

## Issues with the Speaking Section of the International English Language Testing System (IELTS): An Assessment Critique

Solihin<sup>1</sup>, Dita Rezkia Utami<sup>1</sup>, Aris Dwi Intan Aprianti<sup>1</sup>, Mayang Sari<sup>2</sup>

<sup>1</sup> STKIP Paracendekia NW Sumbawa, Indonesia

<sup>2</sup> SMPIT Sumbawa, Indonesia

\*Correspondence: [solihin.mohe@gmail.com](mailto:solihin.mohe@gmail.com)

### ABSTRACT

*The International English Language Testing System (IELTS) is one of the most widely used tests of English internationally. It is a testing system designed or owned by three partners: University of Cambridge ESOL Examination, British Council, and IDP Australia. This research is to critique the speaking section of the IELTS based on the experience of some test-takers involved as the subjects of this research. The method used was descriptive qualitative. Data collection techniques applied were content analysis, interview, and self-reflection. Two IELTS test-takers were interviewed to get their views on the three phases of the speaking section. The authors' own experiences when taking the IELTS were also reported. It was found that there are issues with the speaking section of the IELTS. Those issues are its: fairness, authenticity, validity, and the appropriateness of interview as a sole tool to assess speaking ability. This paper also suggests a number of actions to be taken in order to make the speaking section fair, authentic, and valid. Additionally, other ways of assessing speaking ability are proposed. This paper then reveals what these issues mean to test-takers, examiners, and organizations running IELTS.*

### ARTICLE HISTORY

Published September 18<sup>th</sup> 2023



### KEYWORDS

Assessment Critique; IELTS; Issues; Speaking.

### ARTICLE LICENCE

© 2023 Universitas Hasanuddin  
Under the license CC BY-SA  
4.0



### 1. Introduction

The assessment policy or procedure that is critiqued or evaluated in this paper is the International English Language Testing System (IELTS) as one of the most widely used tests of English internationally (Karim & Haq, 2014). It is a testing system designed or owned by three institutions: University of Cambridge ESOL Examination, British Council, and IDP Australia. The number of test-takers who have participated in this testing system has been over 100,000 at 251 approved centers in more than 105 nations (Roshan, 2013).

The International English Language Testing System (IELTS) is an English language test designed to measure a person's English language skills, especially for those who are not native English speakers and wish to continue their studies or work in English-speaking countries (Read, 2022; Anada et al., 2018). IELTS measures four key English skills: 1) Listening: This tests your ability to understand spoken English in various contexts, such as everyday conversations and lectures. 2) Reading: This tests your ability to understand written text in a variety of text types, including news articles, essays, and reports. 3) Writing: This tests your ability to express your thoughts in writing well in two tasks, namely task 1 (writing a report) and task 2 (writing an essay). 4) Speaking: This tests your ability to speak English clearly and effectively in an interview. The majority of the candidates taking this test are usually those who are planning to study or work in countries where English is used as a communicative language. Moreover, IELTS score is one of the basic prerequisites stipulated by most higher education institutions holding majors for international students. Besides, there are also candidates who just simply want to measure the level of their English ability. The skills that are tested in this testing system are the four language skills: listening, reading, writing, and speaking. It tests both academic and general training versions of language proficiencies of the test takers (Cambridge University Press, 2006).

The tested skill that will be focused on, in this occasion, is the speaking skill. The purpose of speaking skill is to interact in the target language that involves comprehension as well as production. Therefore, from this point of view, IELTS is unique because it measures speaking skill directly through live interaction (Karim & Haq, 2014; Ulker, 2017). There are three parts of the speaking section in the IELTS which all together take between 11 to 15 minutes in total. Part

I (around 4 to 6 minutes) is about one's personal life such as job, hobby, interest, where they live, and other familiar topics. Its purpose is to get the candidate settled or make him or her comfortable. Part II is about a more specific topic provided on a cue card. It takes 1 minute for preparation and after that, the candidate must talk around 1 to 2 minutes. The topics can be about shopping, travel, favorite TV shows, etc. Part III is about deeper questions related to Part II. For example, if Part II is about a holiday, the questions in part III may be about tourism or travel. It takes around 4 to 5 minutes. The candidates are required to discuss the topics in a more abstract view (Issit, 2008).

## **2. Methodology**

### **2.1. Design**

This study was designed as a descriptive qualitative. Qualitative research (Lincoln & Denzin, 1994 in Anderson, 1998) involves the collection of a variety of empirical materials-case study, personal experience, life story, interview, historical, and visual view. This research aims to give some critiques towards the speaking section of the IELTS. These critiques were based on some facts that the subjects of this research experienced.

### **2.2. Participants**

The participants of this research were two test takers of the IELTS and the authors' own experience. Their experiences regarding how many times they took the test, how they felt during the test, etc. were asked during the interview. The researcher's own experience in taking the test was also reported.

### **2.3. Instruments**

The tools used in this study were voice recorder, semi-structured interview guide, and a number of speaking questions of the IELTS, including phases I, II, and III. The voice recorder and interview guide were used during the interview with the participants. The speaking questions were referred to during the content analysis activity.

### **2.4. Data Collecting Techniques**

The methods of data collection used were content analysis, interview, and self-reflection. Content analysis is the systematic examination of texts and visuals (e.g., newspapers, magazines, speech transcripts), media (e.g., films, television episodes, internet sites), and/or material culture (e.g., artifacts, commercial products) to analyze their prominent manifest and latent meanings (Saldana, 2011).

### **2.5. Data Analysis Techniques**

Techniques of data analysis were following the ones proposed by Miles and Huberman (1994). They are data reduction, data display, and conclusion drawing. The steps of data analysis were, first the researcher identified the issues regarding the assessment, particularly the speaking section of the IELTS. After that, critiques or evaluation were given on the basis of what the subjects have gone through. Interview is one way to get in-depth data in which the researcher asked the subjects to give opinions or ideas more openly (Sugiyono, 2010). According to Frank (2016), self-reflection as a method is used to observe what one says, how she or he behaves, what and how he or she writes, then reflects upon these as a way of learning or innovation.

## **3. Result and Discussion**

### **3.1. Findings**

#### **1) Assessment Issues**

Although IELTS is a very widely used test, there seems to be problems related to the arrangement of the sections of the speaking part. They are pertaining to the fairness of the topic in phase two, the authenticity of the topic, and the validity of the speaking test results. Besides, there is also an issue related to interview as the only sample eliciting tool of speaking skill. The first three issues (fairness, authenticity, and validity) are caused by the unpredictability of the topic given on a cue card in phase two (Part II) of the speaking test and the last one is seen in relation to the representativeness of speaking skills assessed through interview only. These issues are against the principles of language assessment proposed by many experts. For example, Brown and Abeywickrama (2010), in their book titled "Language Assessment Principles and Classroom Practices" suggest that language assessment should be fair, authentic, and valid. Brown et al., (2012) further mentioned that fairness is critically important in an assessment process for both students/test-takers and teachers/examiners. Similarly, Darr (2005) contended that what is tested should be a fair sample of the area of competence we are interested in. According to Ashwin (2015), some of the features of a good

assessment task are that it should be seen as a realistic and a valid representation of what they are learning or want to achieve.

### **3.2. Discussion**

#### 1) Critiques/Evaluation

##### a) Fairness

Most candidates consider phase II of the IELTS speaking test is more about getting their luck rather than showing their ability. Based on the researcher's own experience, this seems to be true. When the researcher asked some candidates who were taking the test on the same day as researcher was, "how do you think you will go?", most of them said, "I hope to get a topic that is familiar to me, otherwise I may be in trouble". This does not mean that the candidates come unready to the test. In fact, they have done an IELTS course and therefore, of course, they come with some preparation. However, sometimes, in phase II, some of them are given a topic about the things they are not familiar with, even very completely different from that in phase one. This gets worse when they are under time pressure during the one-minute preparation. They have to think about what to say, what words to use and how to express the ideas for the things they actually have only few ideas about.

This situation makes the fairness of the test questionable. Some candidates may be given topics which are in line with their background knowledge, but some others are forced to talk about topics they have no idea about (Bruning, 2010; Nunan, 1999). In other words, the test requires the candidates to talk about a topic they have no schemata or prior knowledge about. The researcher himself experienced this when he was facing the speaking test of the IELTS. He was given a topic about a music show that he had been to. In fact, he had never been to any music show. During the one-minute preparation, he could not do much. He had to think about what music show he would mention, what phrases to use, and how to express them. As a result, he could not answer the questions on the cue card optimally.

The responses given by one the subjects of this study indicate that she received different marks in the speaking skill for two different topics of two different tests. In her first test she was handed out the cue card that contained a familiar topic in speaking part II, but in her second test, she was given an unfamiliar topic. She said that the score for her speaking is different. The first speaking score was higher than her second one. This is also the case that the authors faced. He took the test two times. In speaking part II in his first test, he was given a familiar topic. The second one was not very familiar. Although the score was the same, he did not feel confident with the second topic because he had never experienced it.

##### b) Validity

Related to the unpredictability of the topic in phase II, when test takers are presented with a topic that they have no idea about, the validity of the test results may be questionable. As is known, validity is the state in which a test really measures what is intended to be measured (Brown & Abeywickrama, 2010; Hughes, 2003). If the candidates do not really know about what they have to say pertaining to the given topic, they definitely will talk in an undue manner. This means that the test measures the candidates' specific domain knowledge, not their speaking ability. This is contradictory to the objective of the speaking test in the IELTS which is "assessing speaking ability".

The doubt about the validity of the test can get worse in part III. This is because the questions that are asked in this part are related to the topic provided in part II. If in part II itself the candidates are given a topic that they are not familiar with, how can they respond to the deeper and more abstract questions about it optimally in part III? Therefore, this type of test may not be valid. This is especially important in relation to content validity which is about a test that contains items that can elicit the representative sample of a particular skill. According to Hughes (2003), the content of a test should not be seen from the point of view of what is easy to assess but, instead, what is important. For example, the set of items used to test undergraduates should not be the same as those used for postgraduates. IELTS provides the same sets of tests to all candidates without considering their background knowledge, education level, and age. Thus, the content validity of IELTS is still questionable.

##### c) Authenticity

If the questions in phase two are not at all related to the candidates' real life, for example as mentioned previously about a music show, then the topic is not authentic. To make it clear, authenticity is about the realness of the language and task used in an assessment or a test (Flinders School of Humanities & Creative Arts, 2017). According to Brown and Abeywickrama (2010), when checking the authenticity of a test, we have to ask ourselves a number of questions. Those

questions are: 1) is the language as natural as possible? 2) do the questions/tasks closely reflect real life? 3) is the topic interesting, meaningful, and relevant to the candidates' life? Hence, all of these questions tell us that if the test questions are not related to the test-takers' real life, then the test is not authentic. It is not meaningful to their life.

Accordingly, to deal with the concerns about the unfairness, inauthenticity, and invalidity of the test results, the questions provided on the cue card in part II should be pertaining to the ones familiar to the test-takers' lives.

d) The appropriateness of interview as a sole technique of measuring speaking ability

In relation to the appropriateness of interview as the only sample eliciting tool of speaking skill, IELTS does not seem to consider the varied language we use in real-life communication (Nunan, 1999). IELTS, through interview, seems to only cover the formal context of the language we use. This does not completely represent the real-life situations, where we do not only use formal language, but also the informal one and the language we use in both contexts vary considerably. This is also very clear when we look at the two types of IELTS, namely, IELTS for general training and the one for academic purposes. There is no difference in the speaking section of it in reference to the two different contexts. Therefore, IELTS, in this case, may not elicit or analyze the speaking skill in its true sense. Besides that, though interview is mostly used for assessing speaking skills, relying solely on it may not cover the other speaking skills needed in other contexts. This is because the interview used in IELTS is still in a traditional form where the interviewer is very dominant in taking all initiatives, but the interviewee remains as the one to respond to the questions being asked, especially in part I. Other speaking skills such as asking questions and initiatives to start discussion remain untested. This is in line with the research findings done by Karim & Haq (2014). It was found that the interviewer remained dominant in the interview while the interviewee only had to respond to the questions being asked.

e) How might the speaking test of the IELTS be improved?

This part suggests some revisions in phase II of the speaking test of the IELTS as it affects or is related to the other phases (phases I and III) in order to improve the test quality. More particularly, this part elaborates how it can be made fair, authentic, valid, and elicit the appropriate sample of speaking skill.

To make the test fair, authentic, and valid, the question topic in phase two should be made predictable. It should be based on the information provided by the candidate in phase one, not providing completely a new topic. It can be related to the candidates' job, hobby, and interest provided in phase one. For example, in phase one, the question about what the candidate does is almost always asked. If, for examples, he works as a teacher, the interviewer should hand out a cue card about the things he really likes in teaching. This way then can lead the candidate to only focus on this topic. As the topic in phase number two is related to teaching, then, the questions in phase three are (supposed to) related to teaching. By arranging the questions this way (thematically related from phase one to phase three), then the speaking test of the IELTS can be said authentic, fair, lead closely to validity. It is said authentic because the topic is about the interviewee's real life. It is considered fair as it tests something that is familiar to him or her, not a strange one. It is valid because it really measures what it is intended to measure, not something made up.

More specifically, regarding content validity, there should be more variations in terms of contents being presented to candidates that come from different backgrounds and for different purposes of taking the test. Those who come to take the speaking test for general training purposes should be given a general topic, but those who want to take the one for academic purposes should be given a topic related to their academic life. Giving the same contents for different groups may contribute to invalid results of the test.

Regarding the use of interview as a sole sample eliciting tool of speaking test, there should be other techniques that can go along with it such as role playing, real-world conversation, or picture cued tasks. By doing so, the other speaking skills such as asking information and more interactive skills can be measured. Moreover, the different types of learners such as visual and tactic learners can be aided in doing the test.

f) What might this mean for test-takers, examiners, and organizations running IELTS?

By doing so, it may mean differently to test takers, examiners, and IELTS organizations. In connection with the test takers, they may have to prepare themselves with more focus on the topics related to their daily lives, with proper languages, phrases, sentences, and other language elements needed to get a good score. For the examiners, they may have to prepare questions based on the backgrounds of the candidates, not like choosing a lottery. This is because the examiner is the one who is responsible for taking all the initiatives during the test. For IELTS organizers or teams, they have to develop questions/tasks in the speaking test using the resources that exist in the countries and centers. For

example, if Indonesia is the location, the IELTS test is usually handled by Indonesia Australia Language Foundation (IALF) which are located in different cities. They should base their topics according to the resources that are available in Indonesian cities, but still retain the international standards of the language ability based on the rubric.

#### 4. Conclusion

To conclude, the assessment policy that is critiqued in this paper is the speaking test of the IELTS. The assessment issues are related to the unfairness, inauthenticity, and invalidity of the test results which are caused by the unpredictability/unfamiliarity of the cue card questions given in phase two. Besides, there is an issue associated with the use of interview as the only technique of assessing candidates' speaking skills which does not elicit the real language use in the daily lives of the test-takers. These all are against the language assessment principles stipulated by many scholars. Therefore, to cope with the issues of the unfairness, inauthenticity, and invalidity of the test results, the questions given on a cue card in phase two should be made predictable or familiar to the test-takers. One of the ways can be, for example, they are made connected to the candidates' jobs, hobbies, or hometowns provided in phase I. To tackle the issue of interview as the sole eliciting tool of speaking ability, other techniques should accompany it such as role play and real-world conversations. By doing so, it means the candidates can focus more on the topics related to their daily lives, examiners have to prepare questions related to the backgrounds of the candidates, and IELTS running companies should produce materials based on the resources existing around the candidates' environments. It is admitted that the critiques given in this research are still limited due to the number of sources it involved. There were only two test takers interviewed besides the authors' own experiences. Therefore, further research is needed for deeper analysis of the critiques. It is also acknowledged that the issues raised may vary in significance depending on each test-taker's experiences and abilities as different individuals taking the test may not have the same concerns.

#### References

- Anada, R. P., Arsyad, S., & Dharmayana, I. W. (2018). Argumentative Features of International English Language Testing System (IELTS) Essays: A Rhetorical Analysis on Successful Exam Essays. *International Journal of Language Education*, 2(1), 1-13.
- Anderson. (1998). *Fundamentals of educational research* (2nd ed.). London & New York: Falmer.
- Ashwin, P. (2015). *Assessment: how does it make a contribution to learning*. In Ashwin, Paul (Ed.) *Reflective teaching in higher education* (pp. 249-273). Bloomsbury, London.
- Brown, H. D., & Abeywickrama, P. (2010). *Language assessment: principles and classroom practices* (2nd ed. ed.). White Plains, NY: White Plains, NY : Pearson Education.
- Brown, Sally, Race, & Phil. (2012) *Using effective assessment to promote learning*. In Hunt, Lynne & Chalmers, Denise (Eds.), *University teaching in focus: a learning-centred approach* (pp. 74-91). Camberwell, Vic.: ACER Press.
- Bruning, R. H. (2010). *Cognitive psychology and instruction* (5th ed. ed.). Boston: Boston : Pearson.
- Cambridge University Press. (2006). *Cambridge IELTS 5*. Cambridge: Cambridge University Press.
- Darr, C. (2005). A hitchhiker's guide to validity. *Journal issue*, (2).
- Franks, T. M. (2016). *Purpose, practice, and (discovery) process: When self-reflection is the method*. *Qualitative Inquiry*, 22(1), 47-50.
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Issit, S. (2008). Improving scores on the IELTS Speaking Test. *ELT Journal*, 62, 131-138. <http://dx.doi.org/10.1093/elt/ccl055>
- Karim, S., & Haq, N. (2014). An assessment of IELTS speaking test. *International Journal of Evaluation and Research in Education (IJERE)*, 3(3), 152-157.
- Miles M.B. & Huberman, A.M. (1994). *Qualitative data analysis* (2nd Ed.). California: SAGE Publications Inc.
- Nunan, D. (1999). *Second language teaching and learning*. Boston, Mass.: Boston, Mass. : Heinle & Heinle.
- Read, J. (2022). Test Review: The International English Language Testing System (IELTS). *Language Testing*, 39(4), 679-694.

- Roshan, S. (2013). A critical review of the revised IELTS speaking test. *International Journal of English Language Education*, 2(1), 120-127.
- Saldana, J. (2011). *Fundamentals of qualitative research*. New York: Oxford University Press.
- School of Humanities and Creative Arts, Flinders University. (2017). ESOL methodology study guide, Semester 2. [ESOL8004]. School of Humanities and Creative Arts: Flinders University.
- School of Humanities and Creative Arts, Flinders University. (2017). Second language testing study guide, Semester 2. [ESOL8002]. School of Humanities and Creative Arts: Flinders University.
- Sugiyono. (2010). *Metode penelitian kuantitatif, kualitatif dan R & D*. Bandung: Alfabeta.
- Ulker, V. (2017). The Design and Use of Speaking Assessment Rubrics. *Journal of Education and Practice*, 8(32), 135 - 141.