

Tikrit University
College of Education for Humanities
English Department



PhD (Second Course) 2025-2026

**An Advanced Course in EFL Communicative Testing,
Measurement and Evaluation**

Assessment Concepts and Issues (Part Two)

(2)

Prof. Dunia Tahir (PhD)

Brown, H. D., & Abeywickrama, P. (2018). *Language Assessment Principles and Classroom Practices* (3rd ed.). Pearson Education.

2026. A.D.

1447.A.H.

Behavioral Influences on Language Testing

Through the middle of the twentieth century, language teaching and testing were both strongly influenced by behavioral psychology and structural linguistics. Both traditions emphasized sentence-level grammatical paradigms, definitions of vocabulary items, and translation from first language (L1) to second language (L2) and placed only minor focus on real-world authentic communication. Tests often consisted of grammar and vocabulary items in multiple-choice format along with a variety of translation exercises ranging from words to sentences to short paragraphs.

Such discrete-point tests still prevail today, assessments were designed on the assumption that language can be broken down into its component parts and that those parts can be tested successfully. These components are the skills of listening, speaking, reading, and writing and the various units of language.

Q1: What influenced language teaching and testing in the mid-twentieth century?

Q2: What did traditional language tests mainly focus on?

Q3: What is meant by "discrete-point testing"?

Q4: Where are discrete-point tests still commonly used?

Q5: What assumption underlies discrete-point testing?

Q6: What language skills and components were tested in this approach?

Q7: What approach did Spolsky call this testing tradition?

Q8: Name two popular standardized tests that emerged from this approach?

Integrative Approaches

Pedagogy was rapidly moving in more communicative directions, and testing specialists were forced into a debate that would soon respond to the changes. The discrete-point approach presupposed a decontextualization that was proving to be inauthentic. So, as the

profession emerged into an era emphasizing communication, authenticity, and context, new approaches were sought. John Oller (1979) argued that language competence was a unified set of interacting abilities that could not be tested separately.

His claim was that communicative competence is so global and requires such integration that it cannot be captured in additive tests of grammar, reading, vocabulary, and other discrete points of language. Others soon followed with their support for what became known as integrative testing.

Q1: Why did the discrete-point approach begin to be questioned?

Q2: What shift occurred in language pedagogy during this period?

Q3: What was John Oller's (1979) main argument?

Q4: Why did Oller reject additive testing of language skills?

Q5: What is integrative testing?

Q6: Who supported the integrative testing approach?

Q7: What is an integrative test?

Q8: What two types of tests were considered examples of integrative tests?

Q9: What is a cloze test?

Q10: Why did Oller (1979) believe cloze tests measured overall language proficiency?

Q11: What abilities does cloze test tap into?

Q12: What is dictation in language testing?

Q13: Why was dictation considered an integrative test?

Q14: What cognitive skills are required for success in dictation tests?

Q15: Why were dictation tests considered practical for large-scale testing?

Q16: What is the unitary trait hypothesis?

Q17: What language components were believed to be inseparable under this hypothesis?

Q18: Why was the unitary hypothesis abandoned?

Communicative Language Testing

By the mid-1980s, especially in the wake of Canale and Swain's (1980) seminal work on communicative competence, the language-testing field had begun to focus on designing communicative test tasks. Bachman and Palmer (1996, 2010) included among "fundamental" principles of language testing the need for a correspondence between language test performance and language use.

And so a quest for authenticity was launched, as test designers centered on communicative performance. Following Canale and Swain's (1980) model, Bachman (1990) proposed a model of language competence consisting of organizational and pragmatic competence, respectively subdivided into grammatical and textual components and into illocutionary and sociolinguistic components.

Bachman and Palmer (1996, 2010) also emphasized the importance of strategic competence (the ability to use communicative strategies to compensate for breakdowns and to enhance the rhetorical effect of utterances) in the process of communication. All elements of the model, especially pragmatic and strategic abilities, needed to be included in the constructs of language testing and in the actual performance required of test-takers.

Test designers began to identify the kinds of real-world tasks that language learners were called on to perform. The assessment field also became more concerned with the authenticity of tasks and the genuineness of texts.

Q1: What major shift occurred in language testing by the mid-1980s?

Q2: What influenced the development of communicative language testing?

Q3: According to Bachman and Palmer (1996), what is a fundamental principle of language testing?

Q4: What problem did language assessment expect to face with earlier tests?

Q5: What is language competence?

Q6: What is organizational competence and grammatical competence?

Q7:What is sociolinguistic competence ?

Q8:What is strategic competence?

Traditional and “Alternative” Assessment

Q1:What is meant by traditional assessment?

Q2:What is alternative assessment?

Q3:Why did alternative assessment develop?

Q4:Can a clear line be drawn between traditional and alternative assessment?

Q5:What caution is given about comparing traditional and alternative assessment?

Q6:What do Brown and Hudson (1998) say about assessment traditions?

Q7:What are some challenges of alternative assessment?

Q8:What is meant by offering feedback?

Q9:What is the payoff of alternative assessment?

Q10: Why is alternative assessment considered more motivating?

Q11:How does performance-based assessment help learning?

Q12:What is the goal of educational reform?

Q13:How does alternative assessment differ from traditional testing?

What are the differences between Traditional and alternative assessments?

No.	Traditional	Alternative assessment
1	One-shot, standardized exams	Continuous, long-term assessment
2	Timed, multiple-choice format	Untimed, free-response format
3	Decontextualized test items	Contextualized communicative tasks
4	Scores sufficient for feedback	Individualized feedback and washback

5	Norm-referenced scores	Criterion-referenced scores
6	Focus on discrete answers	Open-ended, creative answers
7	Summative	Formative
8	Oriented to product	Oriented to process
9	Noninteractive performance	Interactive performance
10	Fosters extrinsic motivation	Fosters intrinsic motivation

Performance-Based Assessment

During the past two decades, an increasing number of educators and advocates for educational reform have argued to de-emphasize large-scale standardized tests in favor of contextualized, communicative assessments that better facilitate learning in our schools.

The argument was that standardized tests do not elicit actual performance on the part of test-takers. Performance-based assessment would require the performance of the actions which would be systematically evaluated through direct observation by a teacher and possibly by self and peers .

Instead of offering paper-and-pencil selective-response tests of a plethora of separate items, performance-based assessment of language typically involves oral production, written production, open-ended responses, integrated performance (across skill areas), group performance, and other interactive tasks. A characteristic of many (but not all) performance-based language assessments is the presence of interactive tasks-and hence another term, task-based assessment,

Because performance assessments are key tools for linking classroom practices to real-world activities, such assessments are considered ideal for formative assessment practices in classroom instruction and are part of the wider concept of classroom-based assessment in the field of education

Q1:What is performance-based assessment?

Q2:How does performance –based assessment differ from paper-and-pencil tests?

- Q3: Why is performance-based assessment considered more accurate?
- Q4: What is meant by content validity?
- Q5: Why is performance-based assessment time-consuming and expensive?
- Q6: In what teaching context is performance-based assessment commonly used?
- Q7: What is task-based assessment?
- Q8: What kind of tasks are used in performance-based assessment ?
- Q9: Why do paper-and-pencil tests fail to measure communicative performance?
- Q10: What is a prime example of an interactive language assessment?
- Q11: What skills are required in an interview assessment?
- Q12: How can performance-based assessment achieve authenticity?
- Q13: What role do interactive tasks play in performance-based assessment ?
- Q14: What is the main goal of performance-based assessment?

CURRENT “HOT TOPICS” IN LANGUAGE ASSESSMENT

Three important topics include (1) dynamic assessment, (2) assessing pragmatic competence, and (3) the increasing use of technology in assessments of various kinds.

Dynamic Assessment

The focus and emphasis on assessing for learning, which aligns closely with formative assessment practices, draws parallels to dynamic assessment (DA), a prolearning form of assessment conceptually based on Vygotskian approaches to education. The zone of proximal development (ZPD) considers a learner’s potential abilities beyond the actual performance in a task, what the learner can do when others give assistance.

Poehner and Lantolf (2003) argue that an assessment of a learner is “not complete until we observe how the person involved behaves in response to assistance. In other words, to fully understand a person’s potential to develop (i.e., her/his future), it is necessary to discover her/his ZPD”.

DA, as its name suggests, contrasts sharply with traditional assessment, which is static or stable over time. Instead, in DA, learner abilities are considered malleable, not fixed.

Classroom practices and assessments may include:

- 1- providing clear tasks and activities
- 2- posing questions that prompt students to demonstrate understanding and knowledge
- 3- interventions with feedback and student reflections on their learning

Q1:What is Dynamic Assessment ?

Q2:Which theory is DA based on?

Q3:What is the ZPD?

Q4:How does DA differ from traditional assessment?

Q5:What classroom practices are associated with DA?

Q6:What is the main goal of DA?

Assessing Pragmatics

Pragmatics is the use of language in interactions and is a field that studies language choices made by users, the constraints they encounter, and the impact of the language used on other interlocutors. Although research on L2 pragmatics has developed over the years, assessment of pragmatic competence is an underexplored area in language assessment research. Tests of pragmatics have primarily been informed by research in interlanguage and cross-cultural pragmatics. The most widely used types of research instruments include discourse completion tasks.

Q1:What is pragmatics?

Q2:Why is assessing pragmatic competence important?

Q3:What types of tasks are commonly used to assess pragmatics?

Q4:What speech acts are often studied in pragmatics research?

Q5:What additional aspects of pragmatics should be assessed?

Use of Technology in Testing

In recent years there are more technological innovation and applications of technology to language learning and teaching. a number of language courses use some form of computer-assisted language learning (CALL) or mobile assisted language learning (MALL) to achieve their goals.

Technology-assisted testing offers unique **advantages**:

- * a variety of easily administered classroom-based tests
- * self-directed testing on various aspects of a language (vocabulary, grammar, discourse, one or all of the four skills)
- * practice for upcoming high-stakes standardized tests
- * individualization with customized, targeted test items
- * large-scale standardized tests that can be administered easily to thousands of test-takers at many different stations, and then scored electronically to report results quickly
- * improved technology for automated essay evaluation and speech recognition

Of course, some **disadvantages** are present in our current predilection for all things technological:

- * Lack of security and the possibility of cheating are inherent in unsupervised computerized tests.
- * Occasional “homegrown” quizzes that appear on unofficial Web sites may be mistaken for validated assessments.
- * The multiple-choice format preferred for most computer-based tests contains the usual potential for flawed item design.
- * Open-ended responses are less likely to appear because of (a) the expense and potential unreliability of human scoring or (b) the complexity of recognition software for automated scoring.

- * The human interactive element (especially in oral production) is absent.
- * Validation issues stem from test-takers approaching tasks as test tasks rather than as real-world language use

Q1: Why has technology become important in language testing?

Q2: What is CALL? Why is it widely used?

Q3: What is MALL?

Q4: How has technology affected language assessment?

Q5: What is computer-Adaptive testing? What is the goal of CAT?

Automated Scoring and Innovation

Q6: What major tests use automated scoring for writing and speaking?

Q7: How does corpus linguistics improve language testing?

Q8: Why might video listening tests differ from audio-only tests?

Q9: What are the main advantages of technology-assisted testing?

Q10: What is a major disadvantage of technology-assisted testing?

Q11: What types of tests can be easily administered using technology?

Q12: Why is technology useful for large-scale standardized tests?

Q13: What role does technology play in scoring writing and speaking?

Q14: What are open-ended responses less common in technology-based tests?

Q15: What important element is missing in technology-assisted testing?

Q16: Why are tests important in curriculum?