

IELTS vs. TOEFL: A Comparison Study on Format, Content and Grading Systems

Maryam Kashani Nazari¹

Ahamd Pourgasem^{2*}

Elham Gerami³

Abstract

The reason why International English Language Testing System (IELTS) and the Test of English as a Foreign Language (TOEFL) became and remain powerful in society is a crucial question. Both large-scale tests were purposively developed for assessment of test-takers' English language proficiency who intend to obtain admission to academic institutions or to succeed in immigration purposes. As both tests are considered as high-stakes tests affecting individuals' future lives, they need to gain sufficient knowledge on each test to make a perfect choice. As a result, the current research contributes to enhancing test-takers' knowledge by comparing IELTS and TOEFL in terms of their reliability, validity, format, content and grading systems. A qualitative comparison was used to identify the similarities and differences between both tests. IELTS has two versions of general and academic and TOEFL has three versions of paper-based, computer-based and internet-based. IELTS assesses four traditional skills, and TOEFL-CBT includes listening, structure/writing combination, reading and speaking. IELTS grading ranges between zero to nine while TOEFL scores are cumulative. IELTS takes two hours and half but TOEFL CBT takes four hours. Considering reliability, in IELTS there is lack of rater reliability and writing and speaking are ratered in case of inconsistencies. In TOEFL the emphasis on the selected-response item is to assure the high reliability is met. Base on the findings, the two tests are similar as they are intended to test individuals' language

¹ **Maryam Kashani Nazari**, The Student of IELTS FOR TEACHERS Course, Amirkabir University of Technology (Tehran Polytechnic), International Language Center, Tehran, Iran
Email: maryam.kashani57@gmail.com

² **Ahmad Pourgasem (PhD)**, Amirkabir University of Technology (Tehran Polytechnic), TESOL Trainer and Staff Member of International Language Center, Tehran, Iran (Corresponding Author). E-mail: a.pourgasem@aut.ac.ir

³ **Elham Gerami**, Amirkabir University of Technology (Tehran Polytechnic), IELTS Trainer and Staff Member of International Language Center, Tehran, Iran. E-mail: elhamgr@aut.ac.ir

proficiency and the two tests are different in terms of their format, length and content and grading systems.

Keywords: IELTS, TOEFL, format, content, grading, length, similarities, differences

1. Research Question

There are different English language proficiency tests used nationally. Language tests are used for several aims in many contexts and are considered as one of the vital aspects in individual's professional life. Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS) are most widely accepted language tests worldwide. Both tests are used for English as Foreign Language (EFL) as well as English as Second Language (ESL) learners with the purpose of measuring their English language skills and abilities. The results of these two test can be used in different situations. For instance, for an international student to be approved for admission into an educational institution, the student must have the necessary English language proficiency score that meets the minimum requirements. Moreover, they are compulsory for some people who are applying for a job or an immigration visa. In both cases, merely those individuals who could obtain the required score are successful.

The current study investigates on the similarities and difference between IELTS and TOEFL tests. This study compares these two high-stakes test by considering the similarities and differences on format, content and grading systems as well as the reliability and validity of both test. It is important for test-takers to know how these two test can be similar or different when they plan to choose one over the other. Hence, this study is significant as it provides necessary insights for test-takers to decide which test to take. Therefore, the researcher attempts to answer the question that whether there are similarities and differences between IELTS and TOEFL tests in terms of their format, content and scoring systems.

2. Literature Review

In this section of research elaboration on the background of TOEFL and IELTS including their development and their aims of language assessment are given respectively.

2.1 Background of TOEFL

The TOEFL was developed in 1964 as a tool to assess English language proficiency for admission purposes. It was created by the Education Testing Service (ETS), which is a non-profit organization that focused on developing a language assessment instrument that would be used universally. More than eight thousand educational institutions recognize and accept TOEFL scores in over 130 countries including Australia, Canada, New Zealand, the United Kingdom, and the United States (TOEFL Program History, 2017). The TOEFL test is used at measuring English proficiency of English as a Foreign or English as a Second Language students who plan to continue their education at higher institutions. It assesses the test takers' English language skills including listening, reading, speaking, and writing and identify how they can perform on academic tasks. Some major revisions have been done on the TOEFL since it has been launched which were motivated by advances in theories of language ability and changes in English teaching practices. ETS has designed three versions of TOEFL: paper-based (1964), the Computer-based (1998) and the Internet-based (2004). TOEFL comprises a number of innovative design features, such as integrated tasks that involve many skills to simulate language use in academic settings, and test materials that imitate the reading, listening, speaking, and writing demands of real-world academic environments (TOEFL Program History, 2017).

Moreover, beyond changes in the medium of test, there were some improvements in theories of language proficiency that have motivated the redesign of the construct, content, and types of items in the test. There are significant challenges in test development which relates to retaining a high level of standardization and test's reliability. (TOEFL Program History, 2017).

2.2 Background of IELTS

The IELTS was developed and administered by the British Council, the University of Cambridge Local Examinations Syndicate and the International Development Program of Australian universities and colleges. The IELTS is administered in over 135 countries and the scores are accepted by more than 7,000 educational institutions. As TOEFL or IELTS score is obligatory for international students' admission to colleges and universities, having knowledge about the similarities and differences of both tests are necessary. The IELTS measures general and academic English language proficiency. The academic test is used for those test takers that

intend to study at a post-secondary institution in an English-speaking country. While, the general test is employed for people who would like to work, or migrate to English-speaking countries (IELTS, 2013). The IELTS have four modules namely listening, reading, writing and speaking. The listening and speaking modules are same in both general and academic versions, however, the reading and writing sections differ in each version with the aim that in the academic version, they examine test-takers' skills in academic disciplines which are very similar to their courses of study. While, in the general version reading and writing mostly relates to the general condition where immigrants might face in their daily lives in an English speaking country.

The two tests have been used for numerous goals which give them a powerful position in the world. As they are high-stakes tests, test takers attempt to accomplish well on the tests since the higher scores provide them a better opportunity to apply for jobs, universities and scholarships. Zahedi and Shamsaee (2012) claimed that IELTS and TOEFL tests are two examples of high-stake standardized tests which have significant influence in determining the future life of the test takers. Additionally, tests have power to control what should be taught and what should be learned that is they have power to control the knowledge of people. This is something important about testing where it is used as a disciplinary tool which means that test takers, whether willing or not, have to change their behaviour in order to meet the tests demands (Shohamy, 2001a).

3. Research Method

This paper employed a qualitative comparative approach to examine the issues related to the two large-scale English language tests namely International English Language Testing System (IELTS), and the Test of English as a Foreign Language (TOEFL) in terms of their format, content, grading system, reliability and validity. A number of papers have examined a variety of issues related to these language test instruments, however, they were intended mainly for language testing researchers. The current paper addresses the need of test-takers who wish to enhance their knowledge about IELTS and TOEFL before they take these tests for screening, admission, and immigration purposes. Due to the significant role of these two large-scale test on the lives of participants these test should be constructed carefully and must be valid and reliable as important decisions are being made based on them. Therefore, first the reliability

and validity of these test are discussed in this research. Then the two tests are elaborated in terms of their format, content and grading systems.

4. Reliability and Validity Concepts

Reliability is to the degree to which test scores signify true scores of test-takers. When examining test reliability, the following issues should be considered: a) the degree to which the conditions under which the test is directed are conducive to ideal performance. In a second or foreign language test, any variable other than the language ability that influences test scores is considered a possible cause of measurement error. Errors can limit the reliability and generalizability of scores. b) The psychometric properties, such as difficulty and discrimination indices of test tasks or items, and the internal consistency of test tasks or items; for open-ended types of items, internal consistency is examined in terms of rater agreement in scoring. In the case of objective types of items consistency refers to the extent to which items measuring a particular aspect of the language construct intercorrelate with each other (AERA et al., 1999, as cited in Chalhoub-Deville, & Turner, 2000).

Validity is the most important issue to be considered in developing and evaluating tests. Validity refers to the degree to which the interpretations of test scores are supported by evidence and theory according to what has been proposed by uses of tests. Validity highlights the following issues: a) the validation includes collecting evidence about the relevance of the content covered from the specified construct domain. Moreover, validity stresses both theoretical and empirical evidence to support the interpretation of the test score (Messick, 1996 as cited in Chalhoub-Deville, & Turner, 2000). b) the purpose of the test should guide the test developer with consideration of all aspects of test construction to support the meaning and use of test scores. c) the validation process is an ongoing process. Hence, validation emphasizes a continuing and a systematic research program that documents the interpretations of test scores and support their use according to what has been stated in the test purpose.

4.1 Reliability of IELTS

Manuals of IELTS describe a detailed method for certification of interviewers for the speaking test and raters for the writing parts which needs recertification procedures every two years. This process of training raters is worthy, but the information about rater reliability is lacking. IELTS developers claim that writing and speaking samples are re-rated when facing

inconsistency in the scores and that centres are monitored as part of regular reliability studies conducted by the developers. IELTS publications need to provide more records about raters' reliability the reliability of the instrument, and the scores (Chalhoub-Deville, & Turner, 2000).

Another issue of reliability about the admission tests is that decisions are made on the basis of cut off scores. Some institutions do not use any local investigations for making decision about the required cut-off scores and made their decisions based on to what other institutions are doing. This is not recommended as different institutions would have various requirements and their cut-off scores should reflect their local needs. Although IELTS Annual Review provides information about accepted scores by different institutions, the publication asks test-users to make sure about the appropriateness of getting a particular band scale for admission. Through local research they can determine the suitable cut-off score which is appropriate and required (UCLES, 1999b, as cited in Chalhoub-Deville, & Turner, 2000).

IELTS developers claim to institutions that IELTS scores only show English language proficiency and do not exactly predict about test-takers academic success or failure (UCLES, 1999b). Developers agree that the appropriate approach to making decision about the admission demands institutions to examine various variables other than only English language ability of test-takers.

4.2 Validity of IELTS

According to the IELTS manuals internationalization of the test states about the partnership of the British Council and UCLES, and also the International Development Projects (IDP) Education Australia. Additionally, the test materials are generated in both the UK and Australia which ensures the content of each test reflects an international dimension. Through such collaboration to test development a specific lexical or cultural knowledge of a country that might disadvantage test-takers who do not have that specific knowledge and information will be avoided. In another matter, the appropriateness of using test scores in those contexts for which they are not planned for should also be taken in to consideration. IELTS' scores have been intended primarily for use in the UK and Australia. Although, IELTS is marketing the test in North America. The vital issue to consider is then the comparability of language use in North American academic institutions to the language use in the other two countries. In other words, it is crucial to examine the appropriateness of scores obtained from IELTS as measures of academic language use in North America. Without such investigation it is not easy to

determine what the subsequent test scores in this context mean and how they should be used (Chalhoub-Deville, & Turner, 2000). In 1980s and early 1990s, main changes have been done on IELTS which has changed it from a test with three academic subject modules to its current form. According to Clapham (1996), IELTS had changed to be more of a revision. According to various examinations, comprehensive changes has been made to the IELTS test. Hence, the validation process is not a one-time activity but it is an ongoing process. Developers of IELTS need to continue their validation research and document the properties of their test and the resulting scores.

4.3 Reliability of TOEFL

TOEFL has been described as the ideal psychometric-based English language screening test (Spolsky, 1995). According to Pierce (1994), the process of development of TOEFL items and revision are dictated to psychometric analyses. Moreover, the emphasis on the selected-response item type is to contribute to assure the high reliability standards. Considering the scores, TOEFL is based on a norm referenced approach where results are represented as numeric scores showing the comparative standing of students in relation to a criterion group performance. Annual reports and manuals provides extensive information that presents the reliability of the scores and the performance of test-takers by consideration of various background variables including their native language, gender, and educational level. The program also provides several publications a year within the TOEFL Research Report and Technical Report series that examine a variety of issues to increase researchers' knowledge about and practitioners' use of test scores.

According to the definition of reliability, any variable that affects test scores other than the language ability being measured, is considered a possible source of measurement error. Measurement error can limit the reliability and generalizability of scores. Given the computer format of TOEFL-CBT, a significant issue of consideration is whether test-takers' performance is unpleasantly affected because they are not familiar with the computer medium. Before TOEFL was administered in its CBT format, test takers' experience with computers and its influence on TOEFL performance was investigated (Kirsch et al., 1998; Eignor et al., 1998; Jamieson et al., 1998; Taylor et al., 1998). Approximately 90,000 sample of TOEFL test-takers was surveyed regarding computer familiarity. Findings showed that 16% of the test-takers were have low computer familiarity. Consequently, TOEFL researchers developed a tutorial that test-takers took before starting a TOEFL-CBT.

Furthermore, findings showed no specific differences in scores between computer-familiar and computer-unfamiliar test-takers. Though, as the researchers claimed, more investigation is required to examine the relationship between various background variables and CBT performance. In addition to the mandatory tutorial included in TOEFL-CBT, the instructional CD-ROM TOEFL sampler is provided to increase test takers' familiarity with TOEFL-CBT. CBT represents the next generation of tests and the challenges faced by the TOEFL program which are expected. As more testing organizations use computer medium, the effort of researchers should help solve many of the current issues. However, computer familiarity and should become standard practice for any testing organization planning to use computers as a medium of test delivery (Chalhoub-Deville, & Turner, 2000).

4.4 Validity of TOEFL

Preoccupation with the psychometric qualities of TOEFL helps confirm good testing practices. Yet, it has made the TOEFL resistant to and slow in including changes that might threaten its high reliability standards. While the validity of test scores is damaged when reliability standards are not maintained, reliability records alone cannot make up for insufficient validity. In other words, a strong reliability is not enough to guarantee meaningful interpretations made from TOEFL scores. (Spolsky, 1995).

Anastasi (1986) emphasized on the importance of validity. The author claimed that Validity should be built into the test from the beginning rather than limited to the last stages of its development.

The psychometric-structuralist representation of the language construct as operationalized in TOEFL has changed little over the years. Spolsky (1995) claims that changes made are normally determined by factors such as the maintenance of the test and increasing its marketing advantage rather than in response to theoretical developments and research investigations. Although improvements made to TOEFL-CBT, it is still fall short in its representation of a communicative language construct. For example, the writing prompts, while it needs extended responses, and do not have writers contextualize their essays. In another matter, the focus in the scoring of writing is more on grammatical and textual competence and not on sociolinguistic competence. It should be mentioned that in recent years TOEFL have invested considerable resources to conceptualize, design and construct a new battery of tests. They are considering various integrative models of communicative competence, examining academic

language use, a various types of items, and examining different forms of score reporting to accommodate the needs of the various TOEFL constituents.

5. IELTS vs. TOEFL Format

IELTS is an instrument which is used for university admission or immigration purposes. It is projected to measure both academic and general English language proficiency. IELTS includes six sections which are called modules. All test-takers are administered the same Listening and Speaking modules. While test-takers can decide to take either the General or Academic which have are different in terms of Reading and Writing modules. The General Training modules measure test-takers' language proficiency for the purpose of working in English language environments, undertaking work-related training, or providing language ability for the purpose of immigration. On the other hand, the Academic modules measure test-takers' academic readiness to study at the undergraduate or graduate level where the medium of instruction is English language.

TOEFL is another instrument which intends is to measure the English proficiency of non-native speakers who intend to study in higher learning institutions in the USA and Canada. As with TOEFL scores are increasingly being used by institutions, private organizations, and government agencies in other countries as well (Chalhoub-Deville, & Turner, 2000).

6. IELTS vs. TOEFL Content

The IELTS length is around two hours and thirty minutes. The modules includes the traditional skills: reading, writing, listening, and speaking. The modules have no central theme or topic. IELTS includes various types of tasks and response types within each module. For instance in the IELTS Listening module includes multiple-choice items, taking notes, labelling diagrams, short answers, summarizing information, and matching lists. The four parts within the Listening module are progressively difficult. But the information on how difficulty level is determined is not provided. In order to help familiarize test-takers with the different task formats, IELTS developers provide practice and preparation materials.

The Listening module has various variety excerpts. Excerpts are played only one time. Questions are provided in test booklets. Test-takers are given time to preview the questions. They are asked to respond to the questions while the tape is playing. They write their answers directly in the test booklet. This is needed, because some of the tasks require labelling diagrams,

matching figures with words, etc. Hence, test-takers are given time to transfer their answers to an answer sheet.

The Reading module has text sources from the UK and Australia. According to Clapham (2000), these texts are planned for a non-specialist test-takers. Similar to the Listening module the parts within the Reading module, become increasingly difficult. However, it is not clear how difficulty level is defined.

The Writing module requires two writing samples for both Academic and General training Writing. For the Academic Writing, the first writing task requires test-takers to describe, summarize or explain the information presented in a table, graph, chart or diagram. They might also be asked to describe the data, explain stages of a procedure, the working process of a machine or to define an object or an event. The second task asks test-takers to write an essay on an only given topic. This is a fairly usual procedure on standardized tests which asks test-takers to write an essay in response to a point of view, argument or problem solution (IELTS Exam, 2019). For general writing there are two tasks and topics are of general interest. The first task asks test-takers to write a letter inquiring information, or describing a situation. The style of the letter might be personal, semi-formal or formal. The second task requires test-takers to write an essay by providing their own point of view, argument or problem. (IELTS Exam, 2019)

The IELTS Speaking module is around fifteen minutes interview between an examiner and the test-taker. The interview requires test-takers to describe, and to explain a lot of personal and general interest topics. The interview also includes an elicitation task that is like a role-play. Interviews are recorded for the time that double check is necessary.

Considering the content for TOEFL-CBT, it should be mentioned that, TOEFL was converted from a paper-and-pencil to a CBT in 1998. TOEFL-CBT includes three sections: Listening, Structure/Writing combined, and Reading. Speaking is assessed separately by using the Test of Spoken English (TSE). While Test of Spoken English can be administered along with TOEFL, it is an independent test with different procedures and scheduling. The length of TOEFL CBT takes about four hours without the speaking part. It includes the mandatory tutorial to help test-takers gain knowledge with necessary computer functions and test skills. A main difference between the TOEFL-CBT and other language tests is the computer delivery system and the adaptive algorithm in the Listening and Structure sections. An adaptive test is different from a traditional test due to the reason that an item is selected based on a test-taker's

performance on previous items. In other words, a computer-adaptive test improves the testing condition by targeting each test-taker's ability level.

The Listening section uses visuals that asks test-takers to view a picture and listen at the same time. The listening is played only one time. Test-takers cannot preview the questions, and cannot see them while the listening is being played. They are not also allowed to take notes. Once the listening is finished, the question is heard and both the question and response options are displayed on the screen. In addition to traditional multiple-choice items, the test includes item types that ask test-takers to select two options, match or order objects, and select a visual.

The Structure section includes two types of multiple-choice items which are selecting the option that best completes a sentence and identifying an incorrect option. Both the Listening and the Structure sections are adaptive which means test-takers must answer each question before the next item is administered. Additionally, test-takers cannot return to previous items.

The Reading section includes four to five texts with 10-14 items per text. Test-takers are administered linear sections of reading texts and items. The computer algorithm manages test-takers individualized sets of texts and items which is according to the content and statistical requirements of the test., it is possible for test-takers to return to previous items because no adaptive algorithm is used for this section.

In the writing section, test-takers are requested to construct a response. Test-takers are required to write an essay on a general topic. No information is given to help test-takers contextualize their essay and only one topic is provided. Test-takers can type or use hand writing. The handwritten essays are scanned before they are sent to raters for evaluation.

7. IELTS vs. TOEFL Grading Systems

IELTS scores are between the range of 0 and 9 which are called band score. The score report provides separate band scores for the four modules. Moreover, the average of module scores are gain to develop an overall band score. It is also possible to get 0.5 scores as well (for example, 6.5 or 7.5). Hence, there is a band score for each skill (listening, reading, writing and speaking). The overall band score is the average score of all skills. To make it more clear, if the overall score is an average of 7.25, then the score will be increased to 7.5 and if it is 7.75 then it will be increase to 8, however, if the overall is 7.1 it will go down to 7. Moreover, for IELTS, there is a band score description. The table gives a description of each band score with

information about the level of English required for band scores 5 to 9. IELTS developers do not explain how the Listening and Reading raw scores are converted into band levels. Furthermore, the criteria scales which are employed for Writing and speaking rating are rare, and there is lack of information on how ratings are converted to band scores. However, these information is crucial to test-users and helps in the interpretation of the obtained ratings.

In the case of TOEFL, the grading system for Listening and Structure sections which are adaptive is cumulative. The final section score is determined by item difficulty and the number of items which are answered correctly. The important issue of consideration is that correct answers to more difficult questions has more weight comparing to the correct answers to the easier questions. Item difficulty is assessed with the use of three-parameter item response theory model. Scores in the Reading section are based on the number of correct answers as it is linear not the adaptive system, but they are adjusted for potential inconsistencies in the individualized sets of reading. There are two independent raters that evaluate essays in the Writing section. Essays are rated using a six-point scale. Section scores are converted into scaled scores (ETS, 1998). Scaled section scores across the three tests contribute equally to provide a total scaled score.

8. Discussion

IELTS and TOEFL exams are equal based on their purpose. Although institutions of British colonies use IELTS primarily, it is becoming as ubiquitous as the TOEFL. The two high-stakes test are developed to examine language proficiency of test-takers either if they are intended to attend in higher education or for immigration purposes. Considering the comparative studies on the IELTS and TOEFL-CBT, there are certain similarities and differences that will be discussed in the following paragraphs.

8.1 IELTS and TOEFL Similarities

In terms of IELTS and TOEFL similarities, the first important issue is that both are large-scale test which are required for immigration purpose or admission purpose to colleges or universities. Both are important test with their own rules and policies. For the academic purpose mostly institutions publish their requirement criteria, although some do not mention and the test-takers should decide themselves, where they might not be able to choose between the two tests without appropriate and sufficient knowledge about them. Additionally, both tests are

based on examining test-taker's language ability and skills in reading, speaking, listening and writing skills and they're assessed on the basis of standard scores in all the countries. Moreover, both tests are used to evaluate the minimum level of English skills of test-takers to be accepted in academic institution or get a visa as an immigrant to a country (IELTS vs TOEFL: Similarities and Differences, 2019).

8.2 IELTS and TOEFL Differences

8.2.1 Difference in Speaking

In IELTS test, speaking is like an interview with a real person instead of talking to a machine. The major difference in the speaking of both tests is that in the TOEFL test, the entire speaking part takes place through an automated software. The IELTS speaking test seems to be more feasible than talking to a computer, however, some people think that talking to a machine is a better option to avoid interacting with a stranger and they prefer TOEFL (IELTS vs TOEFL: Similarities and Differences, 2019).

8.2.2. Difference in Writing

The difference between IELTS and TOEFL test with respect to the writing part is that first, in TOEFL test-takers are required to write an essay on a general topic, while IELTS have different tasks for Academic and General tests and for each one there are two types of task which are different from each other. As mentioned previously, in Academic writing the first task asks test-takers to describe a table, graph, chart or diagram and the second task is an essay. In general writing there are again two tasks but with it differs from Academic writing in that the first task asks test-takers to write a letter describing a situation which might be personal, semi-formal or formal and the second task is again an essay. Second, in TOEFL test-takers can type or use hand writing while in IELTS only handwriting is allowed. Some individuals are not really good typists and they prefer handwriting while others find typing easier than handwriting (IELTS vs TOEFL: Similarities and Differences, 2019).

8.2.3 Difference in Listening

The duration of listening session is different for IELTS and TOEFL. In TOEFL the listening session lasts for 60 minutes while IELTS is only 30 minutes long. Additionally Then there's a difference in speakers' accents is observable in IELTS. In the IELTS a variety of accents including USA, Australia, Canada, Wales, Ireland etc. are used in IELTS whereas in TOEFL

listening only standard American accent can be heard. In another matter, in TOEFL, the Listening section uses visuals that asks test-takers to view a picture and listen at the same time. Test-takers cannot preview the questions, and cannot see them while the listening is being played, however, in IELTS they can see the questions. In TOEFL they are not also allowed to take notes, but in IELTS it is possible to take notes in the booklet. In TOEFL In addition to traditional multiple-choice items, the test includes item types that ask test-takers to select two options, match or order objects, and select a visual, while in IELTS different types of tasks are given.

8.2.4 Difference in Reading

Reading session in TOEFL is much longer and includes multiple choice questions only that require a lot of reading which can be boring, whereas reading section in IELTS includes different types of tasks. All in all, IELTS is regarded relatively easier and more engaging than TOEFL mainly because it takes shorter time and involves more interaction with people.

9. Conclusion

Language proficiency scores obtained from IELTS and TOEFL are used to make decisions regarding admission into academic institutions, work-related training, or providing language ability for the purpose of immigration. Therefore, test scores provide critical information. As a result, test developers of high-stakes tests have the responsibility to construct instruments that meet professional standards, to investigate about their instruments and assure about the scores, make test manuals, user guides and research documents available to the public in order to enhance the test-takers knowledge and information about the tests they need to take. Test-users need to be cognizant of the properties of the instruments they employ and ensure appropriate interpretation and use of test scores provided. Test-users also need to search for information necessary to succeed in admission, working or immigration programs as they are based on the obtained scores from these tests. In sum, it should be noted that it is not possible to say which test is the best. Each test might be more appropriate for and matches better to the particular needs of each institution or organization and test-takers should make decision on selection of each test based on their own needs.

References

- Chalhoub-Deville, M., & Turner, C. E. (2000). What to look for in ESL admission tests: Cambridge certificate exams, IELTS, and TOEFL. *System*, 28(4), 523-539.
- Eignor, D., Taylor, C., Kirsch, I., Jamieson, J. (1998). Development of a Scale for Assessing the Level of Computer Familiarity of TOEFL Test Takers (TOEFL Research Report No. 60). Educational Testing Service, Princeton, NJ.
- Jamieson, J., Taylor, C., Kirsch, I., Eignor, D. (1998). Design and Evaluation of a Computer-based TOEFL Tutorial (TOEFL Research Report No. 62). Educational Testing Service, Princeton, NJ.
- IELTS Exam. (2019). Retrieved from <https://www.ielts.org>
- IELTS vs TOEFL: Similarities and Differences. (2019). Retrieved from basic-concept: <http://www.basic-concept.com/c/ielts-vs-toefl-similarities-and-dissimilarities>
- Kirsch, I., Jamieson, J., Taylor, C., Eignor, D. (1998). Computer Familiarity among TOEFL Test Takers (TOEFL Research Report No. 59). Educational Testing Service, Princeton, NJ.
- Pierce, B., 1994. The test of English as a foreign language: developing items for reading comprehension. In: Hill, C., Parry, K. (Eds.), *From Testing to Assessment: English as an International Language*. Longman, New York, pp. 39-60.
- Spolsky, B., 1995. *Measured Words*. Oxford University Press, Oxford.
- Taylor C., Jamieson, J., Eignor, D., Kirsch, I. (1998). The Relationship Between Computer Familiarity and Performance on Computer-based TOEFL Test Tasks (TOEFL Research Report No. 61). Educational Testing Service, Princeton, NJ.
- TOEFL Program History (2017). Retrieved from: https://www.ets.org/s/toefl/pdf/toefl_ibt_insight_slv6.pdf
- Zahedi, K., & Shamsaee, S. (2012). Viability of construct validity of the speaking modules of international language examinations (IELTS vs. TOEFL iBT): evidence from Iranian test-takers. *Educational Assessment, Evaluation and Accountability*, 24(3), 263-277.
- UCLES, 1999b. *IELTS Annual Review: (1998/1999)*. University of Cambridge Local Examination Syndicate, the British Council, and IDP Education Australia, Cambridge.