English Language Assessment: Innovations, Validity, and Reliability

Dr. Asha Kiran

Department of English, Banaras Hindu University, Varanasi

ABSTRACT

This paper explores the contemporary landscape of English language assessment, focusing on innovations, validity, and reliability within the field. As English continues to serve as a global lingua franca, the demand for robust and accurate language assessment tools has intensified. Innovations in assessment methodologies, driven by advancements in technology and pedagogical approaches, have reshaped the landscape, offering diverse means to measure language proficiency. This paper reviews various innovative assessment formats, including computer-adaptive testing, performance-based assessment, and multimodal approaches, highlighting their potential to provide more authentic and comprehensive evaluations of language skills. Moreover, it examines the critical concepts of validity and reliability in language assessment, emphasizing their indispensable role in ensuring the accuracy and fairness of assessment outcomes. Through a synthesis of recent research and theoretical frameworks, this paper elucidates the complex interplay between validity, reliability, and innovation in English language assessment, offering insights into best practices and future directions in the field.

Keywords: English language assessment, Innovations, Validity, Reliability, Assessment methodologies.

INTRODUCTION

English language proficiency assessment plays a pivotal role in various contexts, including education, employment, immigration, and international communication. As English continues to assert its dominance as a global lingua franca, the need for effective, reliable, and innovative assessment tools becomes increasingly pronounced. This introduction sets the stage by highlighting the significance of English language assessment in contemporary society and outlines the key themes that will be explored in the paper: innovations, validity, and reliability. Firstly, the introduction provides an overview of the evolving landscape of English language assessment, contextualizing the growing demand for accurate and comprehensive evaluation of language skills. It emphasizes the diverse purposes for which English proficiency assessment is utilized, ranging from academic admissions and language certification to workforce recruitment and immigration eligibility.

Secondly, the introduction introduces the concept of innovation in language assessment, acknowledging the transformative impact of technology and pedagogical advancements on assessment methodologies. It underscores the importance of embracing innovative approaches to assessment, such as computer-adaptive testing, performance-based assessment, and multimodal evaluation, to better capture the complexities of language proficiency. Thirdly, the introduction discusses the fundamental principles of validity and reliability in language assessment, elucidating their significance in ensuring the fairness, accuracy, and authenticity of assessment outcomes. It emphasizes the need for assessments to be valid, measuring what they purport to measure, and reliable, yielding consistent results over time and across different contexts.

Overall, this introduction serves to orient the reader to the multifaceted nature of English language assessment, laying the groundwork for a comprehensive exploration of innovations, validity, and reliability within the field. Through a critical analysis of recent research and theoretical frameworks, this paper aims to deepen understanding and provoke discussion on best practices and future directions in English language assessment.

LITERATURE REVIEW

English language assessment has been a subject of extensive research and scholarly inquiry, reflecting its central role in various domains of society. This literature review synthesizes key findings and insights from existing literature to provide a comprehensive understanding of innovations, validity, and reliability in English language assessment.

Innovations in Assessment Methodologies: Innovations in English language assessment have been driven by advancements in technology, pedagogy, and assessment theory. Computer-adaptive testing (CAT), for instance, has gained prominence for its ability to tailor test content to the proficiency level of individual test takers, thereby optimizing assessment efficiency and accuracy (Weigle, 2017). Performance-based assessment approaches, such as portfolios, projects, and oral presentations, have emerged as alternatives to traditional standardized tests, allowing for more authentic and holistic evaluation of language skills (Weir, 2005). Furthermore, multimodal assessment methods, which integrate various modes of communication (e.g., written, spoken, visual), offer new possibilities for assessing complex language competencies (Pachler et al., 2010).

Validity in Language Assessment: Validity, as a central concept in assessment theory, pertains to the extent to which an assessment instrument measures what it purports to measure (Messick, 1989). In the context of English language assessment, validity encompasses various aspects, including content validity (alignment of test content with language proficiency objectives), construct validity (degree to which the test measures the intended construct of language proficiency), and criterion-related validity (relationship between test scores and external criteria, such as academic success or communicative ability) (Bachman & Palmer, 2010). Ensuring validity in language assessment involves rigorous test development processes, validation studies, and ongoing evaluation of test performance (Shohamy, 2013).

Reliability and Fairness in Assessment: Reliability refers to the consistency and stability of assessment results over time and across different contexts (Bachman & Palmer, 2010). Reliability is essential for making valid inferences about test takers' language proficiency levels and for ensuring fairness in assessment outcomes. Various sources of measurement error, such as test administration conditions, scoring procedures, and test-taker characteristics, can impact the reliability of language assessments (Kane, 2006). Therefore, efforts to enhance reliability often involve standardization of test administration protocols, training of raters, and the use of statistical techniques to estimate and minimize measurement error (Bachman & Palmer, 2010).

Challenges and Future Directions: Despite significant progress in English language assessment, challenges persist, including issues related to test fairness, accessibility, and cultural bias (Alderson, 2005). Moreover, the rapid evolution of technology and the increasing diversity of language learners pose ongoing challenges for assessment developers and practitioners (Chapelle & Douglas, 2006). Future directions in English language assessment may involve further integration of technology, development of innovative assessment formats, and greater attention to the socio-cultural dimensions of language testing (Douglas, 2010).

In conclusion, this literature review highlights the dynamic nature of English language assessment, characterized by ongoing innovation, validation, and reliability efforts. By critically engaging with existing literature, this review provides valuable insights into current trends, challenges, and future directions in the field of English language assessment.

THEORETICAL FRAMEWORK

The theoretical framework underpinning this study draws upon key concepts and frameworks from the fields of language assessment, psychometrics, and educational measurement. The integration of these theoretical perspectives provides a comprehensive lens through which to examine innovations, validity, and reliability in English language assessment.

Classical Test Theory (CTT): Classical Test Theory serves as a foundational framework for understanding the psychometric properties of assessments, including reliability and validity. According to CTT, a test score consists of a true score component and an error component, with reliability quantifying the proportion of true score variance in observed scores (Crocker & Algina, 1986). CTT provides insights into traditional reliability estimates, such as internal consistency (e.g., Cronbach's alpha) and test-retest reliability, as well as validity coefficients derived from correlations with external criteria.

Item Response Theory (IRT): Item Response Theory offers a sophisticated framework for assessing test item performance and calibrating item difficulty and discrimination parameters. Unlike CTT, which focuses on test-level properties, IRT models individual item characteristics and the probability of a correct response as a function of test takers' latent trait levels (Embretson & Reise, 2000). IRT enables the development of computer-adaptive tests and facilitates the estimation of item and test information functions, which inform decisions about item selection, test length, and score interpretation.

Validity Frameworks: Validity frameworks, such as the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014) and Messick's (1989) framework, guide the evaluation of

validity evidence across multiple sources and domains. Messick's framework conceptualizes validity as a unitary concept comprising five sources of validity evidence: content, substantive, structural, generalization, and external. This framework emphasizes the importance of interpreting test scores within a broader context and considering the consequential validity of assessment interpretations.

Socio-cultural Perspectives: Socio-cultural perspectives highlight the influence of sociocultural factors on language learning and assessment practices (Lantolf & Thorne, 2006). These perspectives emphasize the dynamic interplay between language, culture, and identity in shaping language assessment practices and outcomes. Socio-cultural theories inform understandings of fairness, equity, and inclusivity in assessment, prompting considerations of the cultural relevance of assessment content, the accessibility of assessment accommodations, and the validity of assessment interpretations for diverse language learners.

Innovative Assessment Approaches: Theoretical frameworks related to innovative assessment approaches, such as performance-based assessment (Weigle, 2002) and dynamic assessment (Lantolf & Poehner, 2004), inform the design and implementation of alternative assessment formats that capture the complexities of language use and development. These frameworks emphasize the importance of authenticity, interaction, and contextualized language tasks in assessing language proficiency and promoting language learning.

By integrating these theoretical perspectives, this study seeks to provide a robust analytical framework for examining innovations, validity, and reliability in English language assessment. This framework facilitates a nuanced understanding of assessment practices, challenges, and implications for language learners, educators, and policymakers. Additionally, it guides the development of research questions, data collection instruments, and analytical strategies to address the study's objectives effectively.

RECENT METHODS

Computer-Adaptive Testing (CAT): Computer-adaptive testing (CAT) has emerged as a popular method for individualized and efficient assessment of language proficiency. CAT dynamically selects test items based on the test taker's responses, presenting items that are most informative for estimating the test taker's proficiency level (Chapelle & Douglas, 2006). This adaptive approach allows for precise estimation of proficiency levels with fewer items, reducing test duration and minimizing test anxiety (Weir, 2005). CAT systems often use item response theory (IRT) models to calibrate item parameters and adaptively select items that match the test taker's ability level (van der Linden & Glas, 2010).

Performance-Based Assessment: Performance-based assessment methods focus on evaluating language proficiency through real-world tasks and activities that simulate authentic language use situations (Weigle, 2002). These methods include tasks such as oral presentations, role-plays, portfolio assessments, and project-based assessments, which require test takers to demonstrate their language skills in contextually meaningful ways (Fulcher, 2010). Performance-based assessments offer advantages in terms of authenticity, ecological validity, and alignment with communicative language teaching principles (Brown, 2004). They provide insights into test takers' ability to use language in real-life situations, beyond mere knowledge of grammar and vocabulary (Weigle, 2002).

Multimodal Assessment: Multimodal assessment approaches integrate multiple modes of communication, including written, spoken, visual, and interactive elements, to capture the complexity of language use and expression (Pachler et al., 2010). These approaches recognize that language proficiency extends beyond linguistic competence to include socio-pragmatic and discourse-related skills (Kress & van Leeuwen, 2001). Multimodal assessments may involve tasks such as creating multimedia presentations, participating in online discussions, or analyzing and interpreting visual texts (Levy & Stockwell, 2006). By engaging test takers in diverse modes of communication, multimodal assessments provide a more comprehensive and authentic representation of language proficiency (Kress & van Leeuwen, 2001).

Integrated Skills Assessment: Integrated skills assessment combines the evaluation of multiple language skills (e.g., reading, writing, listening, speaking) within a single assessment task or test format (Bachman & Palmer, 2010). Unlike traditional discrete-skill tests, which assess language skills in isolation, integrated skills assessments require test takers to integrate and apply their language skills in integrated tasks that mirror real-world communication demands (Weir, 2005). For example, test tasks may involve listening to a lecture, reading a related text, and then writing a summary or giving a presentation on the topic (Bachman & Palmer, 2010). Integrated skills assessments promote the development of communicative competence and provide a more authentic measure of overall language proficiency (Bachman & Palmer, 2010).

Task-Based Assessment: Task-based assessment focuses on evaluating language proficiency through the performance of communicative tasks that reflect real-world language use situations (Nunan, 2004). Test tasks are designed to elicit language use in meaningful contexts, such as problem-solving tasks, information-gap activities, and decision-making simulations (Ellis, 2003). Task-based assessments emphasize the functional use of language for communication purposes and provide opportunities for test takers to demonstrate their ability to negotiate meaning, manage discourse, and achieve communicative goals (Nunan, 2004).

These recent methods represent innovative approaches to English language assessment that aim to address the limitations of traditional assessment formats and better align with current understandings of language learning and teaching. By incorporating principles of authenticity, validity, and reliability, these methods offer valuable insights into test takers' language proficiency and contribute to more effective language assessment practices.

PROPOSED METHODOLOGY

To investigate innovations, validity, and reliability in English language assessment, a mixed-methods research design will be employed. This methodology integrates both quantitative and qualitative data collection and analysis techniques to provide a comprehensive understanding of the research questions. The proposed methodology consists of the following components:

Quantitative Phase:
[1]. **Sampling**: A stratified random sampling technique will be used to select a diverse sample of English language learners representing different proficiency levels, ages, backgrounds, and contexts.
[2]. **Data Collection**: Quantitative data will be collected through standardized English language proficiency tests, such as TOEFL or IELTS, and innovative assessment methods, such as computer-adaptive testing (CAT) or performance-based assessments. Test scores and demographic information will be obtained from participants.
[3]. **Data Analysis**: Descriptive statistics (e.g., means, standard deviations) will be used to summarize test scores and demographic characteristics of the sample. Inferential statistics (e.g., correlation analysis, regression analysis) will be conducted to examine relationships between test scores, demographic variables, and assessment methods.

Qualitative Phase:
[1]. **Sampling**: Purposeful sampling will be used to select a subset of participants from the quantitative phase for in-depth interviews or focus group discussions. Participants will be selected based on their performance on the assessments and their willingness to participate in qualitative data collection.
[2]. **Data Collection**: Qualitative data will be collected through semi-structured interviews or focus group discussions, exploring participants' experiences, perceptions, and attitudes towards different assessment methods, their perceived validity and reliability, and their preferences for assessment formats.
[3]. **Data Analysis**: Qualitative data will be analyzed using thematic analysis to identify recurring themes, patterns, and insights related to innovations, validity, and reliability in English language assessment. Transcripts will be coded and categorized to identify key findings and interpretations.

[1]. **Integration of Data**:
[2]. Triangulation: Quantitative and qualitative data will be triangulated to corroborate findings and provide a more comprehensive understanding of the research questions.
[3]. Comparative Analysis: Quantitative and qualitative findings will be compared and contrasted to identify consistencies, discrepancies, and complementary insights.
[4]. Interpretation: Integrated data analysis will be used to generate overarching interpretations and conclusions regarding innovations, validity, and reliability in English language assessment.

Ethical Considerations:
[1]. Informed Consent: Participants will be provided with clear information about the study's purpose, procedures, risks, and benefits, and their voluntary participation will be obtained through informed consent.
[2]. Confidentiality: Participants' confidentiality and anonymity will be ensured by assigning pseudonyms to participants and securely storing and anonymizing data.
[3]. Research Ethics: The study will adhere to ethical guidelines and standards for research involving human participants, including obtaining approval from relevant institutional review boards (IRBs) or ethics committees.

By employing a mixed-methods approach, this study aims to provide a comprehensive and nuanced understanding of innovations, validity, and reliability in English language assessment, thereby informing best practices and future directions in the field.

COMPARATIVE ANALYSIS

The comparative analysis will involve examining and contrasting the findings from the quantitative and qualitative phases of the study to identify consistencies, discrepancies, and complementary insights regarding innovations, validity, and reliability in English language assessment. Here's how the comparative analysis may unfold:

Innovations in English Language Assessment:
[1]. Quantitative Phase: The quantitative analysis may reveal statistical trends regarding the performance of participants on traditional standardized tests versus innovative assessment methods (e.g., CAT, performance-based assessments). It may provide insights into the effectiveness of these methods in capturing different aspects of language proficiency.
[2]. Qualitative Phase: The qualitative analysis may uncover participants' perceptions of the strengths and limitations of traditional versus innovative assessment methods. Participants may offer insights into their preferences for specific assessment formats and their perceived authenticity and relevance.

Validity of English Language Assessment:
[1]. Quantitative Phase: Quantitative analyses may examine correlations between test scores obtained from different assessment methods and external criteria (e.g., academic success, communicative ability). These analyses may provide evidence of criterion-related validity for various assessment formats.
[2]. Qualitative Phase: Qualitative data may shed light on participants' interpretations of validity in English language assessment. They may discuss their understanding of what constitutes a valid assessment, how well different assessment methods measure language proficiency, and the extent to which assessments align with their learning goals.

Reliability of English Language Assessment:
[1]. Quantitative Phase: Quantitative analyses may assess the internal consistency reliability of traditional standardized tests and innovative assessment methods using measures such as Cronbach's alpha. Additionally, test-retest reliability may be examined for both types of assessments.
[2]. Qualitative Phase: Participants' experiences with different assessment methods may provide qualitative insights into the reliability of English language assessments. They may discuss factors that influence their performance consistency across assessments and their perceptions of the fairness and consistency of scoring procedures.

Integration of Findings:
[1]. Triangulation: The comparative analysis will involve triangulating quantitative and qualitative findings to corroborate or challenge each other. Consistencies across data sources may strengthen interpretations, while discrepancies may prompt further investigation.
[2]. Complementary Insights: The integration of quantitative and qualitative data will provide a more holistic understanding of innovations, validity, and reliability in English language assessment. Quantitative analyses may provide statistical trends, while qualitative insights may offer nuanced explanations and contextualized perspectives.

Implications and Recommendations:
Based on the comparative analysis, implications and recommendations for English language assessment practice, policy, and research will be derived. These may include suggestions for the integration of innovative assessment methods into existing practices, strategies for enhancing validity and reliability, and considerations for addressing learners' diverse needs and preferences.

Through the comparative analysis, this study aims to contribute to a deeper understanding of English language assessment and inform evidence-based practices that support learners' language development and proficiency evaluation.

SIGNIFICANCE OF THE TOPIC

The significance of the topic of innovations, validity, and reliability in English language assessment lies in its implications for various stakeholders, including language learners, educators, policymakers, and assessment developers. Here are several reasons why this topic is significant:

Impact on Language Learning Outcomes: English language assessment plays a crucial role in shaping language learning experiences and outcomes. The validity and reliability of assessment methods directly influence the accuracy

of proficiency measurements, which, in turn, inform learners' progress tracking, instructional planning, and goal setting.

Equitable Access to Opportunities: Fair and valid English language assessment is essential for ensuring equitable access to educational and professional opportunities for diverse populations. Language assessments often serve as gatekeepers for admission to academic programs, employment opportunities, and immigration processes. Therefore, addressing issues of validity and reliability in assessment is critical for promoting social justice and equal access to opportunities.

Quality Assurance in Education: Valid and reliable assessment practices are central to ensuring the quality and integrity of English language education programs. Educational institutions rely on assessment data to evaluate program effectiveness, identify areas for improvement, and make evidence-based decisions about curriculum development, instructional strategies, and resource allocation.

Innovation and Advancements in Assessment: The topic of innovations in English language assessment reflects ongoing efforts to leverage technological advancements, pedagogical insights, and assessment theory to enhance assessment practices. Exploring innovative assessment methods and approaches can lead to the development of more effective, efficient, and authentic assessment tools that better align with contemporary language learning theories and practices.

Global Significance of English Proficiency: English proficiency has become increasingly important in the globalized world, with English serving as a primary medium of communication in various domains, including business, academia, diplomacy, and entertainment. Therefore, assessments of English language proficiency have implications not only at the individual and institutional levels but also at the societal and international levels.

Policy and Decision-Making Implications: Policymakers and education stakeholders rely on assessment data to inform policy decisions related to language education, immigration, and workforce development. Understanding the validity and reliability of assessment methods is essential for designing effective policies, implementing equitable practices, and allocating resources strategically.

Research and Professional Development: Engaging with the topic of English language assessment provides opportunities for research, professional development, and knowledge dissemination within the fields of language testing, psychometrics, linguistics, and education. Research on innovations, validity, and reliability contributes to advancing theoretical understanding, refining assessment practices, and informing evidence-based interventions.

In summary, the significance of the topic of innovations, validity, and reliability in English language assessment lies in its far-reaching implications for language learners, educators, policymakers, and society at large. By addressing issues related to assessment quality, equity, and effectiveness, this topic contributes to promoting inclusive and effective language learning environments and fostering meaningful linguistic and communicative competencies.

LIMITATIONS & DRAWBACKS

While the topic of innovations, validity, and reliability in English language assessment holds considerable significance, it's important to acknowledge several limitations and drawbacks associated with research and practice in this area:

Complexity of Language Proficiency: English language proficiency is a multifaceted construct encompassing various skills, such as reading, writing, listening, and speaking, as well as socio-pragmatic and discourse competencies. Assessing these diverse aspects of language proficiency comprehensively poses challenges, as no single assessment method can fully capture the complexity of language use and communication.

Cultural and Contextual Factors: English language assessment is influenced by cultural and contextual factors that may impact test performance and validity. Assessment tasks, content, and scoring criteria may not always be culturally appropriate or relevant for diverse test takers, leading to potential biases and unfairness in assessment outcomes.

Validity Erosion: The validity of English language assessments may be compromised by factors such as test preparation strategies, coaching programs, and test format familiarity. Test takers may engage in test-taking behaviors aimed at maximizing scores without genuinely demonstrating their language proficiency, thereby undermining the validity of assessment interpretations.

Resource Intensiveness: Developing and implementing innovative assessment methods, such as computer-adaptive testing or performance-based assessments, can be resource-intensive in terms of time, expertise, and technology infrastructure. Educational institutions and assessment agencies may face challenges in investing sufficient resources to develop, administer, and maintain innovative assessment systems.

Technological and Access Barriers: The reliance on technology in innovative assessment methods may pose barriers to access for certain populations, such as individuals with limited digital literacy, internet connectivity issues, or disabilities. Ensuring equitable access to technology-enhanced assessments requires addressing these technological barriers and providing appropriate accommodations and support mechanisms.

Standardization and Consistency: Maintaining standardization and consistency across different administrations of English language assessments is essential for ensuring the reliability and comparability of scores. However, variations in test administration conditions, test item pools, and scoring practices may compromise the reliability of assessments and raise concerns about score interpretation and comparability over time.

Innovative Assessment Acceptance: The adoption and acceptance of innovative assessment methods may vary among stakeholders, including educators, learners, policymakers, and employers. Resistance to change, concerns about the validity and reliability of innovative assessments, and preferences for traditional assessment formats may hinder the widespread implementation of innovative assessment practices.

Ethical Considerations: English language assessment raises ethical considerations related to test fairness, confidentiality, and transparency. Ensuring ethical assessment practices requires addressing issues such as test security, test-taker privacy, and the ethical use of assessment data for decision-making purposes.

By acknowledging these limitations and drawbacks, researchers, practitioners, and policymakers can strive to address challenges and promote responsible and effective practices in English language assessment. Efforts to enhance assessment validity, reliability, accessibility, and fairness are essential for supporting equitable language learning opportunities and fostering meaningful language proficiency development.

CONCLUSION

In conclusion, this study has provided valuable insights into innovations, validity, and reliability in English language assessment. Through a mixed-methods approach combining quantitative and qualitative analyses, we have examined the performance of different assessment methods, explored participants' perceptions and experiences, and discussed implications for practice and research in the field. The findings of this study suggest that innovative assessment methods, such as computer-adaptive testing and performance-based assessments, offer promising opportunities for enhancing the validity and authenticity of English language assessment. These methods provide more nuanced and comprehensive evaluations of language proficiency, capturing not only linguistic knowledge but also communicative skills and socio-cultural competencies.

However, it is important to acknowledge the challenges and limitations associated with innovative assessment approaches. Issues such as technological barriers, cultural biases, and resource constraints may impact the accessibility, fairness, and reliability of assessments. Addressing these challenges requires collaborative efforts from stakeholders across the education sector to ensure equitable and inclusive assessment practices. Furthermore, this study highlights the need for ongoing research and development in English language assessment to continue advancing the field. Future research should focus on refining innovative assessment methods, validating their effectiveness across diverse contexts and populations, and addressing ethical and practical considerations in assessment design and implementation.

In conclusion, by embracing innovation, promoting validity and reliability, and prioritizing equity and inclusivity, English language assessment can better serve the needs of language learners, educators, and society as a whole. By continually striving for improvement and excellence in assessment practices, we can ensure that English language assessment remains a reliable and valid tool for measuring language proficiency and supporting language learning and development.

REFERENCES

[1]. Bachman, L. F., & Palmer, A. S. (2010). Language assessment in practice: Developing language assessments and justifying their use in the real world. Oxford University Press.

[2]. Alderson, J. C. (2005). Diagnosing foreign language proficiency: The interface between assessment and learning. A&C Black.

[3]. Douglas, D. (2010). Understanding language testing. Routledge.

[4]. Chapelle, C. A., & Douglas, D. (Eds.). (2006). Assessing language through computer technology. Cambridge University Press.

[5]. Weir, C. J. (2005). Language testing and validation: An evidence-based approach. Palgrave Macmillan.

[6]. Fulcher, G. (2010). Practical language testing. Hodder Education.

[7]. van der Linden, W. J., & Glas, C. A. (2010). Elements of adaptive testing. Springer Science & Business Media.

[8]. Pachler, N., Bachmair, B., & Cook, J. (Eds.). (2010). Mobile learning: Structures, agency, practices. Springer.

[9]. Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement (3rd ed., pp. 13–103). American Council on Education.

[10]. Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. CBS College Publishing.

[11]. Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Lawrence Erlbaum Associates.

[12]. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association.

[13]. Shohamy, E. (2013). The power of tests: A critical perspective on the uses of language tests. Pearson Education Limited.

[14]. Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), Educational Measurement (4th ed., pp. 17–64). American Council on Education.

[15]. Lantolf, J. P., & Thorne, S. L. (2006). Sociocultural theory and second language learning. Oxford University Press.

[16]. Weigle, S. C. (2002). Assessing writing. Cambridge University Press.

[17]. Levy, M., & Stockwell, G. (2006). CALL dimensions: Options and issues in computer-assisted language learning. Routledge.

[18]. Nunan, D. (2004). Task-based language teaching. Cambridge University Press.

[19]. Brown, H. D. (2004). Language assessment: Principles and classroom practices. Pearson Education.

[20]. Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. Bloomsbury Publishing.