On the Authenticity of Academic IELTS Tests Practiced in Iran

Goodarz Shakibaei

Department of Humanities

College of English language

Ahvaz Azad University Branch

Islamic Azad University, Ahvaz

Iran.

ABSTRACT

The purpose of this paper was to investigate the authenticity of academic IELTS tests. Beckman and Palmer (1996, p23) 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 the study included 70 Iranian IELTS test takers who took academic IELTS test in 2012 to 2013, and 100 IELTS teachers. Twenty IELTS test takers also used to collect the qualitative part of data through interviewing. . A mixed design study, both qualitative and quantitative, was used. The data were gathered through a semi-structured interview and authenticity checklist. The checklist was translated into Persian. The reliability of translation was checked through backward translation technique. Descriptive statistics used to analyze the collected data. The findings of the study showed that academic IELTS tests are not authentic in terms of tone, format, channel, form and scope of interaction. Qualitative findings also showed that number of IELTS tests items is not enough.

Keywords: Authenticity, Academic IELT tests, Real life, Test takers

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, p23) 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 assumed that language testing scale could include 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, Davit 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 1996:128).

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

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

f. Authenticity relates to assessment (Bachman 1991).

e. Authenticity relates to the social situation of the classroom (Breen 1985, Arnold 1991).

In a test, authenticity may be presented in the following ways:

a. The language in the test is as natural as possible.

b. Items are contextualized rather than isolated.

c. Topic are meaningful (relevant, interesting) for learner.

d. Tasks represent or closely approximate, real life tasks.

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 distinguished: situational and international authenticity (Lewkowicz 2000:48). While situational authenticity was considered with the relationship between test task and the criterion behavior, international authenticity indicated the international between test task and test user (Authenticity in language assessment by: Katharine 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 authenticity, only a few studies have been carried out to investigate academic IELTS tests in terms of authenticity. To investigate this quality of academic IELTs tests, the following question is formulated and answered in this study.

-To what extent are the contents and tasks of the current IELTS tests practiced for entrance to national and international universities authentically representative of the tasks in target language situation"?

2. Review of Literature

Carroll Brendan J. (1980) believes that: the issue of authenticity must always be an important aspect of any discussion on language testing. A full application of the principle of authenticity would mean that all the tasks under taken should be real-life, interactive communicative operations and not the typical routine examination response to the teachers stimuli, or a part a stimulus – response relationship that the language of the test should by day-to-day discourse not edited or decorated in the interests of simplification but presented with all its expected irregularities, that the contexts of the interchanges are realistic, with ordinary interruptions, background noises and irrelevancies found in the airport or lecture-room, and the rating of a performance, based on its effectiveness and adequacy as a communicative response, will rely on non-verbal as well as verbal criteria.

In applied linguistics, the term authenticity originated in the Mid 1960s with a concern among material writers such as Close (1965) and Broughton (1965) that language learners were being exposed to texts which were not representative of the target language they were learning. Authenticity at the time was seen as a simple notion distinguishing text extracted from real life scares from those written for pedagogical popular Modern language test theory places great emphasis on the usefulness of which consists of six test qualities, that is, reliability, validity, authenticity, instructiveness, wash back, and practicality.

Nematullah Shomoossi & 2. Mansoor Tavakoli (2010) reviewed 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 require a dynamic and operational model to include the human elements and all stages of designing authentic test tasks, which may not be considered as the end point 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.

Winnie Yuk- chun (1995) Lee 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.

D. E. Ingram (2003) studied authenticity in language testing. He briefly discussed the progress in language testing (especially proficiency assessment) towards more authenticity and referred 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, 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 19 contexts in which the language was tested resemble those of real-life, the more accurately is the language test likely to predict how the candidate will cope, at least linguistically, with real-

life activities. The geographical limitations of the International Second Language Proficiency Ratings (ISLPR) imposed by the need for candidates to be able to access a centre for a face-to-face interview and the limitations imposed by the relationships and situations that can be reproduced in an interview room are further reduced by the ISLPR Global as we search for ways to bridge the gap between language tests and real-life language performance and move towards more authenticity in language testing.

Heidar Abdi and Bahare Asadi(2013) investigated the authentic materials and their advantages and disadvantages. They explained how authentic materials can 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 gain confidence in being able to function in an English-speaking society.

Ebrahim Khodadady and Mahmood 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.

Marjan 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.

Hsin-min Liu (2005) studied the methods for assessing authenticity in computer-assisted language learning and assessment. The results that authenticity has become a multi-componental 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.

Malcolm MacDonald , Richard Badger, Goodith White(2000) conducted a research on usefulness of the criterion of authenticity for the selection and evaluation of EAP listening materials. They found that the texts that related to the students experience and permitted learner interaction appeared to have more potential for language learning than those which merely replicated the discourse of the target situation. The analysis has brought them to the practical conclusion that the use of video and especially simulated teacher presentation within the EAP classroom can lead to greater engagement of the learners than using audiotapes of actual lectures, although the latter might accord more with crude notions of text authenticity in EAP.

Weiping M. Wu Charles W. Stansfield (2008) described a working model used to determine the Target Language Use (TLU) (Bachman and 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.

Bernand 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.

Elana Shohamy and Thea Reves (1985) discussed authentic language tests: where from and here to? They reviewed the trend of development towards authenticity in language testing. They showed how language tests moved from the indirect era towards becoming more direct and more authentic, attempting to elicit the language used by real people in real life. It was pointed out, however, that although these tests are more face valid, i.e. they seem to resemble more the actual language used in real life, they still have some major deficiencies. One such deficiency is the lack of measurement and statistical analysis and the limited empirical evidence to show their psychometric qualities. They also pointed out that the language of 'authentic tests' is not a true representation of real-life language.

Christine Klein-Braley (1985) in a paper entitled "A cloze-up on the C-Test: a study in the construct validation of authentic tests" claimed that C-Tests are authentic tests of the construct of general language proficiency.

Tim Morore and Janne Morton (2007) conducted 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.

Maria Luisa and Ana Almagro 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, 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.

Robyn Spence-Brown (2001) studied the authenticity in an embedded assessment task in a university-level Japanese language course in Australia. The article examined a teaching and assessment activity based around interviewing native speakers outside the classroom, which was designed to optimize the authenticity. The article argues that the addition of an assessment dimension fundamentally changes the nature of a task and thus compromises authenticity. It suggests that authenticity must be viewed in terms of the implementation of an activity, not its design, and examines various factors that affect the degree of authenticity experienced by individual students.

Mohmood Reza Atai and Masoud 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. Lewkowicz, Jo. A (2000) also studied examined the concept of authenticity and the way it evolved in language testing and general education. The researcher looked at the importance of authenticity for test takers, and showed that test takers are willing and able to identify the attributes of a test likely to affect their performance. Bernand Spolsky believed that lack of authenticity in material or method used in a test weakens the generalizability of results and only part of the problem can be overcome by authentic-seeming tasks. Observation of authentic behavior is another partial solution.

Sasan Bagherzadeh 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 module 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.

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 70 Academic IELTS test takers from different provinces of Iran received the authenticity checklist questionnaire. The completed questionnaires were used to investigate the authenticity quality of academic IELTS 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 academic IELTS test. These academic IELTS 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 academic IELTS 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 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 were identified

through colleagues at the universities and educational institutes which 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, 190 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 used to develop a checklist to identify the authenticity of academic IELTS 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 on four 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' view about IELTS test modules including 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 modules and then I asked them follow-up questions to explore more about IELTS components, the relationship between the four modules including writing, reading, speaking and listening.

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 questionnaire, 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 academic IELTS 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 70 IELTS test takers and 100 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:

The results of authenticity

The results of authenticity including quantitative and qualitative are presented in different parts. At first, the quantitative results are presented.

3.4.1. 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.

Table1: 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

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, which is 3. Therefore, it could be strongly argued that academic IELTS 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 indicate 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 academic IELTS 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 indicate that IETS test is more authentic in terms of items7&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.

3.4.2. Authenticity from IELTS test takers' views

Table2: Descriptive statistics

	N	Mean	Std. Deviation	Std. Error Mean
q1	50	2.240	1.21744	.14058
q2	50	2.840	.97315	.11237
q3	100	2.960	.77877	.08992
q4	100	2.440	.70212	.08107
q5	100	2.640	1.06085	.12250
q6	100	2.400	.90045	.10398
q7	100	3.013	1.06450	.12211
q8	100	2.960	1.00593	.11615
q9	100	2.560	.80940	.09346
q10	100	2.480	.94954	.10964
q11	100	2.760	.81936	.09461
q12	100	3.040	1.08354	.12512
q13	100	2.64	.747	.086

q14	100	2.99	1.133	.131
q15	100	2.960	.8767261	.1012356
q16	100	2.80	1.027	.119

The results of the above table show that the mean scores of the IELTS test takers on items 1,2,4,5,6,9,10,11,13,16 is less than the cutoff point which is 3. Therefore, it could be strongly argued that academic IELTS tests are not authentic in terms of the following items on authenticity checklist questionnaire:

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

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

- 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.

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

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

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

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

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

The results also indicate that the mean score of IELTS test takers on items 3, 7, 8, 12, 14, 15 is 3 or larger than 3 (some scores of items are rounded off) which is the same as the cutoff score. Therefore, it could be argued that these academic IELTS tests are somehow authentic in terms of the following items of authenticity checklist:

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

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

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

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

The results also indicate that IETS tests are more authentic in terms of items7&12 with the greatest means of 3.0132 and 3.0400 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 academic IELTS tests are good indicators of your ability to use English in your studies? Why?

The 20 participants consisted of academic IELTS 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 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 nice but 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.

●**IELTS tests are needed but they should not serve as selection test.**

Another extracted theme was that IELTS tests are needed, but a change in their function is required. Almost all participants stated that IELTS tests should be administered to diagnose the students' weakness and strength in the use of academic language in TLU situations; whereas, almost all IELTS tests administered in Iran are used either to measure the students' ability to continue their studying at an English speaking university or to work in an English speaking country. Participants argued that even though the students may pass these tests, it does not mean that they can do all TLU situation tasks in English.

● **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 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 which 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.

   REFERENCES

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

   University Press.

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

   671-704.

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

   University Press.

   3.   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.

   4.   Bachman .L .F (2000), Modern language testing at the turn of the century: assuring that what we count counts, Language Testing 17 (1) 1–42.

   5.   Bernard Spolsky (1985),The limits of authenticity in language testing, *Language Testing*,

   SAGE Publications, P. 31-40.

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

   Carroll Brendan J. (1980), Testing Communicative Competence, Pergamon Press, Oxford, 1980,

   P.11 f.

      8. Christine Klein-Braley (1985), a study in the construct validation of authentic tests,

      Language   Testing, SAGE Publications, P.75-104.

   9.   Douglas K. Stevenson (1985), Authenticity, validity and a tea party, *Language Testing* ,

   SAGE Publications, P. 41-47.

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

   11. Elana Shohamy and Thea Reves (1985) Authentic language tests: where from and where to?

   , *Language Testing* , SAGE Publications, P. 48-59.

   12. Herbert W. Seliger (1985), Testing authentic language: the problem of meaning,

   Language Testing, SAGE Publications, P. 1-15.

   13. Jo A. Lewkowicz(2000), Authenticity in language testing: some Outstanding questions

   , Language Testing, 17 (1) 43–64.

   14. Lee, W. (1995). Authenticity revisited: text authenticity and learner authenticity. *ELT Journal, 49 (4),* 323-328

   15. Miller, L. (2003). Developing listening skills with authentic materials. *ESL Magazine, 6 (1),* 16-19.

16. Lewkowicz, J.A. (2000): Authenticity in language testing: some outstanding questions. *Language Testing* 17, 43–64.

17. Lynch, A.J. 1982: 'Authenticity' in language teaching: some implications for the design of listening materials. *British Journal of LanguageTeaching* 20: 9–16.

18. Malcolm MacDonald *, Richard Badger, Goodith White (2000) , Authenticity and Academic listening, English for Specific Purposes 19. 253-267.

19. Nematullah Shomoossi and Mansoor Tavakoli(2010), Authenticity and Authentication in Language Testing: an Operational Perspective, MJAL 2:1 JANUARY 2010, 1-26.

20. Porter, D. and Roberts, J. (1981). Authentic listening activities. English Language Teaching

 Journal,   36 (1), 37-47.

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

22. Time More and Janne Morton (2007), Authenticity in IELTS writing Task 2, www.ielts.org.

23. Weiping M. Wu and Charles W. Stansfield(2001), Towards authenticity of task in test development, *Language Testing* ; 18; 187,p. 186-206.