

REVIEW

**Xun Yan, Slobodanka Dimova and
April Ginther: *Local Language
Testing: Practice across Contexts***

(Cham: Springer Nature Switzerland AG, 2023. 252 pages)

Jianshun Lu

Hebei North University
China
luhaiyao005@163.com

Zhiqiang Wang

Hebei North University
China
wangzq0016@163.com

Even though Local Language Testing (LLT) is the most accepted pattern for language evaluation practice, it has not been appropriately distinguished until very lately. In spite of the omnipresence of LLT processes, local tests barely stand in the spotlight. Nevertheless, interest in LLT surveys among academics and authors are remarkable. Studies in LLT show enormous potential to contribute to our perceptions of the assessed structures and our theoretical arguments about validity in the field. Nonetheless, the principal objective of local language testers' engagement in research validation aims at enhancing the LLT practices and providing effectual arguments in debating with local stakeholders. Although the diverseness in Language Testing (LT) practices might bring about issues in establishing specific methodological procedures in LT, if the publication trend of LLT studies boost over time, definite 'contextual, practical, and methodological patterns' are presumably to emerge.

LLTs are promoted for utilization within a particular context and to present numerous objectives and goals, containing "entry-level proficiency testing, admissions testing, placement testing (PT), international teaching assistant testing, testing for immigration purposes, and program evaluation" (2). These objectives might not be directed and applied effectively by dint of either widely-purposed, regulated, commercial implements or particularly-intended classroom evaluations. Challenges linked to LLT practices are appropriate across various languages, contexts, and scope. Thanks to the expansion of English, it is undeniable that the greater number of chapters displayed in the book concentrates on the evaluation of English.

This edited, salient, concise yet grounded book by the editors is constructed into 11 chapters plus a prologue by the editors. The book might target a broad readership including teachers-researchers with the need to establish and expand LLTs. It clearly delineates LT research and practices across 'linguistic, cultural, socio-political and geographical contexts' as well. Additionally, it encompasses practical topics directly applicable to test development validation and use. It is noteworthy that the present book is part of the book series of volume 61 in Educational Linguistics (EDUL). EDUL is dedicated to innovative studies of language use and language learning.

The chapters tackle with pressing needs for evaluation that appear in the local context and may not be applied perfectly with the existent large scale, international English tests. Whilst only a single chapter (chapter 6) is concerned with English testing in mandatory education, majority of them cope with ESL at tertiary education level, where the writers deal with multiple issues, including the association between "test analysis, test quality, and test administration" in The Oral English Test (OET) (chapter 1), validation of the Czech Language Certificate Exam (CCE) with respect to the local context (chapter 2), a report on an exploratory CEFR mapping study involving a local ESL proficiency and PT at a Canadian university (chapter 3), designing a new diagnostic reading test for a local post-admission assessment program (chapter 4), provision of an account of a locally-developed Elicited Imitation (EI) test for instructional program evaluation (chapter 5), a review of PT procedures for French

and Chinese languages at a large Australian university (chapter 7), the Italian LT in chapter 8, as well as pertinent LLT related to particular disciplinary fields and courses (chapters 9 and 10), and, eventually, the assessment of L2 Learners' Oral Skills in a Local LT context as evidence of Tourism English ability in chapter 11.

Thematically, each chapter showcases description of the LT and/or instructional context; it struggles to diagnose the local need for test design or revision, and furnishes perceptions and outlooks obtained from test development practices and data analyses. The chapters are carried out through numerous layers to serve and concentrate on particular issues such as (a) Test Purposes (chapters 2 to 6), (b) LLTs as Policy Instruments (chapters 7 and 8), and (c) Assessment of Language for Specific Purposes (LSP) (chapters 9,10 and 11). Technically, this book elaborates LLT practices that occur in the intermediate space between large-scale standardized testing and classroom evaluation. Drawing experimental studies on various languages, the book argues' capability to indicate that test results have accurately and deliberately lodged local contexts and values within instructional practice, and have applied data for program assessment and surveys. The researchers in this volume supply the theoretical outlooks and concepts with practical instances of how LTs can be constructed to efficiently function within and across various institutional contexts.

Chapter 1 by Ren and others explores the resemblance of prompts utilized in an office-hour role-play task on an in-house International Teaching Assistants (ITA) test. It is concerned with the application of LLT for monitoring a standardized, locally designed PT for ITAs in one of the largest public institutions in the US. As the method of analysis, Many-Faceted Rasch Measurement (MFRM) is applied to probe “prompt difficulty, rubric consistency, and selected interactions” (13). Their study furnishes useful insight into practices that combine “test analysis, test quality, and test administration in a local context” (28). Their probe indicates that raters showcase preferences for special prompts when choosing prompts.

Chapter 2 by Hulešová and Vodičková tackles three inter-connected and fundamental issues in their assessment of the CCE. Their research, carried out over 7 years, was performed (1) to highlight advances proposed by an ALTE certification project concerning re-arrangement for the CEFR, and (2) to offer an account of how CCE level cut scores to adjust with the local, university-wide requirement that all passing cut scores be set at 60 %, and (3) to inspect/organize stable item difficulty estimates, permitting their involvement in a CCE item bank. For the listening and reading subsets, they have chosen the Direct Comparison procedure and, for the writing and speaking subsections, they have applied a modified Body of Work method.

It is argued that one way to improve the appropriateness of the LT for diversified aims is to adjust the test with nationwide or worldwide standards. In their chapter 3, Suzanne Springer and Martyna Kozłowska indicate the alignment processes conducted in order to (1) address a descriptive CEFR mapping survey including a local ESL proficiency and PT at a Canadian

university, (2) to corroborate their contingent claim that there are interconnections between the CEFR benchmarks and their course content and purposes, and (3) to assist instructors to promote CEFR benchmarked materials (as a programmatic use).

In the next chapter, chapter 4, Liu and Read manifest that it is both necessary and attainable to construct and/or re-design LLT based on local needs and contexts. Primarily, they bring up issues concerning a mismatch between the current DELNA reading assessment and the more expansive set of academic reading skills. Then, they proceed to re-design the given test, incorporating learners' needs in a principled way, so that more fine-grained information could be presented to students and language advisor.

Cheng and Li in chapter 5 address another challenge requiring an extension of test's purposes; their survey accentuates the significance of the association between testing and instruction for a local program. Indeed, the enduring presence of the scheme, in which the test was entrenched, was anticipated on learners' demonstration of acquisition in proficiency after a two-term, two-course sequence of teaching.

Härmälä and Hilden in chapter 6 consider a wider scope of LLT by investigating the alignment of the written, official Finnish curriculum and a national test of English-language learning outcomes at the end of 6th grade. The requirement for expansion of LLTs often proceeds as a consequence of the application and fulfilment of a covert or an overt educational or Language Policy (LP) in a specific context.

Even though each LLT displayed in the chapters of the book respond to LPs, two chapters' (7 and 8) particular focal point is on challenges related to the role of LLTs in the implementation of fair and just LPs. In their chapter 7, Chen and others inspect the role of LLT to formulate and design fair placement policies of learners at suitable proficiency level courses at the University of Melbourne for Chinese and French. The chapter accentuates problems arising from stakeholders' conflicting conceptions of the roles of 'equity and equality', which are the main concerns for the formulation of fair placement policies. The stipulate that the optimal and perfect upshot of proper and impartial placement in language courses would contain learners' contentment and confidence with placement along with homogenous classes so that the classes are not too challenging for some and too simple for others. Machetti and Masillo in the next chapter, chapter 8, on the other hand, pose a significant topic regarding the role of LLT as a policy tool in the European context. They voice concerns about an employment of LTs as part of an immigration policy because the misuse of scores may give rise to 'discrimination, exclusion, and injustice' particularly if they lack validity.

LSP tests, based on their scope and stakes, can be either local or large-scale. These tests tackle with LLT local test traits if they are constructed to evaluate domain-specific language in a specific Target Language Use (TLU), with the purpose of impacting upon the local context. The three remaining chapters (9, 10 and 11) in the volume cope with local LSP testing

practices with the intention to evaluate to what extent the course objectives and aims were reached in various disciplinary courses.

As a devised video-based mock tests revealing phonological errors, Larose Luigi in chapter 9 poses a consistent issue that LLT examiners, instructors, and program administrators encounter, i.e. how to make scores on LLTs to be fruitful both for teaching and learning. In his research, Larose addresses this issue in assessing British Sign Language (BSL) in his university context, where the learners make slow progress on promoting their phonological skills.

In chapter 10, Wudthayagorn and Limgomolvilas, by creating an indigenous rubric, concentrate on testing English LSP in pharmaceutical sciences for Thai pharmacy students' dispensing drug skills. Then, to further validate the efficiency of the rubric, they conduct MFRM on scores raters assigned "based on the rubric and qualitatively analyzed interview data with the raters about their perceptions of the rubric and students' performance" (223). The results of the MFRM were in line with the positive effects reported in rater interviews.

Based on a mixed-method research design, in the last chapter, chapter 11, Karatay, primarily, designed a task-based Tourism English oral performance assessment utilizing a particular spoken dialogue system (SDS) in which a computer program was created to act as a hotel guest and examinees respond as a hotel employee. Results indicated that a dialogue-based oral evaluation in ESP programs might culminate positive washback in teaching and learning by presenting more communicatively-oriented language instruction than had been the case prior to the introduction of the assessment.

Even though per chapter in this edited book manifests a particular context with specific LLT needs, the prevalent characteristic among most chapters is the constant promotion of LLTs based on urgent feedback from multiple local actors or based on the altering needs in the local context. The integral procedures of construct testing and definition plant firm groundwork for consequent score use and interpretation. Eventually, it is argued that LLT practice, at the same time, calls for an ample investment of time and labor, and all of the improvement attempts inspected and elaborated in this volume were performed across multiple years.

Funding

The work was supported by Hebei North University 2022 City-School Deep Integration Special Project for Research and Practice of Education-Teaching Reform: "Research on the Construction of Foreign Language Education Community in Zhangjiakou" (SXRHJG202205); Hebei North University 2023 Provincial University Basic Research Business Fee Project: "Course Construction of 'Introduction to Chinese Culture (Russian)' --- based on International Online Collaborative Teaching Model" (JYT2023025).