Abstract

Authenticity has always been the concern of test developers in the history of second language assessment. This study was an attempt to investigate the authenticity of the present researchers’ innovative idea of “Language Town” as a method for assessing learners’ speaking ability. To this end, a simulated town was designed like a real town in an outdoor space of about 400 square meters. The participants in this study were 31 undergraduate students of Translation Studies at Jahrom University who were taking the 4-credit course of Conversation 2. First, the students’ speaking ability was measured based on the IELTS testing system. Then each student was sent to the Language Town with a few definite missions determined in advance by the teacher; e.g. ordering food in the restaurant, depositing/withdrawing money in a bank, etc. Using IELTS band descriptors, each student’s speaking ability was measured by two raters in both tests. Then the correlation between the two sets of scores obtained from the IELTS test and the one in the Language Town were calculated. Using open-ended questions, a survey was also conducted to extract the students’ attitudes towards the Language Town. The results of the statistical analyses showed a weak correlation (0.36) between the two sets of scores. The survey also revealed that almost all the students were interested in and had positive views towards Language Town as an authentic method of assessment. A Virtual Language Town (VLT) could be a solution to the practicality problems of the Language Town.

Keywords: Language Town, Authentic Assessment, Speaking Ability
Introduction

The major problem of teachers and learners of a second language, particularly those who deal with the language for the purpose of using it in the real context of the target language community, is the lack of authenticity of the teaching and testing methods. As a result, such approaches do not prepare the students appropriately for the real and challenging tasks which, in addition to syntax, language functions, fluency and coherence, pronunciation, range of lexis, demand other vital context-dependent competencies, such as pragma-linguistics, sociolinguistics, etc. In fact, the history of language testing is, to a large extent, the history of attempts to bridge the gap between tests and real-life language use. It is the history of progress towards more authenticity in language testing (Ingram, 2003, p. 4).

Considering the IELTS test as one of the most recent developments in the field of language testing, second language teachers have a tendency to use this test for assessing their students’ general competence, including their speaking ability. However, different studies have revealed that in spite of an overall positive response to the test, the results of the IELTS tests cannot truly predict the learners’ success in the real context of the target language community (Allwright & Banerjee, 1997; Cotton & Conrow, 1998; Hill, Storch, & Lynch, 1999; Kerstjen & Nery, 2000; Moore & Morton, 1999). This situation is even exacerbated when it comes to testing the learners’ speaking ability. For example, Paul (2007) suggested that language production at a micro level similar to that in IELTS Speaking Test tasks is not necessarily an indicator of overall language adequacy at a macro level. Most directly, Rea-Dickins, Keily, and Yu (2007) found the Speaking Module to be a poor predictor of test takers’ future academic performance. Rea-Dickins et al. (2007) added that even when students’ scores on the sub-skills of the test fulfill a program’s admission requirements, students are found to lack critical thinking and evaluative skills.

Furthermore, Ducasse and Brown (2011) indicated that the structured nature of the IELTS speaking interviews appears to limit the
ability of interviewers to elicit a broader range of interactional functions, even with more proficient speakers. They argued that as candidates are only required to respond to direct questions or propositions, no evidence of their ability to participate actively in oral discussions – to find a way to share and discuss their own knowledge or ideas, express their opinions, and challenge, support, or evaluate others’ contributions – was gathered. According to Ducasse and Brown (2011), studies concerned with the predictive validity of IELTS have generally found little or no significant relationship between IELTS scores and subsequent general and academic performance. This lack of relationship is even more evident in IELTS Speaking Tests.

According to Pinner (2016), the problem is directly related to the issue of authenticity of language tests, including the IELTS. In one of his most recent works, Pinner (2016) replaced the ‘classic’ definition of authenticity with a reconceptualized version, which, as he claims, is more inclusive to other varieties of English. He poses the ‘paradox of authenticity’ arguing that authenticity is not something absolute, but “rather relative to the learner and their unique and individual beliefs” (p. 1). He states that in the real world, learners’ motivations, in interaction with multiple contextual factors, make the task of communication quite complex and challenging; a condition which hardly exists in current language tests. Consequently, such tests are not authentic enough to guarantee the learners’ successful task accomplishment in the real world. He tries to discuss authenticity in light of emergent theories of language acquisition such as chaos/complexity theory and dynamic systems approaches and consequently, introduces the Authenticity Continuum, which is a framework for treating authenticity as a “socially mediated” and “contextually dependent dynamic process of investment”. Based on this definition, one can easily recognize the fact that the designers of language tests, including the IELTS, have largely ignored the important factors of context and society in the administration of their tests. In fact, most of the tests are administered in artificial situations which lack the necessary conditions of the real world contexts.
Concerning the authenticity limitations of the IELTS tests, particularly the speaking module, this study was an action research to practice the present researcher’s innovative idea of “Language Town” as an authentic method to replace IELTS speaking test for assessing learners’ speaking ability in her Conversation classes.

**Language Town**

Language Town is a simulated model town which is designed like a real town consisting of different places (e.g. a bank, a restaurant, shops, hospital, bus station, etc.), with some real primary facilities in each location and some proficient English language speakers in charge of each location. These proficient speakers play the role of attendants who interact with customers to fulfill their needs, just like what happens in the real world. The area of this town could be variable depending on the available space. The most important thing to consider is to design the town as real as possible to give the examinee the feeling of an authentic context. Each examinee is given a few missions, such as shopping, ordering food in a restaurant, depositing money in a bank, or other similar tasks that an individual is required to do in the real world. Those examinees that succeed to accomplish the task through effective communication, are considered as successful language users. The missions that are assigned to the students are selected based on the needs analysis carried out in advance.

**Objective of the Study and Research Questions**

This study was intended to investigate the authenticity of the present researchers’ idea of “Language Town” as a method for assessing learners’ speaking ability. More specifically, this study tried to answer the following questions:

1. Is there any relationship between the learners’ speaking scores obtained through Language Town method of assessment and those achieved using the IELTS testing system?

2. What are the examinees attitudes towards Language Town as a method of assessing speaking ability?
Design of the Study
This study used a mixed method to evaluate the authenticity of Language Town as a method for assessing learners’ speaking ability. In the first phase of the study quantitative research was carried out to determine whether there is any correlation between the IELTS scores and Language Town scores. In the second phase of the study, a survey was conducted to elicit the participants’ feelings and attitudes towards Language Town.

Participants
The participants in this study were 31 undergraduate freshman students of Translation Studies (11 males and 20 females) at Jahrom University who were taking the 4-credit course of Conversation 2 with the present researcher. It was the second semester of the academic year.

Instruments
IELTS speaking module questions were used to assess the learners’ speaking ability in the first phase of the study. Different sets of questions were used for different participant to avoid practice effect.

IELTS Band Descriptors was used to assess the students’ performance on both speaking tests (Appendix A).

Open-ended questions were used in the survey to elicit the learners’ attitudes towards Language Town as a method for assessing their speaking ability (Appendix B).

Data Collection
In the first phase of the study, the students’ speaking ability was evaluated based on the IELTS testing system in two sessions. Conversation 2 is a 4-credit course which is held two sessions a week. The assessment took about 4 hours in each session. In this phase each student was interviewed by the researcher using the questions of the three parts of the IELTS test. In addition to the researcher who rated each student, another rater who was trained in advance was present to evaluate the students’ performance using IELTS Band Descriptors. The average score of the two raters were calculated for each student to represent their IELTS speaking score. In the following week, Language
Town method was used to assess their speaking ability. As a result, during the two sessions in the following week, each student was sent to the Language Town, which was designed in advance. Each student was given a few definite missions determined by the teacher; e.g. shopping, ordering food in a restaurant, depositing/withdrawing money in a bank, chatting and discussing a topic with friends, talking to university professors in their office, etc. In the present study, proficient senior students were appointed as attendants in different locations of the Language Town. Each student’s speaking ability was measured by two raters who followed the students in each location in the town, observed them, and based on the IELTS Band Descriptors, evaluated their communicative competence through interaction with the attendants in each location. Just like the IELTS scores, the average scores of the two raters gained from the Language Town method were calculated for each student to represent their speaking ability.

In the second phase of the study, a survey was conducted to extract the students’ attitudes towards the Language Town as an authentic method for assessing their speaking ability. To this end, a set of open-ended questions were designed in advance to be asked from the students. The first two questions asked about the examinees feelings towards their Language Town experience. And the second two questions focused on the authenticity of the tasks the examinees were assigned to accomplish. The questions were written in Farsi (the students’ mother tongue) to avoid any probable misunderstanding. Also the students were asked to write their answers in Farsi to let them express their actual attitudes freely and comfortably. The examinees wrote their answers to the questions on sheets of papers which were collected for later analysis.

**Data Analysis**

Pearson Correlation was conducted between the two sets of scores obtained from the IELTS test and the Language Town method.

The answers to the survey questions were read and analyzed carefully to find about the students feelings and attitudes towards Language Town approach. Also the students’ opinions towards the
authenticity of the Language Town were elicited through the analysis of the answers.

**Results and Discussion**

Regarding the first question of the study, whether there is any relationship between the learners’ speaking scores obtained through Language Town method of assessment and those achieved using the IELTS testing system, Tables 1 and 2 provide the answer. Having a look at the mean scores of the tests in Table 1, one can immediately notice that the mean of Language Town scores (4.8387) is smaller than that of the IELTS scores (5.8548).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>IELTS</td>
<td>5.8548</td>
<td>.91464</td>
<td>31</td>
</tr>
<tr>
<td>LGTOWN</td>
<td>4.8387</td>
<td>1.59367</td>
<td>31</td>
</tr>
</tbody>
</table>

Furthermore, Table 2 shows a weak correlation index (0.36) which is significant at the 0.05 level. This means that there is not a strong relationship between the participants’ IELTS speaking scores and their scores obtained via the Language Town.

<table>
<thead>
<tr>
<th></th>
<th>IELTS</th>
<th>LGTOWN</th>
</tr>
</thead>
<tbody>
<tr>
<td>IELTS</td>
<td>Pearson Correlation</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>31</td>
</tr>
<tr>
<td>LGTOWN</td>
<td>Pearson Correlation</td>
<td>.361*</td>
</tr>
</tbody>
</table>
One possible interpretation for the lower mean score of the Language Town test compared to that of the IELTS test might be the fact that Language Town tasks were more difficult for students to accomplish than the IELTS speaking module questions. This confirms Pinner’s (2016) argument that in the real world, learners’ motivations, in interaction with multiple contextual factors, make the task of communication quite complex and challenging; a condition which hardly exists in current language tests, including the IELTS tests. In fact, excluding the context and simply sitting in front of the examiner and answering a set of questions in IELTS tests is far easier than being situated in a real context which demand more complex and multidirectional interactions both with different participants and the real objects present in the environment.

On the other hand, the weak correlation in Table 2 can imply that IELTS tests cannot truly predict the learners’ success in the real context of the target language community. This is exactly in line with the previous studies (Allwright & Banerjee, 1997; Cotton & Conrow, 1998; Hill, Storch, & Lynch, 1999; Kerstjen & Nery, 2000; Moore & Morton, 1999; Paul, 2007; Rea-Dickins, et al. 2007) which indicated a low predictive validity for the IELTS tests. One good justification for the weak correlation between the two sets of scores could be what Ducasse and Brown (2011) argued that the structured nature of the IELTS speaking interviews appears to limit the ability of interviewers to elicit a broader range of interactional functions, even with more proficient speakers. In fact, since candidates are only required to respond to direct questions or propositions, there is no evidence of their ability to participate actively in oral discussions, such as finding a way to share and discuss their own knowledge or ideas, expressing their opinions, and challenging, supporting, or evaluating others’ contributions. This is also confirmed by the lower mean of the Language Town scores.
compared to the IELTS scores, showing the fact that Language Town tasks probably demanded a higher-order and more complex cognitive ability compared to those in the IELTS test.

Furthermore, the results of the survey revealed that almost all the students were interested in and had positive views towards Language Town as an authentic method of assessment. For example, one of the students stated, “testing in a … mmm… real context like Language Town is more exciting and …mmm… more meaningful.” At the same time, they argued that the tasks in the Language Town were more challenging to accomplish. For example, when one of the participants was asked why he thought Language Town tasks were more demanding, he answered, “in the IELTS speaking test, the only thing we …mmm… had to do was to answer the questions the interviewer asked… but in the Language Town I have to speak with a person, act physically, and work with real objects to complete the tasks… so I need more concentration.” Some other students argued that since it felt like a real context and that they had to interact with different people, speaking and communication were more difficult in the Language Town compared to the IELTS test. In fact, most of the answers implied the fact that the multiple contextual factors involved in the Language Town had made it more challenging and demanding.

On the whole, considering the results of both the correlation and that of the survey, one can infer that the Language Town has provided an authentic context which, as Pinner (2016) stated, replaced the ‘classic’ definition of authenticity with a reconceptualized version, which, as he claims, is more inclusive to other varieties of English. In fact, Language Town created a real context in which the learners’ unique and individual beliefs were involved in doing the tasks. Language Town, due to its similarity to the real world situations, entangles learners’ motivations, in interaction with multiple contextual factors and consequently, makes the task of communication quite complex and challenging; a condition which hardly exists in other language tests, like IELTS. In fact, a testing context like the Language Town can introduce tasks which are, as

Limitations of Language Town Idea
One obvious limitation of the Language Town idea as an authentic method for assessing the speaking ability is its low practicality, especially for high-stake tests. The need for a large space, decoration, various objects, and proficient attendants in each location are both expensive and time-consuming. In fact, it is not economical to develop such a Language Town for every speaking test and by every institute independently. However, one possible solution to this problem is to replace current language labs in schools, universities, and language institutes with small Language Towns. That is, we can devise one large, fixed, and permanent Language Town as a center specifically allocated to testing learners’ speaking ability. In this way, the Language Town will be designed and equipped once and used by different groups of learners who should reserve the Language Town in advance to be used for their testing purposes.

Another idea, which could be even more economical than those mentioned above, is to develop a Virtual Language Town (VLT). This is what can be used easily in high-stake tests as well. In the following section, an imaginary VLT has been described and the way it functions has been elaborated.

Virtual Language Town (VLT): An economical solution to current inauthentic language tests
Wherever it is too dangerous, expensive or impractical to do something in reality, Virtual Reality (VR) is the answer (Virtual Reality Society, 2017). Considering the practical limitations of the Language Town, VLT can be considered as an innovation in the field of language testing. VLT is a virtual interactive Language Town utilizing virtual reality technology and computer programs which place the examinees in simulated 3D locations with smart virtual participants with whom test takers can communicate and be evaluated by automatic and intelligent raters based on predetermined rubrics and criteria defined for the program. In fact, VLT can make use of VR technology to place
language test takers in different locations and situations to interact with various virtual characters in each place. This is exactly what happens in interactive computer games. This can be done through collaboration among TESL, linguistics, and computer program experts.

One can think of an examinee standing in front of a 3D screen which displays a Language town. She starts walking in the streets looking for a bank nearby. She finds the bank, enters the building, takes a turn, and goes to the virtual banker. She starts communicating with the banker to do her banking affairs. After she completes her mission in the bank, she comes out of the building and since she feels a little bit tired, she goes to a coffee shop for some drink. She sits at the table, picks the menu and selects her favorite drink. Then she asks the waiter to come to her table and write down her order. After having her drink she asks for the bill and pays the bill. These are some sample missions that are given to the examinee to be accomplished within a predetermined amount of time to test her language use competence and in particular her speaking ability. The VLT is programmed in such a way that only those examinees who can communicate effectively can complete each task successfully and be considered as competent language users. Different levels of competencies can be defined in advance for the program and each candidate, based on their knowledge, can be assigned to their appropriate levels. Depending on the institute’s goals, VLT can be programmed for both prognostic and evaluation of attainment tests. Furthermore, VLT can be administered simultaneously for a large number of students depending on the number of 3D screens available in the testing center.

**Conclusion**

Regarding the results of the present study, Language Town, especially its VLT version, can be considered as an authentic and economical method for testing all language skills, including the speaking ability. In fact, Language Town can replace language labs in different institutes including schools and universities to solve the problem of current language tests which suffer from inauthenticity and low predictive validity. Being situated in a simulated context with those characteristics
of the real world, both the examiners and examinees can much better evaluate one’s competencies in a second or foreign language. Language Town method of assessment is especially useful for those teachers and learners of a second language, who in particular deal with the language for the purpose of using it in the real context of the target language community. Considering vital context-dependent competencies, such as pragma-linguistics and sociolinguistics, in addition to the knowledge of syntax, language functions, fluency and coherence, pronunciation, range of lexis, Language Town can prepare learners appropriately for the real and challenging tasks they encounter in the real world.

References


**Appendices**

**Appendix A**
سوالات بخش نظرسنجی:

1. احساس خود را در مورد تجربه شهرک زبان توصیف کنید.

2. بطور خاص و با ذکر دلایل توضیح دهد آیا نظر شما نسبت به این شیوه‌ی ارزیابی مثبت است یا منفی؟

3. کدام شیوه ارزیابی را بیشتر منطبق بر نیازهای ارتباط شفاهی خود در دنیای واقعی می‌دانید، روش آیلتس یا شهرک زبان؟ با ذکر دلیل توضیح دهد.

4. از نظر شما کدام شیوه ارزیابی دشوارتر و چالشی تر می‌باشد، آیلتس یا شهرک زبان؟ با ذکر دلیل توضیح دهد.