A HYPOTHETICAL FRAMEWORK ON ASSESSMENT: TESTING OF ENGLISH LANGUAGE LEARNERS

Dr. T. Raja Sekhar

Reader, Department of English, Hindi Mahavidyalaya, Nallakunta, Hyderabad-500044

ABSTRACT

The present article provides an overview of language testing and assessment as a discipline rooted in applied linguistics, with a focus on English language learners and test-takers. It highlights the role of various professionals involved in language testing, including curriculum and syllabus designers, test designers, textbook authors, teachers, and researchers in English language teaching-learning-testing. The theoretical framework for language testing and evaluation is drawn from theories in applied linguistics, particularly Canale and Swain's model (1980) and educational measurement-based test theory as proposed by Bachman (1990). These theories form the basis for designing and conducting language tests to assess learners' language proficiency and performance. The passage also discusses the historical shift in terminology from "testing" to "assessment." Initially, the focus was on conducting tests to measure learners' language abilities, but over time, the concept expanded to include a broader view of assessment, encompassing various methods and tools used to evaluate language skills. This includes large-scale standardized tests like TOEFL and IELTS, as well as the classroom-based assessments created by teachers. The global trend in language assessment is towards integrating "assessment of learning" and "assessment for learning." "Assessment of learning" refers to evaluating learners' language proficiency after learning has taken place to determine the extent of their learning achievements. On the other hand, "assessment for learning" involves ongoing assessment practices that help teachers and students understand their progress and identify areas for improvement during the learning process.

The article's main focus is on the assessment and testing of English language skills, discussing the various types of language tests, the evolution of testing patterns, and emerging trends in language assessment. Overall, language testing and assessment are crucial aspects of language education, providing valuable insights into learners' language abilities and informing instructional decisions to enhance language learning outcomes. The integration of both assessment of learning and assessment for learning contributes to a comprehensive and holistic approach to language evaluation and teaching.

Keywords: *testing, assessment, learning of English language, kinds of testing, evolution of testing etc.,*

INTRODUCTION

Curriculums in educational institutions around the world frequently assert that textbooks and formal instruction promote neutral and secular knowledge. As one of the most important aspects of education, language acquisition has been utilised as a conduit for efficiently carrying out recursive and coercive activities such as "selection" and "judgement." Therefore, a nation's language policy frequently favours the knowledge domain(s) of the nation's power aristocracy.

In terms of languages, cultures, beliefs, political affiliation, economic and social standing, etc., it does not represent all segments of society in a democratic manner. This form of behaviour may not be manifestly intentional. One may question, however, why and how such methods occasionally find their way into language policies under the pretext of "secularisation" or Weber's "rationalisation." Developing and moulding the consciousness of the dominant class(es) without explicitly employing social control methods is a highly ingenious strategy.

Language policymakers utilise language testing, which has historically proven to be an effective social instrument, to promote a preferred and favoured segment of society as victors and accepted, and the remainder as losers, failures, and rejects. Even though voices have been raised against the gatekeeping aspect of tests and efforts have been made to make testing more student- and learning-friendly, we continue to be distinguished by our performance on high-stakes standardised tests such as the TOEFL and IELTS. In contrast, the records of the English Language Assessment provide a more accurate depiction of such undemocratic practises.

The curriculum emphasizes the importance of teachers reflecting on their past testing and assessment experiences to inform their current and future assessment practices. It discusses the benefits of both formative and summative assessments, which can provide valuable insights for teachers to modify their teaching methods effectively. Teachers need to have a clear understanding of the purpose of the assessment, whether it is for accountability, curriculum improvement, changes to the teaching-learning environment, or policy decisions. This understanding helps in selecting appropriate assessment techniques that align with the specific goals of the assessment.

Quantitative assessments, such as teacher-made tests, standardized tests, and rating scales, provide numerical scores to measure students' performance. There are two types of rating scales: norm-referenced scales, which compare a learner's performance to that of others who took the test, and criterion-referenced scales, which measure learners against specific performance criteria. On the other hand, qualitative assessments involve teacher observations, questioning, interviews, and student reflections. These qualitative methods play a crucial role in gaining a deeper understanding of students' learning processes, experiences, and areas of improvement.

Teachers should consider various factors while developing and designing assessments, such as test administration, structure, content, scoring, and the consequences of test results. Understanding test-taker experiences and providing constructive feedback can positively impact test outcomes and help teachers make meaningful connections between assessment constructs, processes, and uses. Overall, the significance of a balanced approach to assessment, is highlighted, incorporating both quantitative and qualitative methods to gain a comprehensive understanding of students' progress and learning experiences. By reflecting on past assessment experiences and continuously improving assessment techniques, teachers can enhance their instructional practices and better support their students' learning journeys.

STAGES IN EVOLUTION

There is a relationship between the history of English language instruction and English language assessment, indicating that shifts in language teaching methodology have also influenced changes in assessment practices. Spolsky (1975) attempted to categorize the evolution of language assessment into stages, though it is acknowledged that these divisions are not rigid and may overlap.

The emphasis on the most recent advancements in language evaluation is emphasized, particularly regarding the growing awareness of social dimensions in language assessment. This implies considering the broader societal and cultural contexts in which language assessment takes place, acknowledging the impact of cultural diversity and social factors on language proficiency and performance. Furthermore, the passage mentions the study of alternative assessment methods. This refers to exploring and implementing non-traditional approaches to language assessment that go beyond traditional standardized tests. Alternative assessments may include performance-based tasks, projects, portfolios, self-assessment, peer-assessment, and other authentic ways of evaluating language skills and competencies.

Thus, there appears to be an underscore of the importance of keeping up with the changing landscape of language assessment, incorporating a more socially aware perspective and exploring innovative assessment methods to better meet the needs of language learners in diverse contexts. By aligning language assessment practices with current advancements in language teaching and understanding the social dimensions of language evaluation, educators can create more relevant and effective assessment approaches to support learners' language development and overall communication abilities.

THE CONCEPTS OF ASSESSMENT, TEST, MEASUREMENT, AND EVALUATION

Assessment encompasses various methods, such as examinations, assignments, quizzes, and instructor observations, to gauge a student's linguistic proficiency or success. It involves employing diverse measurement instruments, some of which are administered throughout the course, while others are conducted at its conclusion.

A test serves as a tool to measure learning at a specific stage of the learning process, utilizing multiple-choice tests, quizzes, cloze tests, gap-fills, and similar formats.

Evaluation, on the other hand, focuses on the overall language course or program, assessing not just individual students' learning but also how well the course is achieving its goals and outcomes. This involves analyzing interviews, curriculum, materials, and other relevant sources.

Measurement, as defined by McNamara (2000), involves examining scores and their meaning, understanding what these scores reveal about students' learning, and ensuring alignment between assessment and learning objectives.

According to Bachman and Palmer (1996), practising teachers must grasp the attributes of test usefulness, which include:

Validity: Ensuring that a test measures what it is intended to measure accurately and appropriately reflects the test score's application.

Reliability: Establishing the dependability of test outcomes, indicating the consistency of the testing procedure concerning test administration and scoring.

Practicality: Assessing the extent to which test requirements can be met within the available resources, considering factors like time, personnel, and test administration constraints.

ENGLISH LANGUAGE ASSESSMENT AND ITS EVOLUTION

The history of the English language and English language education is closely interconnected, and as a result, English language assessment has evolved in parallel with changes in instructional strategies over time. Scholars such as Spolsky (1975), Brown (1996), and Bachman (2000) have studied the progression of language assessment across different phases and emphasized that these stages should not be rigidly compartmentalized with distinct characteristics. Recent advancements in language testing include a deeper understanding of the influence of social variables and the adoption of alternative evaluation techniques.

THE BEGINNING OF THE ENGLISH LANGUAGE TESTING

In the 15th century, Henry V implemented an English Language policy, replacing French with English as the official language of the Crown in all official communication. This necessitated the education of people in English, which led to Tutors making decisions on teaching and assessment procedures, limiting its accessibility to the general population. The 16th century witnessed efforts to define and conceptualize language, with significant attention given to developing English language teaching techniques.

With the publication of "Johann Christian Fick's 'Practical English Course' (1793) and John Miller's 'The Tutor' (1797)," English language learning became more accessible. This marked a shift towards theoretical considerations and research in English Language Teaching. Modern standardized tests for English language assessment were introduced in 1913. The University of Cambridge began sending examination papers to British colonies to assess candidates, leading to the establishment of the Certificate of Proficiency in English Examination in 1913, granting foreign qualifiers the status of proficient English language users. Since then, standardized tests have played a dominant role in establishing English language hegemony.

The British were pioneers of linguistic imperialism through English Language Education, and the Americans followed suit. Spolsky (1993) explains how the U.S. immigration department used English Language tests as control tools to manage the inflow of immigrant students to the

144

country in the early 1920s. These 'prognosis tests' were segregating and elitist, perpetuating the influence of the U.S. and the U.K. through language tests. Despite the passage of time, the world still experiences the organized filtering impact of language tests established by the U.S. and the U.K., reflecting their historical linguistic imperialism.

GRAMMAR – TRANSLATION METHOD OF TESTING

Before the 1960s, the English language teaching-learning-testing process was significantly influenced by the Grammar-Translation methodology. This approach emphasized the memorization of grammatical rules and their application in translating sentences between the learner's native language and English. As a result, language learning became focused on rote memorization of rules and terminology.

In this era, language testing primarily assessed learners' ability to read, write, and translate accurately and without errors. Oral proficiency and the use of language for communicative purposes were not emphasized in testing by teachers and test designers. Instead, the tests were designed to be straightforward and time-saving, with a focus on evaluating grammatical accuracy and vocabulary retention in the target language. The "Charter's Diagnostic Language Test" and the "Pressey English Test" are examples of exams from this period, which assessed grammar, punctuation, capitalization, and sentence structure errors. However, these tests lacked objectivity and statistical analysis, resulting in their inefficiency. Spolsky (1978) refers to this period as "pre-scientific," highlighting the lack of rigorous scientific methodology in language testing during this time.

Overall, the pre-Lado era was characterized by a dominant Grammar-Translation approach in language teaching, which was mirrored in language testing focusing on rule-based accuracy rather than communicative proficiency. The lack of scientific rigor in designing and evaluating tests limited their effectiveness in truly assessing learners' language abilities.

THE STRUCTURALIST APPROACH TO TESTING

According to Spolsky (1978), the Grammar-Translation method of language assessment was replaced by a "psychometric-structuralist" trend, influenced by behaviorists like Skinner and structural linguists such as Fries and Bloomfield. During this period, students were encouraged to compete, and test scores became the primary focus, overshadowing individual abilities. Discrete-point testing was used to create a fixed standard for evaluating students' language proficiency, prioritizing passing and failing over actual learning. However, this approach led to learners conforming to predetermined behaviors, resulting in a loss of their natural ability to learn. The learners' needs and the significance of teaching-learning contexts were sacrificed in favor of scientific objectivity, validity, reliability, and precision.

The integrative approach to language testing emerged as an alternative to discrete-point testing. Influenced by cognitive psychology and Noam Chomsky's language and learning theories, this approach emphasized functional competence over linguistic accuracy. Language testing shifted its focus to assessing an individual's capacity to communicate effectively and responsibly in

145

social situations, taking into account linguistic features and skills within sociolinguistic perspectives influenced by Hymes (1972). However, even with the inclusion of contextual tests, language testing faced challenges in fully addressing social and cultural dimensions. The contexts in which language items were tested often distanced linguistic, cultural, and social minorities, as it was challenging to accommodate a wide variety of contexts in a single test.

In summary, the shift from discrete-point testing to an integrative approach reflected a move towards assessing functional competence and communication skills in social contexts. However, challenges remained in fully capturing the diverse linguistic and cultural backgrounds of learners in language testing contexts.

COMMUNICATIVE APPROACH TO TESTING

During the 1980s and early 1990s, language testing methods developed by scholars like Morrow (1979), Canale and Swain (1980), and Bachman (1990) gained popularity and significance. These methods focused on assessing learners' language competency. Building on Hymes' work, Bachman and Palmer (1996) emphasized the importance of clearly defining the test concept and context, using authentic materials in the tests, and basing the context of test items on real-life situations.

The researchers highlighted that language tests should go beyond linguistic abilities and also measure cognitive talents such as sociolinguistic proficiency, discourse fluency, and strategic thinking. In adopting this communicative framework, language testing aimed to be more inclusive and reflective of learners' actual language abilities in real-world scenarios.

However, despite the advantages and inclusiveness of the communicative approach to testing, it has not become widely adopted in mass testing practices across the world. Traditional testing methods, including standardized and discrete-point testing, still dominate in many educational systems. Thus, the 1980s and early 1990s saw the emergence of language testing methods that focused on learners' language competency and incorporated the communicative framework. While these methods aimed to be more inclusive and relevant to learners' real-life language abilities, they have yet to become fully integrated into mass testing practices globally. Traditional testing approaches continue to persist in many educational contexts.

LANGUAGE ASSESSMENT AND ALTERNATIVE IN/TO ASSESSMENT

In the 1990s, language testing experts, including Spolsky, McNamara, and Shohamy, began to recognize that language assessments were not merely linguistic exercises but also instruments with broader social, political, cultural, and ideological implications. They emphasized the importance of fairness and ethics in language assessment, and concepts such as "democratic assessment," "alternative assessment," "fairness," and "bias" were explored.

The focus of language testing shifted towards alternative 'in' and 'to' testing approaches, as discussed by Brown and Hudson (1988). These approaches included various assessment

146

methods such as checklists, journals, logs, videotapes, audiotapes, self-evaluation, teacher observations, portfolios, conferences, diaries, and peer and self-assessment. These options went beyond traditional testing and aimed to foster language learning while accommodating different learning styles and linguistic competencies of individual learners.

Language testing is now recognized as a deeply embedded political and ideological issue in society and culture. Scholars have critiqued the power dynamics and disciplinary functions of examinations. Henry Latham (1877) was among the first to question the influence of examinations on education. Michel Foucault (1977) also highlighted the examination's role in constituting individuals as subjects of authority and knowledge through hierarchical surveillance and normalizing judgment. This critique underscores the broader impact of language testing beyond its immediate linguistic assessment role.

In conclusion, the 1990s marked a shift in language testing towards acknowledging its social and ideological implications. Alternative assessment approaches were explored, and scholars began critically examining the power dynamics of examinations in shaping individuals and society. Foucault's work, possibly influenced by Paulo Freire's "Pedagogy of the Oppressed," played a role in shaping the discipline of Applied Linguistics in the late 1970s. However, it took time for critical perspectives to be fully incorporated into language education, particularly language testing. In the 1980s, Norman Fairclough's *Language and Power* brought attention to ideas of anxiety, bias, hegemony, democracy, marginalization, dominance, and ideology, signaling the consideration of critical perspectives in language testing discourse.

Throughout the 1990s, language testing experts such as Spolsky, Tim McNamara, and Elana Shohamy viewed language testing as a powerful educational tool used for social, political, cultural, and ideological control. The field of Critical Applied Linguistics provided substantial backing to the proponents of justice and ethics in language testing. This led to the emergence of concepts such as critical language testing, democratic assessment, alternative assessments, fairness, and prejudice, among others.

The critical examination of language testing has spurred the search for alternative 'in' and 'to' assessment methods. Brown and Hudson (1998) proposed various assessment alternatives, such as checklists, journals, logs, videotapes, audiotapes, self-evaluation, teacher observations, portfolios, conferences, diaries, self-assessments, and peer assessments. These alternatives offer solutions to move beyond traditional testing approaches and view assessment as a tool for promoting learning in formal educational settings. They also allow for a comprehensive framework that accommodates diverse learning styles and preferences while recognizing individual learners' identity and talents.

In language testing and assessment, the focus extends beyond merely measuring language abilities and components. It considers test score interpretation, correctness, appropriateness, and the sociocultural environment in which assessment occurs. This approach emphasizes integrated skills tests, grading, and providing feedback to test-takers, particularly in language classroom exams, to generate relevant and meaningful assessments.

147

Overall, the incorporation of critical perspectives in language testing and assessment has brought attention to the broader social, political, and ideological implications of these practices, leading to the exploration of alternative assessment methods and a focus on promoting meaningful learning experiences for language learners.

EVALUATION AND ASSESSMENT

The terms "evaluation" and "assessment" are often used interchangeably, but they have distinct meanings and purposes in the context of language education.

Evaluation refers to the measurement and observation of the quality of achievement in relation to predetermined objectives and criteria. It investigates and determines the validity and usefulness of the language program as a whole. Evaluation is concerned with judging the overall performance and outcomes to inform decision-making about the language program's effectiveness. Assessment, on the other hand, is the ongoing measurement, monitoring, and improvement of learning and achievement levels. It assesses the extent to which learning objectives and outcomes are being met. Assessment occurs on a frequent basis and involves active engagement from both the teacher and the learner. Its primary goal is to gauge the learners' overall proficiency in the language and identify areas that require improvement.

Evaluation is thus focused on overall results and the effectiveness of the language program, while assessment is more concerned with monitoring progress and supporting continuous improvement in language learning. Both evaluation and assessment are essential components of language education, providing valuable insights into learners' performance and helping educators make informed decisions to enhance language learning outcomes.

SUMMATIVE AND FORMATIVE ASSESSMENT

Assessments can be categorized into two types: Summative and Formative.

Summative assessment involves gathering supporting evidence from examinations or tests, typically at the end of a course or program. This form of assessment is often associated with standardized tests, like the TOEFL, where measures are validated, and criteria are used to evaluate and score the results. The main purpose of summative assessment is to make judgments about students' overall achievement and performance. It helps determine the quality and effectiveness of a course or program.

On the other hand, formative assessment is used to monitor and evaluate students' development and progress during the learning process. This assessment occurs throughout the course or program and involves various methods such as observation, class activities, or providing feedback to students. The main goal of formative assessment is to modify and improve learning while it is still in progress. It provides timely and constructive feedback to teachers, learners, and administrators, helping to identify areas for improvement and guiding instructional decisions. In summary, summative assessment is used at the end to make overall judgments about students' achievement, while formative assessment is ongoing and aims to support and enhance learning throughout the course or program. Both types of assessment are essential

148

components of effective education, providing valuable information for evaluating progress and promoting continuous improvement in learning outcomes.

CONCLUSION

Over the past 50 years, significant progress has been made in language testing and assessment, particularly in three crucial areas: test validity, washback research, and classroom assessment techniques. Test validity, being the foundation of testing, focuses on the accuracy and relevance of evaluating a student's linguistic proficiency. There is a growing emphasis on gathering legitimate evidence from test-takers, including teachers and students, and employing diverse research approaches, such as classroom instructional data, to enhance test validity (Messick, 1989).

Washback, the intricate interaction between instruction, learning, and assessment, has been recognized as a crucial aspect of language testing's impact on language instruction and acquisition (Alderson & Wall, 1993). Researchers like Cheng, Watanabe, and Curtis (2004) have delved into washback research to understand its implications. The findings of such research have demonstrated the influence of diverse teaching-learning environments on fundamental language teaching and learning domains.

Recent advancements in testing include research on the assessment practices of English teachers in the classroom. Limited research indicates that large-scale language testing significantly affects classroom assessment procedures, and teacher preparation programs often lack assessment training (Cheng, 2005). The connection between English language teachers' assessment practices and their teaching and assessment beliefs has been highlighted by other researchers, such as Breen et al (Cumming 2001). Assessment for learning has emerged as a crucial pedagogical approach to enhance student achievement. Effective classroom assessment procedures, which are valid and reliable, are utilized by instructors to support student learning (Black & Wiliam, 1998a, 1998b). While large-scale language examinations are increasingly designed to facilitate learning, it is the responsibility of classroom teachers to implement effective assessment procedures to ensure adequate student learning.

In conclusion, the last five decades have witnessed substantial advancements in language testing and assessment, with a focus on test validity, the impact of assessments on language instruction and learning (washback), and the importance of effective classroom assessment techniques to promote student achievement and learning. These developments have contributed to the overall improvement of language education and have emphasized the significant role of teachers in fostering meaningful and impactful language learning experiences.

REFERENCES

1. Apple, M. W. (1979). *Ideology and Curriculum*. London: Routledge.
2. Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

149

3. Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing*, 17, 1-42.

4. Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

5. Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.

6. Brown, J. D., & Hudson, T. (1998). *Alternatives in language assessment*. TESOL Quarterly, 32(4), 653-675.

7. Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.

8. Fairclough, N. (1989). Language and Power. London: Longman.

9. Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison*. London: Penguin Books.

10. Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Herder & Herder.

11. Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal* 5, 8–11

12. Hymes, D. (1972). *On Communicative Competence. In J.B. Pride & J.Holmes (Eds.). Sociolinguistics*. Harmondsworth, England:Penguin Books.

13. Kunnan, A. J. (Ed.). (2000). *Fairness and validation in language assessment: Selected papers from the 19th language testing research colloquium*, Orlando. New York: Cambridge University Press.

14. Latham, H. (1877). *On the Action of Examinations Considered as a Means of Selection* Cambridge: Deighton, Bell.

15. Morrow, K. (1979). Communicative language testing: revolution or evolution? In Brumfit, C.J. and Johnson, K., (Eds.), *The communicative approach to language teaching*. Oxford: Oxford University Press, 143–57.

16. Shepard, L.A. (1981). *Identifying bias in test items. In B.F. Green (Ed.), New directions in testing and measurement: Issues in testing-Coaching, disclosure, and test bias*. San Francisco: Jossey-Bass, 79-104

17. Shohamy, E. (2001). *Democratic assessment as an alternative*. Language Testing 18, 4, 373 – 392.

18. Spolsky, B. (1978). Introduction: Linguistics and language testers. In B. Spolsky (Ed.), Advances in language testing series: 2. Arlington, VA: Center for Applied Linguistics

19. Spolsky, B. (1978). (Ed.). *Approached to Language Testing. Papers in Applied Linguistics*. ERIC ED 16548

20. Spolsky, B. (1993). *Testing of English of Foreign Students in 1930*. ERIC 2002.

21. Spolsky, B. (1975). Language testing – the problem of validation. In L. Palmer & B. Spolsky (Eds.). *Papers on Language Testing 1967-1974*. Washington, D.FC.: TESOL. 147-53.