Journal of Applied Linguistics and Applied Literature: Dynamics and Advances

Book Review: Anthony Green, *Exploring Language Assessment and Testing: Language in action* (2nd edition), London: Routledge. 2021, 290 pages. ISBN: 978-1-138-38878-9 (Paperback)

Rajab Esfandiari (corresponding author)

Associate professor of English Language Teaching, Department of English Language, Faculty of Humanities, Imam Khomeini International University, Qazvin, Iran, ORCID: 0000-0002-2305-762X
Email: esfandiari@hum.ikiu.ac.ir

When language testing began as a formal activity in the early 1960s, it was primarily concerned with large-scale, standardised tests such as test of English as a foreign language (TOEFL), heavily influenced by the then two dominant schools of thought: structural linguistics and behavioristic psychology. However, since then, much has changed, and language testing has developed fully into a legitimate subfield onto itself in applied linguistics, with its own academic national and international journals, conferences, handbooks, and forums and websites. As Alderson and Banerjee (2001) aptly noted, "the field has become so large and so active that it is virtually impossible to do justice to it … [,] changing so rapidly that any prediction of trends is likely to be outdated before it is printed" (p. 213). One major paradigm shift which language testing and assessment has witnessed in recent years is classroom-based assessment, which primarily focuses on the assessment of the learning processes and objectives, the practical dimensions of L2 assessment, and integration of testing, teaching, and learning than probing into the psychometric properties of language tests. Such a change is best reflected in language assessment textbooks commercially produced and incorporated into educational programmes.

Exploring language assessment and testing: Language in action reflects a classroom-based view of assessment, following a "practice to theory" approach, with a "back-to-front" structure. The book is an accessible introduction to L2 assessment structured into three main parts, each with interdependent chapters. In part I, Green describes the practical aspects of language assessment, explaining the relationship between testing and teaching, outlining the purposes of language assessment, pinpointing the assessment cycle, and finally ending this part with four

Major qualities of classroom-based assessment. Part II presents the techniques and procedures of how to assess receptive and productive skills, with primary focus on applying those principles discussed in part I. Theoretical in nature, Pat III delves into the historical accounts and theoretical perspectives which have influenced the practice of language testing over the years. A full discussion of the contents of the textbook is presented in the following paragraphs.

The first chapter in part I lays the foundations of assessment and situates testing in its broader context. At the opening of the chapter, Green uses the opportunity to argue that the ultimate goal of assessment is to provide evidence for decision-making purposes. It is in line with this argumentation that he defines language assessment simply as "obtaining evidence to inform inferences about a person's language-related knowledge, skills or abilities" (p. 5). Green draws on this simple definition to frame the rest of the textbook and comments that language assessment is inseparable from teaching and learning. The rest of the chapter deals with a clear distinction between testing and assessment and integrates assessment, teaching, and learning.

In chapter 2, the author focuses on the purposes assessments are used for. Such purposes include selection, placement, achievement, diagnosis, proficiency, aptitude, and prognosis. A distinction is made between educational and proficiency purposes, with the former based on the learning processes and course outcomes and the latter concerned with the ability to accomplish certain tasks, irrespective of how the ability is reached. Throughout the chapter, clear examples are given to help readers better undemand the assessment functions.

Chapter 3 examines the stages of test development—the most practical dimension of language assessment in general and in this part in particular. This is the second lengthiest chapter in the book, possibly because of the significance attached to test construction. The author depicts an assessment cycle in which eight interconnected stages should be followed to develop a careful test—whether it is classroom-based or standardised. The assessment cycle includes roles, responsibilities, and documentation, and consists of design, organisation, production, administration, assessment, scoring, score using, and validation. Through schematic representation and use of several analogies and examples, Green painstakingly clarifies the stages, establishes the relationships between the stages, distinguishes between teacher-made and standardised tests when the stages need to be applied, and uses his own experience to present practical advice to trainee teachers how best the stages may be implemented in real-life situations. It should be noted that, as Green himself emphsises, such stages are not monolithic, assessments should viewed vied as "continuous and repeating cycles of activities" (p. 26), and the completion of the assessment cycle does not imply the process ends.

Chapter 4 describes the salient features of an effective assessment—the qualities that result in useful assessments. Four major characteristics are listed and explained, including practicality, reliability, validity, and beneficial consequences. These qualities, as Green remarks, act as quality control indicators which help users to choose assessments or educators to evaluate an assessment system. Green has

intentionally decided to focus on the traditional features of a good language test, because he notes that other characteristics such as authenticity are accounted for in his four-quality control. Such a justification seems reasonable, given the very fact that close attention is paid to classroom-based tests in this textbook. He is also at pains to successfully show these characteristics map onto the stages of test design fully explained in chapter 3. In closing the chapter, the author advises readers to bear in mind that "no assessment can fully satisfy all four essential qualities" (p. 97) and suggests that the utility of effective systems depend on the ability of their users to use them correctly and appropriately.

Chapter 5 analyses the techniques and procedures for assessing receptive knowledge, skills, and abilities—listening and reading. In this chapter, and the following one, a socio-cognitive approach is adopted to explain language skills because it is the social expectations and conventions that shape the way language is used, and it is the psychological insights into mental processes that help us to convey a message, organise an idea, and make sense of the sounds we hear and the language we produce (Weir, 2005). Vocabulary and grammar are treated in chapter 5, and pronunciation in chapter 6, because Green believes that these components are not separable from written and spoken language. This chapter covers conceptual underpinnings of listening and reading, describes the processes, skills, and strategies readers and listeners follow to make sense of what they hear and read, discusses types of reading and listening, and deals with different scoring methods.

In chapter 6, Green turns to productive and interactive skills—speaking and writing, with pronunciation included. Some space is first devoted to distinguishing between production and interaction, because Green believes production involves "planned and rehearsed monologues" (p. 135) such as classroom lectures, while interaction "implies the kind of spontaneous exchange found in conversation" (p. 135). Green is quick to admit that production and interaction are two extremes of a continuum because almost all spontaneous conversations include some sort of planning, and all planned speeches and texts contribute to interaction. In the rest of the chapter, features of spoken and written language are presented, various task types are listed for assessing speaking and writing, different assessment instruments such as rating scales are introduced for recording performance, and well-known scoring procedures, including holistic scales, are used to score performance on speaking and writing. The chapter ends with rater training and the use of computers for scoring productive and interactive performance. In both chapter 5 and chapter 6, a number of examples from internationally recognised tests and frameworks such as CEFR (common European framework of reference) are cited to help readers digest the points.

Chapter 7 in part III is theoretical in nature. This chapter extends our understanding of the points raised and explained in the first six chapters, on the one hand, and serves to explore more theoretical concepts in language assessment, on the other hand. The chapter begins with a succinct introduction of the sixty-year-old history of language testing and is followed by the growth of the professionalisation of the field, with the primary focus on the early years. Next, six phases (or tendencies, to use Green's terminology) in language assessment over the years are

outlined, including pre-scientific/traditional, psychometric/structuralist, psycholinguistic-sociolinguistic, communicative, formative testing, and assessment for learning, with the first four stages reflecting the earlier developments in the field and the last two ones the more recent trends. Green also lists the most influential testing figures such as Robert Lado and the testing textbooks which contributed to the development of the field. The rest of the chapter, one way or the other, focuses on the full treatment of the stages in the history of language testing, and the author pinpoints how the early skills and components theory in language testing in the early years has been replaced by communicative language testing in recent years. Green is very careful to describe the test methods such as discrete-point tests, the influential testing figures, the major commercial tests, the groundbreaking testing textbooks, and testing models and frameworks during each stage to present a more objective view of paradigm shifts over the years, although he personally comments on the developments and shares with readers his own personal experiences.

In addition to the points discussed in the foregoing paragraph, chapter 7 also touches on some concepts which play a more prominent role in the current status quo. Concepts such as fairness, bias, ethics, critical language testing, quality standards, self-reflection and self-regulation reflect the view that the field is growing much faster than ever before and becoming more autonomous and professionalised—a subfield in applied linguistics once looked down on is now appreciated for the major breakthroughs and its interaction with neighbouring disciplines such as measurement theory, sociology, language education, and language teaching. The chapter closes with some emerging trends in language assessment (e.g., computer adaptive tests, internet-based testing, automated scoring, continuous assessment, more sophisticated statistical tests and models, and the role of corpora in language tests) which reflect the growing prominence of technology types, the complexification of educational goals, and the sophistication of learner variables. Just to give readers an example of the fast-spreading nature of the field, in passing, I need to cite Aryadoust and Raquel (2019, 2020), who edited a two-volume handbook specifically devoted to data analysis in language assessment.

Chapter 8, as the concluding chapter of the book, provides answers to the tasks the author has added to each chapter to set readers thinking about the points raised in each chapter and to help them relate these points to the boarder context in real-life contexts. Green's detailed, lucid explanations in this chapter provide another opportunity for readers to learn new points about language assessment.

Exploring language assessment and testing: Language in action is an updated version of its 2014 predecessor. Although this second edition remains intact in its structure and organisation, it is a welcome and timely addition to the field of language assessment, because thematically it reflects the most recent developments in the field. In addition to much new content information, Green has also updated the glossary in which the terminology used in the book is briefly explained, more recent references are added to Further Reading section, and more new references are included in the reference list to reflect the new changes in the field. Green has adopted a pragmatic approach in writing this textbook and presenting the information in that the real-world problems and experiences are introduced first,

followed by practical examples to steer readers towards the discussion of language assessment points, before the readers relate the experiences, problems, and points to theoretical perspectives. Although the textbook is primarily written to be "an indispensable introduction for postgraduate and advanced undergraduate students", in language education and applied linguistics, I highly recommend it to all the members of community of practice in language assessment (and more specifically language teachers) because of the advantages it offers in classroom-based assessment, highly accessible presentation of the assessment points, and the author's lucid, user-friendly style.

References

Alderson, J. C., & Banerjee, J. (2001). Language testing and assessment (part 1). *Language Teaching, 34*(4), 213-236. https://doi.org/10.1017/S0261444800014464

Aryadoust, V., & Raquel, M. (Eds.). (2019). *Quantitative data Analysis for language assessment Volume I: Fundamental techniques*. Routledge.

Aryadoust, V., & Raquel, M. (Eds.). (2020). *Quantitative data analysis for language assessment Volume II: Advanced methods*. Routledge.

Green, A. (2021). *Exploring language assessment and testing: Language in action* (2nd edition). Routledge.

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

Author's Biography

Rajab Esfandiari is an associate professor of English Language teaching at Imam Khomeini International University in Qazvin, Iran. His areas of interest and specialty include teaching and assessing L2 writing, multifaceted Rasch measurement, L2 classroom assessment, and EAP teaching and testing. He can be reached via his e-mail address: *esfandiari@hum.ikiu.ac.ir*