model of authenticity, which consisted of language and TLU domain researchers, policy makers, teachers, testers, material developers and context.

Yuk and Lee (1995) investigated text authenticity and learner authenticity. A distinction was made between text authenticity and learner authenticity. The researcher suggested that there is no single model or framework which will fit all learning situations equally well and we need to be flexible and sensitive to the characteristics of learners if we want our materials to be learner authentic.

Ingram (2003) studied authenticity in language testing. He briefly discussed the progress in language testing (especially proficiency assessment) towards more authenticity and pointed to the variables that make the gap between language testing and real-life use of language difficult to bridge. The paper finally discussed a new approach to proficiency assessment, which is both globally applicable and much closer to focusing on real-life, authentic language performance than previous approaches.

The researcher concluded, to some extent, that gap will always exist because language is only a part of any activity and a person's performance in that activity is influenced to a greater or lesser degree by other variables. Nevertheless, since language is situation-dependent, the more the tasks and contexts in which the language was tested resemble those of real-life, the more accurately is the language test likely to forecast how the candidate will cope, at least linguistically, with real-life activities.

Abdi and Asadi (2013) investigated the authentic materials and their advantages and disadvantages. They explained how authentic materials could be effectively used in the classroom. They concluded that authentic texts increase students motivation when reading in, English and provides opportunities for them to reduce important cultural and lexical information when used effectively, authentic materials help bring the real world into the classroom and significantly enliven the class and students learn how to

disregard what is not relevant and get confidence in being able to function in an English-speaking society.

Khodadady and Hashemi (2011) examined the relationship between text authenticity and test takers' performance on C-Tests. They developed single C-test on an authentic text and administered it to one hundred and thirty five junior and senior undergraduate university students majoring in English language and literature, English translation and teaching English as a foreign language in two universities in Iran. They came to conclusion to that although high correlation coefficients obtained between the authentic and standard CTests (r=0.87, p<.01) showed that they could be used interchangeably, the AC-Test enjoyed slightly higher internal validity and noticeably higher reliability.

Vosoughi (2013) also examined authenticity factors in reading comprehension achievement tests in Iranian high schools through Co- matrix2. The results showed the optimum case for only two measures of Type-token ratio and Connectives. The two measures of Latent Semantic Analysis and Frequency of Content Words were low indicating immediate attention by the language testers.

Liu (2005) studied the methods for assessing authenticity in computer-assisted language learning and assessment. The results indicated that authenticity has become a multi-componential theory-based concept, and therefore investigation of the authenticity of a task needs to be based on analyses from multiple perspectives, consisting of the target language use situation, test/task characteristics, learner perception, learner involvement, and language production.

Lewkowicz (2000) studied authenticity in language testing in University of Hong Kong and focused on some outstanding questions. He focused on some questions as the followings:

1. Which characteristics are critical for distinguishing authentic from non-authentic test tasks?

2. Are some of these characteristics more critical than others?

3. What degree of correspondence is needed for test tasks and TLU tasks to be received as authentic?

4. Can critical characteristics be identified for all tests, that is, general purposes as well as specific purpose language tests?

5. Does perception of authenticity affect test tasks' performance and, if so, to what extent?

Findings of the study showed that test takers' perceptions of authenticity vary and for some individuals, authenticity is an important aspect of the test. Some factors other than the correspondence between test and TLU tasks must have affected students' perceptions of the test tasks and authenticity of materials only constitutes one set of conditions for authenticity in teaching classroom. This study also indicated the relationship between perception and performance. a range of conditions interacting with the test tasks will affect test takers' perceptions and help to determine whether test tasks are considered authentic or not.

Luisa and Esteban (2005) investigated the authenticity in teaching of ESP. They put forward a proposal and a checklist for evaluation of authenticity to ease the task of selecting and using authentic materials in the field of ESP. They concluded that although authentic materials are very rich source for the selection of teaching materials, the selection and use of this type of material asks the teacher for a thorough appraisal and authenticity in teaching-learning

situation should be based on the correspondence between our students' learning and target needs.

Berardo (2006) discussed the use of authentic material in the teaching of reading. They pointed thatthe student benefit from the exposure to real language being used in a real context. Moreover, they believed that when using authentic materials, the students are highly motivating and giving a sense of achievement when understood and encourage further reading.

In another study, Morore and Morton (2007) ran a research in Australia. They did a comparative study and investigated authenticity in the IELTS academic module writing of task 2 items in IELTS test and university assignments. They selected 155 assignment tasks from a range of undergraduate and postgraduate courses and then compared with a corpus of 20 IELTS tasks 2 items. They chose the tasks and assignment from a range of students at two universities in Australia, namely Monash and the University of Melbourne. They concluded that optimizing authenticity of IELTS with its increasing use at the university as an entrance instrument and its corresponding influence on programs of English for academic purposes is especially pressing and some constrains should be taken into account when we want to optimize the authenticity of writing tasks in IELTS test.

Moreover, Atai and Soleimani (2008) investigated the effect of text authenticity and genre on EFL learners' performance in C-tests. The participants were 120 senior students majoring in English translation. The findings make it clear that authenticity and genre can be considered as "specific aspects of test method". However, the interaction between authenticity and genre did not prove to be significant.

Charles and Stansfield (2008) described a working model used to determine the Target Language Use (TLU) (Bachman & Palmer, 1996) in a

Language for Specific Purpose (LSP) test project: The Listening Summary Translation Exam in Taiwanese (LSTE/T). The discussion focused on the authenticity of task (Douglas, 2000) and how it can be achieved by the development of a taxonomy and a structured verification procedure in the process. The authors suggested that, as one type of LSP testing, job-relevant tests are becoming increasingly important and deserve more attention from practitioners in the testing field.

They concluded that first; verification of authenticity is an ongoing process that should exist in the whole course of the project, not something that is done only at the beginning of the project. Secondly, effective verification of authenticity is interactive. Thirdly, the use of secondary resources is unavoidable in situations where primary resources are not available, such as verification of topics related to FCI work in this project. Ultimately, we are not sure that LSP tests are much different from tests of general language proficiency. However, it is clear that if they are to be considered as valid, the authenticity of their language and tasks must be verified.

Spolsky (2008) explained the limits of authenticity in language testing. The author suggested that the criterion of authenticity raises important pragmatic and ethical questions in language testing. He also believed that lack of authenticity in the material used in a test raises issues about the generalizability of results. Any language test is by its very nature inauthentic, abnormal language behavior, for the test taker is being asked not to answer a question giving information but to display knowledge or skill.

Atai and Soleimany (2009) focused on the effect of text authenticity and genre on EFL learners' performance in C-Tests. The researchers focused on the nature of input and expected response by considering genre and text authenticity in C-tests. The findings of the study revealed a significant effect of

genre and text authenticity on the C-test performances of the participants. However, the interaction between genre and text authenticity did not prove to be significant.

Sabet and Mahsefat (2012) focused on the Impact of authentic listening materials on elementary EFL learners' listening skill at university level. The findings of this study indicated that students who were exposed to authentic materials did better in posttest. The analysis of feedback survey also showed their satisfaction and positive attitudes to authentic listening materials.

Baleghirzadeh (2002) did a comparative study between the test of spoken English and the IELTS interview with respect to their authenticity. He investigated the nature of authenticity in two interviews: the interview section of TOEFL and the speaking skill of the IELTS. Bachman and Palmer's (1996) framework of language task characteristics was used to answer the questions. The result of the study revealed that the IELTS interview is more authentic than testing of spoken English (TSE) in terms of test characteristics. According to this study, the features of test tasks in the IELTS interview are more lifelike and hence this exam is more authentic in terms of its test task characteristics.

In sum, the review of literature on IELTS academic tests has revealed that the authenticity quality of IELTS academic tests has been reviewed thoroughly. The review focused on this quality of IELTS academic test investigated in recent studies and the gap of the research has been indicated that the present study could be significantly provided the IELTS academic practitioners with invaluable information and applicable findings.

3. Methodology

3.1. Participants of the Study

To carry out the present study three groups of the participants were needed. The first group included participants who received authenticity checklist. A total number of 80 IELTS academic test takers from different provinces of Iran received the authenticity checklist questionnaire. The completed questionnaires were used to investigate the authenticity quality of IELTS academic tests from IELTS test takers point of view. The participants majored in different fields such as biology, civil engineering, medicine, chemistry, etc. The participants were identified through contacts of the IELTS teachers from various IELTS training institutes through contacts in Iran. I also conducted any educational institutions whose online publicity indicated that they prepared learners for IELTS exam. The participants were all volunteers.

The participants of the study will be a representative sample of Iranian IELTS test takers, i.e., the language learners who took IELTS academic test. These IELTS academic candidates usually take this version of IELTS test to enter international universities and enroll in other institutions of higher education of professionals such as medical doctors and nurses who want to study or practice in an English-speaking country in which the median of instruction is not in their first language, that is, the courses are in English. Admission to undergraduate and postgraduate courses is based on the results of the IELTS academic test. The participants of the second group of study were the teachers who have been teaching IELTS courses at different language institutions in different cities in Iran. 100 lecturers of university and IELTS teachers who were teaching IELTS courses, i.e. reading, writing, speaking and listening, at English language institutes and universities in Iran participated in this study as the second group of the present study.

The collected data were used to investigate the authenticity from English teachers' point of view. The authenticity checklist questionnaires were given to the IELTS teachers in English language. The participants of this group were also volunteers and the teachers who were identified through colleagues at the universities and educational institutes that work on IELTS programs and the main aim of these institutes is to prepare students for international and national language examinations such as IELTS tests. The third group of participants included 20 interviewees that used to gather the qualitative data of the study. These participants were selected from the first group of participants of study. Totally, 180 participants participated in this study for collecting the required qualitative and quantitative data.

3.2. Instrumentation

Two types of instruments were used to collect the required data for the present study. These instruments include the authenticity checklist questionnaire and interview. They are presented in details as the followings:

3.2.1. Authenticity Checklist

Authenticity framework proposed by Bachman and Palmer (1996) and applied by Douglas (2000) was usedto develop a checklist to identify the authenticity of IELTS academic tests practiced in Iran.The checklist was then translated into Persian. The reliability of translation was checked through backward translation technique. This instrument included 16 items, which were made based onfour main parts of Bachman and Palmers' model. The model consists of 1.characteristics of test rubrics (3 items), 2. input characteristics (5 items), 3. output characteristics (5 items), and 4. the interaction between input and

response (3 items). Participants' perceptions of situational authenticity was assessed on a 5-point Likert scale ranging from (1=strongly disagree) to (5= strongly agree). The reliability of this part was above.9. Construct validity was also estimated through principal components factor analysis. Principal component analysis is a variable reduction procedure. It is useful when one has obtained data on a number of variables (possibly a large number of variables), and believes that there is some redundancy in those variables. In this case, redundancy means that some of the variables are correlated with one another, possibly because they are measuring the same construct. Because of this redundancy, one believes that it should be possible to reduce the observed variables into a smaller number of principal components (artificial variables) that will account for most of the variance in the observed variables.

As each instrument consisted of a group of variables measured by different items, principal component analysis was run to determine the items, which may constitute one factor and delete the items that do not have significant loading factor. Another reason for preferring principal component analysis to principal axis factoring was that there was no significant difference between the loading factors of the items measured through these two methods. In addition, the assumptions of factor analysis, KMO test, for this instrument (as well as the other instruments) were checked. The results of factor analysis indicated that the initial Eigen values for each component was above 5.3 and the loading of each factor was above .7 . A brief look at the loadings shows that almost all of the loadings were high enough to conclude that all items of each component constitute one factor(Authenticity checklist is presented in appendix).

3.2.2. Cued-Recall Interviews

To evaluate the authenticity of IELTS tests and to know about the IELTS test takers' perception about IELTS test skills, i.e. writing, reading, speaking and listening, 20 participants were interviewed. The participants were IELTS test takers from different cities of Iran. They studied in different fields at universities of Iran. I also asked them about their views about IELTS test in general. That is, I asked them seven questions about their views about IELTS and the four skills and then I asked them follow-up questions to explore more about IELTS components, the relationship between the four skills.

3.3. Procedure

Since qualitative and quantitative data were needed to answer the formulated questions in this study, a specific procedure was needed. As a first step, the needed questionnaires were administered to the participants either directly by the researcher or through some colleagues and some were e-mailed to them. After collecting the authenticity checklist, they were analyzed by the researcher and the scores of each participant on each of the measure were calculated. Each questionnaire was given a numerical code to know the respective associated score on each of the measures. After eliciting the scores from the measures, they were entered into SPSS data sheet. In order to answer quantitative part of the research question about the authenticity of IELTS academic test, descriptive statistics of the participants including frequencies, percentage, and mean scores were calculated of the data elicited form authenticity checklist. These questionnaires were completed by 100 IELTS test takers and 110 English teachers who have been teaching English for at least 5 years.

Then, the qualitative data needed for the present study were collected. Data were gathered during face-to-face in-depth interviews. The participants were informed of the purpose of the research and their consent was obtained. The researchers also obtained the participants' permission to audiotape each interview for purposes of content analysis and audit trail. The interviews were conducted in both an unstructured and a semi-structured manner. The interviews lasted on average for about 25 minutes. Interviewing took place during all days over a five-month period, until the data collected were being consistently duplicated. No new information was gained from the last three interviews, thus data saturation was considered to have been achieved. The interview data were immediately transcribed verbatim and analyzed using qualitative content analysis. In this interview the researcher focused of the relationship between four skills, number of items, giving a lecture in International conferences in English, project reporting by IELTS test takers in English, note taking while listening to a lecture in English and understanding English news on T.V and radio. The views were recorded and transcribed and then the researcher did content analysis.

3.4. Data Analysis

3.4.1. Authenticity of IELTS Academic Tests

The question of the present study dealt with investigating the authenticity of the IELTS academic tests administered in Iran. More specifically, the first question was: "to what extent are the contents and tasks of the current IELTS academic tests practiced in Iranian authentically representative of the tasks in target language situation?

3.4.2. Authenticity of IELTS Academic Tests from Test takers' Point of View

The results of this section of the study are presented in the following parts.

Situational authenticity: The required data to examine the situational authenticity of IELTS academic tests were collected through authenticity checklist. This checklist consists of four main parts: characteristics of test rubrics (3 items), input characteristics (5 items), output characteristics (5 items), and the interaction between input and response (3 items). Participants' perceptions of situational authenticity was assessed on a 5-point Likert scale ranging from (1= strongly agree) to (5= strongly disagree).

Rubric characteristics: To determine situational authenticity of IELTS academic tests, the checklist proposed by Bachman and Palmer (1996) defining the English teachers' perceptions of IELTS academic tests rubric (objectives, responding procedure, and number of test tasks), was distributed among 110 English language teachers. The descriptive statistics of the participants' responses to the items measuring their perceptions of the test rubrics are shown in the following table.

Table 3.1. *The Participants' Perceptions of Rubric Characteristics of IELTS Academic Tests*

Items	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1	12%	4%	4%	56%	24%
2	4%	40%	28%	24%	4%
3	0%	32%	40%	28%	0%

Table 4.1.shows that 56% of the participants disagreed and 24% of the participants strongly disagree that the objectives, 24% of the participants disagree and 40% of the participants agree that responding procedure and 28% of the participants disagree and 32% of the participants agree that the number of tasks in TLU situation and IELTS academic tests tasks are the same. The

findings also indicate that 12% of the participants strongly disagreed that the objectives in TLU situations and IELTS academic tests are the same and about 4% of the participants strongly disagree that responding procedure in TLU situation and IELTS academic tests tasks is the same.

Input authenticity: Five questions on tone, format, vehicle of delivery, and language length of TLU situation tasks and IELTS academic test tasks and correspondence between TLU and IELTS academic test tasks were asked to identify input authenticity of IELTS academic test tasks. The participants were invited to check the options based on their perceptions of the characteristics of the tasks of IELTS academic tests tasks and TLU situation tasks.

Table 3.2. *The Participants' Perceptions of IELTS academic Tests Input Authenticity*

Items	Very weak	Weak	Average	Good	Very Good
4	8%	44%	44%	4%	0%
5	16%	28%	36%	16%	4%
6	16%	40%	32%	12%	0%
7	12%	8%	56%	20%	4%
8	4%	36%	24%	32%	4%

The results indicate that approximately 12% of the participants were strongly agree that tone, format, vehicle of delivery, and language length of IELTS academic tests are good representative of TLU situation tasks. About 30% of the participants stated that there is a weak correspondence between TLU situation tasks and test tasks. The table also shows that about 40 % of the participants indicated that tone, format, vehicle of delivery, and language length of IELTS tests are average representative of TLU situation tasks. Approximately 10% of the participants stated the correspondence between these features of IELTS academic tests input and TLU situation is very weak and about 2 % of the participants believed that the correspondence is very good. The results of input analysis are shown in the above table.

Expected response characteristics: This section of authenticity checklist consisted of five items including type, channel, form, format and length of IELTS academic test tasks. These five items used to determine whether there is a correspondence between the characteristics of expected responses in TLU situation and IELTS academic tests tasks.

Table 3.3. *The Participants' Perceptions of Expected Response Characteristics*

Items	Very little	Little	Average	High	Very high
9	0%	40%	40%	12%	8%
10	4%	44%	32%	8%	12%
11	4%	44%	40%	12%	0%
12	4%	16%	32%	36%	12%
13	0%	40%	44%	12%	4%

The descriptive statistics indicate that approximately 40% of the participants believed that there is little similarity between the expected response characteristics of TLU situation and IELTS academic tests tasks. About 38% believed that there is an average correspondence between the expected response characteristics of TLU situation and IELTS academic tests tasks. The results also show that about 2% of the participants stated that there is very little correspondence between the expected response characteristics of TLU situation and IELTS academic tests tasks. All the results are presented in the above table.

Interaction between Response and Input of IELTS academic Tests: this section of authenticity checklist included three items on participants' perception of interaction between response and input. It encompassed three aspects of interaction between IELTS academic test tasks which they took and TLU situation tasks. According to Bachman and Palmer (1996), input and responses interact along at least three dimensions: reactivity, scope and directedness.

241

Table 3.4. *Participants' Perception of Interaction between Response and Input*

Items	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
14	4%	40%	24%	20%	12%
15	0%	36%	36%	24%	0%
16	4%	40%	32%	16%	8%

The results indicate that approximately 3% of the participants strongly agreed and 40% agreed that there is similarity between three dimensions of interaction of IELTS academic test tasks which they took and TLU situation tasks. About 30% of the participants were neutral. The table also shows that about 20% of the participants disagreed and 7% strongly disagreed that there is similarity between three dimensions of interaction of IELTS academic test tasks which they took and TLU situation tasks.

3.4.3. Authenticity from IELTS Teachers' Point of Views

The descriptive data including teachers' mean score and standard deviation on the authenticity checklist are presented in the following table.

Descriptive statistics including mean, SD, and SEM of all items are presented in Table1.

Table 3.5.*Descriptive Statistics*

Item	N	Mean	Std. Deviation	Std. Error Mean
q1	100	2.0000	.55048	.05505
q2	100	2.9400	.73608	.07361
q3	100	2.8900	.88643	.08864
q4	100	2.6900	.90671	.09067
q5	100	2.7600	1.07422	.10742
q6	100	2.6900	.83720	.08372
q7	100	3.0000	.85280	.08528
q8	100	2.7900	.91337	.09134
q9	100	2.7800	.73278	.07328
q10	100	2.70	.835	.083
q11	100	2.81	.873	.087
q12	100	3.06	.818	.082
q13	100	2.8182	.80005	.08041
q14	100	2.9091	.96986	.09747
q15	100	2.7576	.78365	.07876
q16	100	2.9394	.95638	.09612

The results of the above table show that the mean scores of the English teachers on items 1, 4, 5, 6, 9, 10, 11, 15, is less than the cutoff point, whichis 3. Therefore, it could be strongly argued that IELTS academic tests are not authentic in terms of the following items of authenticity checklist:
- There is correspondence between the main objectives of IELTS test tasks and target language use (TLU) situation tasks.
- Context setting of IELTS tests is representative of context setting of TLU situation tasks.
- Tone of IELTS test tasks is representative of TLU situation tasks.
- Format of IELTS test tasks is representative of TLU situation tasks.
- Channel of IELTS test tasks is the same as the channel of TLU situation tasks.

- Format of IELTS test tasks is the same as the format of TLU situation tasks.

- Form of IELTS test tasks is the same as the form of TLU situation tasks.

- The scope of interaction between the input and response of TLU situation tasks is the same as that of IELTS tests tasks.

The results also indicated that the mean scores of English teachers on items 2, 3, 7, 13, 14, 16, 12, is 3 (some scores of items are rounded off) which is the same as the cutoff score. Therefore, it could be argued that these IELTS academic tests are somehow authentic in terms of the following items of authenticity checklist:

- Delivery vehicle of IELTS test tasks is similar to delivery vehicle of TLU situation tasks.

- Language length of IELTS test tasks is the same as the length of TLU situation tasks.

- The responding procedure in IELTS tests tasks and TLU situation tasks are the same.

- Number of tasks and task distinction in IELTS tests and TLU situation use are the same.

- Length of IELTS tests tasks is similar to length of TLU situation tasks.

- Type of IELTS test tasks is the same as the type of TLU situation tasks.

- The interaction between response and input in IELTS test tasks is the same as the interaction between the response and input in TLU situation tasks.

- The interaction between input and response in TLU situation tasks and IELTS tests is always direct.

The results also indicated that IELTS test is more authentic in terms of items 7&12 with the greatest means of 3 and 3.06 and it has the least

authenticity on items 1, 4 and 6 in terms of the following items on the authenticity checklist:

- There is correspondence between the main objectives of IELTS test tasks and target language use (TLU) situation tasks.
- Context setting of IELTS tests is representative of context setting of TLU situation tasks.
- Format of IELTS test tasks is representative of TLU situation tasks.

3.4.3. Qualitative Results of Authenticity of IELTS Tests

In order to determine authenticity of IELTS tests and the participants' perceptions of the authenticity, the following leading questions were asked:

- To what extent are the IELTS tasks the representative of real life?
- *Due to real and daily life, are the IELTS items enough?*
- *Are you, as a person who has taken IELTS test and got favorite score, able to report your projects in English?*

- Do you have any problem in understanding the lectures in international conferences? Note taking? Asking questions?
- *What do you think of authenticity as a critical feature of IELTS tests?*
- *To what extent are the IELTS items the same as the real life tasks?*
- *To what extent are the IELTS tasks the representative of real life?*
- *Do you think IELTS academic tests are good indicators of your ability to use English in your studies? Why?*

The 20 participants consisted of IELTS academic test takers in Iran. They were all interviewed. The semi-structured interviews were transcribed and content analyzed and the following themes were extracted. Interesting commonalities among the interviewees are shown in the following part of study.

• IELTS is a kind of exam, which tests techniques related to four skills

Mostly examinees believed that those who want to take IELTS academic test should know and be familiar with especial techniques of reading, writing and listening and IELTS exam is an exam to test techniques and techniques are really effective in IELTS test.

• The IELTS tasks are not representative of real life

The interviewers indicated that writing and listening tasks are limited to specific topics e.g., in writing module just a graph and a topic are used, but there are a variety of writing models in real life, like letters, reports, etc. Listening is not representative of listening in real life because sometimes strange listening tasks are used which are totally different from daily life listening tasks used by people.

• IELTS items are not enough

Mostly examinees believed that IELTS test items are not enough, especially writing and reading. They stressed that more items are needed to measure candidates language abilities, i.e. reading, writing, speaking and listening ability.

• Reading and writing tasks are not like real life tasks

IELTS examinees believed that reading and writing of IELTS test are not like real life tasks because only a graph and a general topic are used in writing module while there are more than these types in real life. That's why they cannot be the representative of real life tasks.

• Authenticity is absent in IELTS tests.

Almost all participants believed that situational and interactional authenticity are nice characteristics of language tests but the IELTS tests administered in

Iran lack this characteristic. They also argued that a small fraction of real life tasks is covered by the tasks measured by IELTS tests.

• **Understanding lectures in international conferences, Note taking and asking questions**

The participants believed that it depends on the lecturers and the texts used for lectures. Moreover, accent is very important; some lecturers use accents that are difficult to understand, lecturing speed is also important. Basically, as a person who took IELTS test, may come across with some problems in conferences held in English. Getting 6 or 7 score of IELTS does not guarantee the full understanding of lectures in international conferences held in English.

• **Authenticity is the core of any IELTS test construct validity.**

The majority of the participants stated that an IELTS academic test without tasks and contents representative of TLU situation tasks does not make sense. They argued that it is urgent that test takers design test tasks that the test takers will encounter in TLU situations. Otherwise, these tests do not test the communicative competence of the test takers.

• **Authenticity influences generalization validity.**

Almost all participants stated that the inferences and decisions made about their abilities based on their scores on IELTS tests are not valid because there is a very low similarity between the tasks and contents of these tests and real academic non-test tasks. In other words, the participants' high achievements in these tests do not guarantee their success in real academic life situations. The interviewees believed that studying in universities in which English is language of teaching needs more ability in English.

• **IELTS test can not guarantee academic paper writing**

IELTS by itself is not enough for those who want to write their academic papers and report their academic projects. Extra grammar and writing self-study is needed to improve their writing. Most of interviewees said that two writing items are not enough to measure candidates' ability in academic writing.

• **Understanding technical terms requires self-study**

Participants in IELTS test are majoring in different fields such as chemistry, biology, mathematics, computer, accounting, etc. Therefore, understanding technical terms and texts related to examinees field of study depends on self-study because the maximum technical terms used in IELTS texts, books and tests is about 30 to 4o terms for each field of study, but the candidates need to know more technical words and phrase. They can learn these terms from various special English books, papers, technical dictionaries and different internet sites.

4. Discussion

The demand for IELTS in Iran, China, India etc. has led to the globally increase of IELTS test takers who are offered academic and general versions. The growth in test numbers shows the increase in the number of organizations and institutions considering IELTS as a criterion to meet their needs for language proficiency assessment. A fundamental quality of IELTS academic tests is authenticity (Backman, 1999). In the present study, we tried to identify the perceived authenticity of IELTS academic tests through different approaches and techniques. An important common result of these approaches

is that at least those IELTS academic tests administered in 2011-2012 in Iran, which were the focus of this study, lacked this quality.

Test tasks should be based on learners' language needs, that is, test tasks should be correspondent to the learners' language needs. The mentioned correspondence is one of the distinguishing qualities of IELTS academic tests. To determine this correspondence, learners' language needs were compared with the types of tasks and the skills measured by the IELTS academic tests. The outcomes indicated that there was no significant correspondence between test tasks and TLU situation tasks. Therefore, it could be said that there was a great difference between what language learners needed and what was measured by IELTS academic tests. In other words, the frequently reported real life tasks were not covered by the contents of IELTS academic tests. That is, neither did the tasks of these tests correspond to TLU situation tasks, nor did they measure the learners' abilities to use language for their own professional and academic needs.

Correspondence between TLU situation tasks and the contents of IELTS academic tests is another fundamental aspect of authenticity. The results in tables, 4.2, 4.3, & 4.4 show that almost all of the participants believed that there is a very low correspondence between the contents of IELTS academic tests and TLU situation tasks. The comparison of characteristics of test rubrics, input characteristics, output characteristics, and the interaction between input and response of IELTS test tasks and TLU showed that they are not similar to each other. Therefore, it could be strongly argued that the IELTS academic tests administered to in Iran to those who apply for academic version to study in universities where English is used as medium of instruction in 2011-2012 lack the required degree of authenticity.

The function of IELTS academic tests is to measure language proficiency of applicants who are supposed continue their studying at English universities and to see whether they are able to study at universities where the medium of instruction is English or not. Almost all the interviewees believed that this type of test can does not function based on the set objectives. The interviewees stated that those who get the required band score can not meet their needs in academic and real life settings. They indicated that the required band score gotten by candidates as an indicator of entrance to such universities is not enough. IELTS academic test is necessary to see if the candidates need further instruction.

The results of table3.1, 3.2, 3.3, 3.4, 3.5 showed that the IELTS academic tests are nor authentic in terms of correspondence between the main objectives of IELTS test tasks and target language use (TLU) situation tasks, - Context setting of IELTS tests, tone of IELTS test tasks format of IELTS test tasks,- Channel of IELTS test tasks, form of IELTS test tasks and the scope of interaction between the input and response of TLU situation tasks

The finding of this study for the first questions are in line with Alibakhshi (2010), Douglas (2001), Ellis (2003), Vosoughi (2013), Spolsky (2008), Heltai (1991). Based on the results, it could be argued that non-familiar test tasks, lack of correspondence between test tasks and TLU tasks, and mismatch between the learners' academic needs and test tasks decrease the authenticity of IELTS academic tests. Lack of authenticity of IELTS academic tests could be because of the fact that the IELTS academic tests are not developed based on TLU situation tasks characteristics, test takers' specific academic needs and content familiarity. Keep it another way, focusing on specific academic needs, specific relevant topics, and close correspondence between IELTS academic test tasks and TLU situation tasks which could lead to any IELTS academic test

authenticity is ignored by IELTS tests developers. Although matching the test tasks and TLT tasks may lead to situational authenticity, it can not ensure the interactional authenticity.

Authenticity, therefore, provides a means for investigating the extent to which score interpretations could be generalized beyond ESAP learners' performance on the test to language use in TLU domain, or to other non-test real academic situations. Such definition of authenticity links it to construct validity (Bachman, 1996).

References

Abdi, A. & Asadi, B. (2013). The authentic materials and their advantages and disadvantages.

Bachman, L. F, (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Bachman, L. F. (1991). What does language testing have to offer? *TESOL Quarterly*, 25, 671-704.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Bachman, L. F. (2000). Modem language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17*(1), 1-42.

Baleghirzadeh, S. (2002). Comparative study between the test of spoken English and the IELTS interview with respect to their authenticity. *Humanities and Cultural Studies*, 64-82.

Birjandi, P., Alavi, S. M., &Salmani-Nodoushan, M. A. (2002). Text familiarity, reading tasks, and ESP test performance: A study on Iranian LEP and non-LEP university students. Tehran University.

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics*, 6, 60-70.

 Brendan, J. (1980). *Testing communicative competence*. Oxford: Pergamon Press.

Christine Klein-Braley (1985). A study in the construct validation of authentic tests, *Language Testing*, SAGE Publications, 75-104.

Gebril, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Language Testing*, *26*, 507–531.

Ingram, D. E. (2003). Towards More Authenticity in Language Testing, AFMLTA National Conference 2003, *Languages Babble, Babeland Beyond*, Hilton Hotel, Brisbane, 10 – 12 July.

Khodadady, E., &Hashemi, E. (2011). Validity and C-Tests: The role of text authenticity. *Iranian Journal of Language Testing*, 1, 30-41

Lewkowicz, A. (2000). Authenticity in language testing: Some Outstanding questions. *Language Testing*,*17*(1) 43–64.

More, T. & Morton, J. (2007), Authenticity in IELTS writing Task 2, www.ielts.org

Shohamy, E., Reves, T., &Bejarano, Y. (1986). Introducing a new comprehensive test of oral proficiency.*ELT Journal, 40*, 212-220.

Shomoossi, N.,   & Tavakoli, M. (2010), Authenticity and Authentication in Language Testing: An Operational Perspective, MJAL 2:1 JANUARY 2010, 1-26.

Spolsky, B. (1985). The limits of authenticity in language testing. *Language Testing,* 2(1), 31–40.

Spolsky, B (1985). The limits of authenticity in language testing. *Language Testing*, SAGE Publications, 31-40.

Stevenson, K. (1985). Authenticity, validity and a tea party. *Language Testing*, SAGE Publications, 41-47.

On the Authenticity of IELTS Academic Tests

Appendix

Authenticity Checklist

Dear ....

We are making a study on International English Language Testing System (IELTS) .Could you please help us by completing this questionnaire? All information will be treated in the strictest confidence and please do **NOT** write down your name on the questionnaire. Thank you very much!

Instruction: Please express your opinion of the sayings by choosing one of the five levels of agreement.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

1. There is correspondence between the main objectives of IELTS test tasks and target language use (TLU) situation tasks.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

2. The responding procedure in IELTS tests tasks and TLU situation tasks are the same.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

3. Number of tasks and task distinction in IELTS tests and TLU situation use are the same.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

4. Context setting of IELTS tests is representative of context setting of TLU situation tasks.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

5. Tone of IELTS test tasks is representative of TLU situation tasks.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

6. Format of IELTS test tasks is representative of TLU situation tasks.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

7. Delivery vehicle of IELTS test tasks is similar to delivery vehicle of TLU situation tasks.

1. Strongly agree 2. agree 3. Neutral 4.disagree 5.strongly disagree

8. Length of IELTS tests tasks is similar to length of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

9. Format of IELTS test tasks is the same as the format of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

10. Channel of IELTS test tasks is the same as the channel of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

11. Form of IELTS test tasks is the same as the form of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

12. Language length of IELTS test tasks is the same as the length of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

13. Type of IELTS test tasks is the same as the type of TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

14. The interaction between response and input in IELTS test tasks is the same as the interaction between the response and input in TLU situation tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

15. The scope of interaction between the input and response of TLU situation tasks is the same as that of IELTS tests tasks.

    1. Strongly agree 2. agree 3. Neutral   4.disagree 5.strongly disagree

16. The interaction between input and response in TLU situation tasks and IELTS tests is always direct.

    1. Strongly agree 2. Agree 3. Neutral   4. Disagree 5.strongly disagree

Thank You