

On the Evaluation of IELTS Listening Test and its Washback Effects on Listening Teaching and Learning

Xiaoli Wang

School of Foreign Languages, China West Normal University, China

Abstract: *Language Testing and teaching, learning have extremely closed relationship so that they are interrelated and interact on each other. Test is testing what gets taught, so it can reinforce learning and motivate students. Listening comprehension, an important indicator to measure one's foreign language proficiency, is always a headache especially for Chinese students among tests of listening, speaking, reading and writing. As IELTS's success is undeniably and undoubtedly significant, it is high time that we made a detailed evaluation of IELTS listening test to have a deep understanding of its washback effects. This article firstly is going to evaluate the Listening section of IELTS from three aspects: test authenticity, test validity and reliability, and then analyzes its washback effects on listening teaching and learning.*

Keywords: *IELTS listening test, Evaluation, Washback effects, Listening teaching and learning*

1. INTRODUCTION

IELTS is the international English language testing system which tests English proficiency across the globe. It is widely recognized as a reliable means of assessing the language ability of candidates who need to study or work where English is the language of communication. It divides into Academic Module and General Testing Module. Each module consists of four parts: Listening, Reading, Speaking, and Writing. And all candidates take the same Listening and Speaking. Testing of listening comprehension is an indispensable part in language testing. For Chinese students, however, they always feel that listening is the most difficult part of English learning and accordingly they get low scores in IELTS Listening test. Washback or backwash means that a test will influence teaching, a test will influence learning, a test will influence what or how teachers teach and what or how learners learn, and so on. Therefore, the IELTS listening test is needed to be analyzed to improve our listening teaching and learning based on its washback effects which include positive and negative ones.

2. TEST AUTHENTICITY

As to language testing researches, authenticity has always been the key concern. Bachman (1990) defined authenticity of language testing from two aspects: authenticity as real-life language use and authenticity as interactive language use. The former one stresses applying language, say, daily expressions in testing and the latter emphasizes the communicative function of language in testing. The authenticity of language testing refers to the degree of similarity between the subject language the tester uses in the test and the language he uses in reality, which is the degree of unification of language test and language communication. That is to say, the language which the test takers will meet in the test is what they see or they hear in their daily life; the tasks they are required to do are the problems they may need to solve in the real world; and the evaluation criteria of their performance is based on to what degree they have finished the tasks and achieved the goal of effective communication. I will also attempt to analyze authenticity of IELTS listening test from three perspectives—testing materials, testing tasks and testing situations respectively.

2.1. Testing Materials

The IELTS listening module consists of four sections, each with ten questions. The first two sections are concerned with social needs. The first section is a conversation between two speakers and the second section is a monologue. The rest two sections are concerned with situations related to educational contexts. The third section is a conversation among up to four people and the fourth section is a monologue.

First is the authenticity of listening materials. The testing material is a key factor to examine whether the test is authentic or not. The authenticity of listening materials requires the materials themselves be similar to real life. Authentic testing materials are for users of their first language, not materials designed for foreign students (Wilkins, 1976). Materials in IELTS listening tests are largely in accordance to the real-life language use, and cover a wide range of topics from all aspects of academic life. For example, a Chinese student goes to the UK for further education. He needs to rent an apartment to live in (Cambridge IELTS Book6, Test4, Section 1); he might want to apply for a part-time job (Cambridge IELTS Book7, Test3, Section 1); then he attends the orientation organized by the school (Cambridge IELTS Book4, Test2, Section2); during his lectures, he surely needs to discuss with his classmates (Cambridge IELTS Book5, Test4, Section3); he listens to teachers' lectures while taking notes (Cambridge IELTS Book5, Test1, Section4); during holidays, he may want to have a trip (Cambridge IELTS Book4, Test1, Section2), etc. All these contents are covered by the IELTS listening test. In this regard, the materials in the listening module of IELTS are truly from real-life and the topics cover different aspects of life which students may encounter during their stay in foreign countries, therefore they can deeply and comprehensively assess test takers' all-round ability to deal with different situations potentially appearing in their future life.

Furthermore, these materials also reflect the verbal features in real-life communication, such as pauses, repeat, slip of tongue, etc. In natural language, there is 30% to 50% speaking time containing pauses and hesitation. These pauses may be filled or silent pauses. Filled pauses are pauses like *well, I mean, hmm*, and so on, which indicate that the speaker is thinking about a proper expression to continue his/her talk. Filled pause is a significant part in oral language, which is also a distinguishing feature compared to other types of language expressions.

Additionally, the IELTS listening module contains various English accents. This is necessary since it is true among native English speakers. Some regional accents are easily identified by certain characteristics. There is also much room for misunderstanding between people from different regions, as one word that is pronounced in one accent will sound like a totally different word in another accent. Apparently, people may meet different English accents in different places. Thus, to test the real English listening ability, test designers should create difficulties in speakers' accents to examine whether the test takers understand the material or not.

2.2. Testing Tasks

The definition of the criteria of task authenticity is "One of the crucial aspects of task authenticity is whether real communication takes place; whether the language has been used for a genuine purpose."

Apart from the materials, up to 7 testing tasks are utilized in IELTS: multiple choice, notes/form/table/summary/flow-chart completion, short-answer question, sentence completion, labeling a diagram/plan/map, classification and matching respectively. It is apparent that IELTS uses various testing tasks to examine test takers' language ability. According to the criteria mentioned above, all kinds of testing tasks in IELTS can be identified as "authentic" or "non-authentic". Coupled with the data in IELTS papers, the authentic testing tasks usually occupy the vast majority of the total testing

tasks. By completing these testing tasks, test takers' certain abilities are practiced and these abilities are also needed in the real life. In this sense, these testing tasks are comparatively authentic.

2.3. Testing Situations

In the aspect of testing situations, various situations are covered by the IELTS, which are taken from real life. There are chiefly up to 10 different situations involved. Each of them is commonly met in daily life. Obviously, it is a successful duplication of the real life in order to create similar or even same situations in the test to examine test takers' abilities. Thus the IELTS listening module is highly authentic in the aspect of testing situations.

To sum up, from the data analysis above, it is indicated that the IELTS listening test is relatively authentic in the perspectives of testing materials, testing tasks and testing situations.

3. TEST VALIDITY

It is generally agreed that a good test should obtain validity and reliability. Let us first look at validity of a test. According to Brown, validity, which reveals the extent to which the test actually measures what it is supposed to measure and nothing else, is by far the most complex criterion of a good test. The validity of a test consists of face validity, content validity, criterion-related validity, and validity in scoring.

Firstly, the Listening test items of IELTS look right to other tester, teachers, moderators and testees. It can be described as having at least face validity. A test with good face validity can maintain students' motivation of learning while a bad face validity test may make students put less effort into performing, thus affect the reliability.

Secondly, "A test is said to have content validity if its content constitutes a representative sample of language skills, structures, etc, with which it is meant to be concerned"(Hughes,2003). Accordingly, the listening test samples of IELTS actually achieve the content validity successfully.

Thirdly, criterion-related validity is obtained as a result of comparing the results of the test with the results of some other criterion measures. It consists of concurrent and predictive validity. The former deals with the comparison with other test results collected at or about the same time; the latter with those at a significantly later time. Generally speaking, IELTS is a mature test, and it well predicts the future success of the candidates in language aspect. I will just assume that IELTS listening test, to a large extent, is doing a good job in criterion-related validity.

Lastly, validity in scoring. This is worth mentioning because even if the validity of a test is proved, the invalid scoring can still make the test invalidity. As for the scoring of IELTS listening test, it places candidates on a spectrum from zero to nine, candidates' levels are clear and valid.

4. TEST RELIABILITY

Reliability is another important factor in deciding whether it is a good test or not. Reliability refers to the degree to which test scores represent test takers' true score. Reliability is considered as an approach of the coherence of marks gained from various performance measurements (Bachman & Palmer, 1996). Bachman & Palmer (1996) declare that if the same test were to be administrated to the same group of individuals on two different occasions in two different settings, it should not make any difference to a particular test taker.

As the questions of IELTS listening test are very practical and based on real life situations, students can use them in daily life situations. Students' answers will not be influenced by the physical environment because the IELTS listening test is usually held in a quiet classroom. Another issue with

reliability is student-related. For instance, if a student is in a good mood, he/she may perform very well in the test. On the other hand, if the participant is in a bad mood, he/she may get a low score in the test. It can be seen that physical or psychological problems can have an impact on students' real listening ability.

In general, from the analysis above, the listening test of IELTS is relatively valid and reliable in many aspects.

5. DEFINITIONS OF WASHBACK

The term "washback", also known as "backwash", has almost as numerous definitions as the researchers who studied about it (Bailey, 1999). The following definitions chronologically present ranges from simple and straightforward ones to rather complex.

Buck (1988) described washback in a study on listening comprehension test in Japan. It is quoted as follows:

There is a natural tendency for both teachers and students to tailor their classroom activities to the demands of the test, especially when the test is very important to the future of the students, and pass rates are used as a measure of teacher success. This influence of the test on the classroom (referred to as washback by language teachers) is, of course, very important; this washback effect can be either beneficial or harmful.

Then Shohamy (1992) viewed washback as "the utilization of external language tests to affect and drive foreign language learning (in) the school context". Buck and She saw eye to eye on a view that the result of this phenomenon is due to "the strong authority of external testing and the major impact it has on the lives of test takers." Besides, Alderson and Wall (1993) assumed that washback as an important concept in applied linguistics refers to "the extent to which the introduction and use of a test influence language teachers and learners to do things they would not otherwise do that promote or inhibit language learning". Moreover, Messick (1996) claimed the term as "the extent to which the test influences language teachers and learners to do what they would not necessarily do". It is further noted that "evidence of teaching and learning effects should be interpreted as washback"(Messick,1996). Brown and Hudson(1998) also took an essential aspect into consideration that"washback is the effect of testing and assessment on the language teaching curriculum that is related to it".

To sum up, although the original review of the definitions of washback discusses much on its definition and scope, they all agree that washback refers to the influence of test instead of test itself, which should be emphasized by the researchers.

6. WASHBACK EFFECTS OF IELTS LISTENING TEST ON LISTENING TEACHING AND LEARNING

According to the purpose of the listening test of IELTS, it can inform instruction by identifying students' particular profiles of weaknesses and strengths. With regard to individual testees, they become familiar with everything involved in the real IELTS listening test and gain valuable experience in taking tests. Upon reflection, they may also be able to pinpoint the perceptual or cognitive problems that underlie the test score by themselves. The test can arouse students' self-awareness to pay close attention to communicative language use. Thus, this will build positive feedback for individual students, teachers and courses. Upon completion, the listening test has a strong washback effect on classroom teaching and learning. The students with low English proficiency had problems with communicative language use and understanding. In China, students are

expected to learn grammar-translated, audio-linguistic methodology. In this regard, most of Chinese candidates are “dumb” and have a hard time in listening comprehension. Therefore, teachers should spend more time on conversational listening and speaking practice and use more authentic materials.

6.1. The Positive Washback of IELTS Listening Test on Listening Teaching and Learning

The IELTS listening test had indeed exerted positive washback on aspects of teachers’ and students’ psychology, assessment, contents and methods.

First of all, as for teaching or learning psychology, although they affirmed that the IELTS listening test brought anxiety and pressure for listening teaching or learning, most of them insisted that certain stress should be transformed into stimulation which further strengthened their confidence and perseverance. Besides, as for teaching or learning assessment, the positive washback lied in that they all acknowledged the usefulness and necessity of IELTS listening test as a language measurement tool which could be employed to examine their merits and shortcomings in language teaching or learning. What’s more, in the aspect of teaching or learning contents, IELTS listening test exerted positive washback on emphasizing comprehensive language ability since most participants acknowledged that it enriched their teaching and learning contents and firmly believed the effectiveness of authentic materials such as VOA or BBC news and stories. Finally, both teachers and students made efforts to combine test-oriented and ability-oriented teaching or learning method driven by the positive washback of IELTS listening test. Besides, it made their teaching and learning more scientific and valid.

6.2. The Negative Washback of IELTS Listening Test on Listening Teaching and Learning

A sound test also has negative effects because of something done or undone. However, the negative washback of IELTS listening test on teachers and students was not as obvious as the positive one.

Generally speaking, the greatest negative washback on teachers was the difficulty in keeping a balance between the improvement of students’ listening comprehension ability and the training to help them get a higher score in IELTS listening test, especially some students with poor language foundation. A satisfying score in listening test does require test-taking skills, which forces teachers to arrange more class hours to practice it. Without doubt, the action disturbs their teaching planning and teaching methods. However, it is a challenge for teachers to overcome the negative effects and thus succeed in combining test-orientation and ability-orientation.

The strongest negative washback of IELTS listening test on students is the conflict between their expectation of success and their passive learning method. Students know clearly that they do not learn for test and their listening comprehension ability is the priority, whereas their learning attitude and learning objective are easily influenced by tests. For instance, they affirm the effectiveness of IELTS listening test and employ it to enhance their listening ability, but most of them say no, when asked if listening section is not required in the test, would they use it any more. The answer shows strong tendency that students choose learning methods for the test to a large extent.

7. CONCLUSION

The purpose of this article is to evaluate the listening test of IELTS by making a deep analysis from three main aspects—test authenticity, test validity, test reliability and presents its washback effects on listening teaching and learning. It is not difficult to draw the conclusion that the listening test of IELTS, to a large extent, is relatively authentic, valid and reliable and also has more positive but less negative washbacks.

Based on the above discussion, there are some suggestions on listening teaching and learning. For one thing, the teachers should ensure students' access to various kinds of systematical and authentic teaching contents, including cultural background, pronunciation as well as intonation in the IELTS listening test according to students' language competence. And Teachers should change their traditional teacher-centered teaching methods into communicative teaching which enables teachers to have a better understanding through interaction with the students' difficulties in English learning. For another, students must practice more English pronunciation, which is the prerequisite condition in that only if they can pronounce well, can they identify the sounds while listening. Students should also manage to be exposed in more authentic information concerning cultural background and common sense.

At last, it is expected that this article may offer test designers some implications on how to make improvements in terms of test authenticity, validity and reliability and may also shed some new lights on improving listening teaching and learning on the basis of washback effects of IELTS listening test.

REFERENCES

- [1] Alderson, J. C. & D. Wall. (1993). Does washback exist? *Applied Linguistics*, 14 (2), 115-129.
- [2] Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: a study of washback. *Language Testing*, 13(3), 280-297.
- [3] Bachman, L.F. *Fundamental Considerations in Language Testing* [M]. Oxford: Oxford University Press, 1990.
- [4] Bachman, L.F. & Palmer, A.S. *Language Testing in Practice* [M]. Oxford: Oxford University Press, 1996:44
- [5] Bailey, K. M. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing*, 13(3), 257-279.
- [6] Bailey, K. M. (1999). *Washback in Language Testing*. Princeton, NJ: Educational Testing Service.
- [7] Buck, G. (1988). Testing Listening Comprehension in Japanese university entrance examinations. *JALT Journal*, 10(1), 15-42.
- [8] Brown, J. D. & T. Hudson. (1998). Alternatives in language assessment. *TESOL Quarterly*, 32, 653-675.
- [9] Cheng, L.Y. (1997). How does washback influence teaching? Implications for Hong Kong. *Language and Education*, 11(1), 38-54.
- [10] Cheng LY, Watanabe Y, Curtis A. Washback in language testing: Research Contexts and Methods [J] *TESL-EJ*, 2004, 8(1), R-14.
- [11] Cheng, L.Y. (1999). Changing assessment: washback on teacher perspectives and actions. *Teaching and Teacher Education*, 15(3), 253-271.
- [12] Guariento, W. & Morley, J. Text and Task Authenticity in the EFL Classroom [J]. *ELT Journal*, 2001.
- [13] Hughes, A. *Testing for language teachers* [M] Cambridge: CUP, 2003, Chap4,5, 9.
- [14] Hughes, A. (2003). *Testing for language Teachers*. Cambridge: Cambridge University Press.
- [15] Harris, D. F. (1969). *A Language Testing Handbook*. New York: McGraw-Hill.
- [16] *IELTS HANDBOOK 2006*. Available at: <http://www.ielts.org/default.aspx>.
- [17] Kerry o'Sullivan and Steven Thurlow. *Focusing on IELTS Listening and Speaking Skills 2*. 2002.

- [18] McNamara T. Language Testing [M] Oxford: Oxford University Press, 2000.
- [19] Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13 (3), 241-256.
- [20] Madaus, G. (1984). *The influence of Testing on the Curriculum*. Chicago: University of Chicago Press, 1988.
- [21] Official Examination Papers from University of Cambridge ESOL Examinations IELTS 4-7[M]. Cambridge: Cambridge University Press.
- [22] Shohamy, E. (1992). Beyond proficiency testing: a diagnostic feedback testing model for assessing foreign language learning. *The Modern Language Journal*, 76, 513-531.
- [23] Wilkins, D. *Notional Syllabuses* [M]. Oxford: Oxford University Press, 1976. 陈晓扣. 论语言测试的反拨作用[J]. *解放军外国语学院学报*, 2007, 3: 40-44.
- [24] 孙贝妮. *雅思听力全解全析*[M].北京:世界图书出版公司,2006
- [25] 黄大勇. 语言测试的真实性概念[J].*语言教学与研究*, 2004(2):79-82
- [26] 冯小芳. 雅思听力对大学英语听力教学的启示[J].*成都大学学报*, 2007(1):62-64
- [27] 刘芳. 基于雅思反拨效应的中外合作项目英语教学改革——以桂林电子科技大学为例 [J]. *桂林师范高等专科学校学报*, 2015(2) : 154—158.

AUTHOR'S BIOGRAPHY



Xiaoli Wang, is a postgraduate student at the School of Foreign Languages of China West Normal University in China. She majors in curriculum and pedagogy (orientation: English). Her research interests are English teaching, error analysis of English writing, and language testing.