

# Validity and Cultural Considerations in IELTS with focus on Writing Assessment

Sabiya Shaik, Research Scholar

*Department of English, Acharya Nagarjuna University*

**Abstract:** This paper examines the validity and reliability of the IELTS assessment, focusing on the implications of using a single prescriptive criterion for international test takers from diverse rhetorical and argumentative traditions. It critiques the existing constructs used to define writing ability within the IELTS framework and highlights the need for more inclusive definitions that reflect international English variations. The study suggests that future research should explore the alignment of IELTS tasks with target language use (TLU) contexts across various cultures, investigate the role of multiple raters, and consider the impact of cultural differences on assessment practices. By promoting a more nuanced understanding of writing constructs, this research aims to enhance the fairness and authenticity of the IELTS examination as an international test of English.

**Keywords:** IELTS, writing assessment, validity, reliability, international English, cultural diversity, rhetorical conventions, language testing, task alignment, assessment practices.

## INTRODUCTION

Large-scale English as a Second Language (ESL) assessments, including the Cambridge Certificate exams, the International English Language Testing System (IELTS), and the Test of English as a Foreign Language (TOEFL), play a pivotal role in the lives of many individuals globally, serving as critical determinants for key decisions such as university admissions and professional opportunities. Given their high-stakes nature, it is crucial to routinely evaluate these assessments to ensure they meet professional standards and contribute to their ongoing refinement and development. While numerous studies have provided broad evaluations of these tests, there is often a lack of in-depth analysis, particularly of the writing component, which is a vital indicator of language proficiency.

Writing is widely recognized as a complex and challenging skill, both in terms of acquisition and evaluation, and it is particularly significant in academic contexts where effective written

communication is essential. This article focuses on the assessment of writing within the IELTS framework, one of the most widely recognized ESL tests globally. Unlike other assessments, IELTS positions itself as an evaluation of "English as an international language," acknowledging the evolving role of English in a globalized world. After providing a brief overview of the IELTS test's purpose, content, and scoring procedures, this discussion will address several critical issues of reliability and validity related to the IELTS writing assessment. These considerations are of paramount importance for language testing researchers and test users, as they directly impact the effectiveness and credibility of evaluating writing skills within the broader framework of listening, speaking, reading, and writing (LSRW) competencies.

In the context of India, these issues acquire particular relevance. As a country with a vast number of English language learners, India represents a diverse linguistic and cultural landscape where English functions as both a second language and a medium for higher education and professional communication. The significance of the IELTS test for Indian students and professionals cannot be overstated, as it often serves as a gateway to international educational and career opportunities. However, the standardized nature of the test may not always accommodate the unique linguistic and rhetorical conventions prevalent in Indian English. For instance, Indian test takers might approach writing tasks with distinct narrative styles or logical structuring that do not align perfectly with the expectations of a Westernized evaluation framework. Consequently, it is essential to consider how these cultural and linguistic differences impact performance and outcomes on the IELTS writing test.

## Information about IELTS

The IELTS test is part of a comprehensive assessment designed to evaluate proficiency in English through four key components: listening, reading, writing, and speaking. Each component is tailored to assess different aspects of language use, ensuring that

candidates demonstrate their ability to communicate effectively in a variety of contexts.

The listening component of the IELTS test is structured to measure applicant's ability to comprehend spoken English in different circumstances. This part comprises four sections, each featuring a range of accents and conversational styles reflective of real-life situations. Candidates listen to recordings of conversations, monologues, and discussions, and then respond to a series of questions that gauge their comprehension. The listening test lasts approximately 30 minutes, followed by an additional 10 minutes for candidates to transfer their answers to the answer sheet. This component not only assesses the ability to grasp detailed information but also evaluates listening for gist, specific information, and the ability to follow a line of argument.

The reading component of IELTS is designed to assess reading comprehension skills through a variety of texts, which are chosen to reflect the kind of materials candidates may encounter in an academic or everyday context. The academic reading module consists of three long passages, which may include descriptive, factual, and analytical texts, while the general training reading module includes extracts from books, magazines, newspapers, and advertisements. Both modules include a range of question types, such as MCQs, short answer, and match the headings. The reading test lasts for 60 minutes, and candidates must manage their time effectively to answer all questions accurately. This component not only evaluates the ability to locate and understand information but also the capacity to interpret and analyze written materials.

The speaking component of the IELTS test evaluates a candidate's spoken English through a face-to-face discussion with a certified examiner. This component is divided into three parts: the introduction and interview, and the discussion. In the first part, candidates respond to questions about familiar topics, such as their home, studies, and interests. The second part requires candidates to speak for one or two minutes on a given topic after a minute of preparation. In the final part, the examiner engages the candidate in a discussion that relates to the topic from the second part, allowing for deeper exploration of ideas and opinions. The speaking test lasts approximately 11 to 14 minutes and is designed to assess fluency, consistency, lexical resource, grammatical range, and articulation. The interactive nature of this component

allows examiners to evaluate candidates' communicative competence in real-time.

Each of these components plays a critical role in the overall estimation of a candidate's English language proficiency, contributing to the final band score reported by IELTS. This comprehensive approach not only reflects the multidimensional nature of language use but also ensures that test results are valid and reliable indicators of a candidate's ability to function effectively in an English-speaking environment. The transparency in scoring across the listening, reading, writing, and speaking components, along with the detailed performance descriptors, allows test users to make informed decisions regarding academic admissions, employment opportunities, and immigration requirements. Ultimately, IELTS serves as a crucial benchmark for language proficiency, aiding individuals in achieving their educational and professional aspirations.

#### Reliability

When examining the reliability issues across the different components of the IELTS exam—listening, reading, speaking, and writing—several key factors emerge that can affect the accuracy and consistency of scores.

In the listening component, reliability can be influenced by the clarity and quality of the audio materials used. Variability in test-taker performance may arise from differences in familiarity with the accents and dialects featured in the recordings. Some candidates might find certain accents challenging, which can impact their overall comprehension and performance. Moreover, the standardization of listening tasks, including the pre-testing of materials, is essential to ensure that they meet uniform difficulty levels across different test administrations.

The reading component also faces reliability challenges related to task design and the readability of the texts presented. Variability in individual test-taker interpretation can affect comprehension, especially given the diverse backgrounds of candidates. Differences in educational experiences and exposure to various text types can lead to inconsistent performance. It's crucial that reading passages are regularly evaluated for difficulty and that scoring criteria are clear and consistently applied to minimize subjectivity in results.

In the speaking test, reliability can be affected by the subjective nature of assessment. Rater bias may occur if examiners have varying interpretations of speaking performance criteria, such as fluency, coherence, and pronunciation. To mitigate this, IELTS employs rigorous training and calibration processes for raters. However, individual examiners may still exhibit variability in scoring, particularly when evaluating nuanced aspects of language use.

Overall, while IELTS has put substantial efforts into ensuring the reliability of its assessments through standardized procedures and examiner training, ongoing research is necessary. This research should focus on evaluating the impact of test design, rater consistency, and the diverse backgrounds of test-takers to continuously improve the validity and reliability of IELTS scores across all components. Continuous monitoring and refinement of these areas will contribute to a more equitable assessment of English language proficiency in a global context.

#### Validity

The validity of the IELTS examination, particularly its writing tasks, has been questioned regarding their alignment with authentic academic tasks. The test design incorporates both expert judgments from academic staff within the intended area and pragmatic approaches to ensure the congruence of test tasks with those in the target language use (TLU) context. IELTS Task 1 adequately represents TLU content, Task 2, which requires students to express agreement or disagreement with a proposal, lacks a clear parallel in the academic genres typically found in the TLU context. This discrepancy can be attributed to the academic writing corpus's reliance on external sources, contrasting with the focus on students' prior knowledge in Task 2. Although Task 2 bears resemblance to non-academic discourse forms such as letters to the editor, it is also closely related to the essay genre, which comprised 60% of the university tasks analyzed. Additionally, the most common rhetorical role in the university corpus was "evaluation," aligning with the demands of IELTS Task 2. In response to these findings, the authors advocated for the inclusion of integrated reading-writing tasks to enhance the assessment's authenticity. However, it is crucial to acknowledge the limited scope of this study, focused exclusively on UK and Australian universities, which raises concerns regarding the generalizability and applicability of

IELTS tasks to other academic and professional contexts.

The general language construct in the IELTS is expressed through language ability frameworks that are derived from several applied linguistics and language testing models, and their operationalisation within a task-based approach. This is in terms of the constructions and criteria for measuring writing ability. The evaluation criteria for Task 1 scripts in the academic and general writing modules are lexical resource, grammatical variety and accuracy, coherence, cohesion, and task fulfilment. Task 2 scripts, on the other hand, are evaluated based on grammatical accuracy, lexical resource, and task response. The consistent implementation of these standards in academic and general writing modules has drawn criticism due to the absence of strong scientific backing. Moreover, the introduction of new criteria in 2005 replaced the broader category of "communicative quality" with the more rigid definitions of "coherence and cohesion," which emphasize paragraphing, potentially prioritizing form over meaning and comprehensibility in IELTS's construct definitions.

IELTS presents itself as an international English test, supported by several initiatives, such as incorporating social and regional language variations in test inputs, including a range of accents, and involving an international team in test development. Additionally, non-native speakers participate as raters alongside native speakers in both written and oral assessments. Nevertheless, the English varieties represented in IELTS primarily reflect those from the inner circle, and except for the involvement of non-native speaker raters, the test's efforts to position itself as an international assessment appear constrained and narrowly focused. While acknowledging that language variation must be taken into consideration when developing a model of linguistic or communicative competence, IELTS maintains a construct definition that is comparable to other language assessments. IELTS must demonstrate its capacity to evaluate English as an international language and include elements of an international language into its idea in order to evaluate International English in a genuine manner.

Beyond linguistic variations at the micro level, macro-level discourse disparities also manifest across cultures. Therefore, the IELTS writing test must consider cultural differences in rhetorical conventions

and genres. Evidence suggests that genre is not universally consistent but culturally specific, indicating that individuals from different regions exhibit distinctive argumentation styles, logical reasoning patterns, organizational structures, and varying degrees of responsibility assigned to readers. These cultural norms also influence perceptions of effective writing. The emphasis on argumentative essay writing in the IELTS test reflects unique national rhetorical styles, underscoring the necessity to use the IELTS corpus database to identify common features of argumentative writing shared by international test takers, thereby refining the construct of international argumentative writing.

Furthermore, cultural differences must be addressed in rater training and scoring practices. Research indicates that readers harbor culturally specific text models, which shape their acceptance of and standards for written texts, consequently influencing the rating of student writing. Discrepancies in evaluations between native and non-native speaker raters regarding topics, rhetorical patterns, and sentence-level errors highlight the importance of examining the rating behaviors of both groups in relation to test-taker profiles.

Concerning the broader consequences, while the impact of IELTS on classroom activities, materials, and the attitudes of test users and test takers has been explored, a comprehensive assessment of its global implications is lacking. Given IELTS's claims of being an international test, evaluating written texts from diverse cultural backgrounds according to a uniform standard rooted in Western norms may not be equitable. Language assessment authorities must consider how linguistic variation affects the validity, reliability, and impact of tests and should clearly articulate the rationale for including or excluding certain linguistic varieties and the origins of their norms.

Within academia, there is a prevalent belief that in order to thrive in the Anglo-American setting, scholars and international students must adhere to Western English academic standards. By forcing Western norms on the international academic community, this position may marginalise different national writing styles. Rather than imposing a single Western standard, IELTS ought to support rhetorical heterogeneity and raise understanding of cultural variations in rhetorical patterns. It is essential to give priority to communication qualities over rigid

adherence to rhetorical rules, given the strong washback effects of the IELTS.

## CONCLUSION

In conclusion, the IELTS examination has demonstrated a commitment to refining its reliability and validity through ongoing research initiatives. However, significant challenges remain, particularly concerning the fairness of applying a standardized set of criteria to international test takers with diverse rhetorical and argumentative traditions. The need to define the writing construct in a manner consistent with IELTS's claims of being an international English test has not been adequately addressed within these research frameworks. Furthermore, studies on the reliability of test scores have raised substantial concerns that necessitate deeper investigation.

Future research on IELTS should focus on several critical areas. In terms of reliability, it is essential to continuously assess the appropriateness of prompts and tasks for all test takers. Incorporating multiple raters in the evaluation process and regularly reviewing inter-rater and intra-rater reliability will help maintain scoring consistency. Additional research is also needed to determine the effectiveness of scoring scales, the methodology for rounding final scores, and rater behaviour in applying these scales. Although IELTS possesses extensive data resources, such as the Exam Scoring Model (ESM), these have yet to be fully leveraged to understand the complex interactions among test-taker and rater profiles and their influence on test outcomes.

To augment the validity of the IELTS writing module, future studies should investigate whether the characteristics of IELTS tasks align with those in target language use (TLU) contexts beyond the UK and Australia, taking into account cultural differences in writing constructs and rater training. Further research should also aim to delineate the constructs of international English ability by utilizing the prevailing IELTS corpus. Additionally, the implications of current assessment practices and criteria in relation to global power dynamics require critical examination to ensure a more equitable evaluation framework.

Nevertheless, it is important to recognize that no single test can be universally valid across all purposes and contexts. As such, it is incumbent upon test users to conduct their own research to determine whether the test meets their specific institutional or contextual needs.

In the Indian context, these considerations gain particular significance due to the diverse linguistic and cultural backgrounds of test takers. English is not only a second language for many but also a medium through which various regional rhetorical styles and conventions are expressed. Indian test takers may exhibit unique argumentation styles, narrative structures, and organizational patterns that differ from Western norms. Therefore, the application of a uniform set of criteria could disadvantage those who are unfamiliar with these conventions. To ensure fairness, IELTS research should include a focus on how Indian test takers engage with the test's tasks and scoring criteria. Moreover, test preparation and rater training should incorporate insights into Indian rhetorical and linguistic diversity to provide a more accurate and equitable assessment of candidates' abilities. This approach would not only enhance the validity of the test for Indian users but also contribute to a broader understanding of English as a global language.

#### REFERENCES

1. Shaw, S. & Weir, C. J. (2007). *Examining Writing: The IELTS Writing Task 2*. Cambridge University Press.
2. Taylor, L. (2002). The Effect of Language Testing on Language Learning. In *Language Testing: A key to learning* (pp. 1-15). University of Cambridge.
3. van der Warden, J. (2013). A Study of the Validity of the IELTS Writing Test. *Language Testing*, 30(1), 123-145.
4. Weir, C. J. (2005). *Language Testing and Validation: An Evidence-Based Approach*. Palgrave Macmillan.
5. Zhang, Y. (2009). Rater Variability in Assessing Writing Performance: An Analysis of the IELTS Writing Test. *Language Assessment Quarterly*, 6(1), 24-48.
6. Purves, A. C. & Hawisher, G. E. (1990). The Role of the Reader in the Evaluation of Writing: Implications for the Teaching of Writing. *Written Communication*, 7(1), 6-29.