

BOOK REVIEW



Review of *Technology-assisted language assessment in diverse contexts: Lessons from the transition to online testing during COVID-19*

Mina Hashemi, University of Western Ontario

Technology-assisted language assessment in diverse contexts: Lessons from the transition to online testing during COVID-19

Karim Sadeghi (Ed.)

2023

ISBN: 9781032117690

US \$ 35.96

274 pp.

Routledge

New York, NY, USA



The onset of the COVID-19 pandemic posed new challenges for second language (L2) education and language assessment in particular. The global health crisis required educators to move from on-site- to emergency-remote teaching (ERT), which was characterized by the rapid design of temporary instructional interventions and the search for solutions to immediate problems (Bond et al., 2021). In the edited book, *Technology-Assisted Language Assessment in Diverse Contexts: Lessons from the Transition to Online Testing During COVID-19* (Sadeghi, 2023), the editor and the authors present a series of conceptual discussions and empirical studies about the ways in which language assessment practices were directly impacted and shaped by the pandemic. The volume consists of 16 chapters grouped into three sections: theoretical and methodological concerns in online L2 assessment (Chapters 1–6), reactions to L2 e-assessment during the COVID-19 pandemic (Chapters 7–11), and managing L2 assessment at the time of the crisis: the way forward (Chapters 12–16).

In Chapter 1, Sadeghi provides a historical overview of technology use in second language assessment, pointing out that the coupling of assessment and technology is not new. The author also discusses the challenges and affordances of technology-mediated language assessment. Whereas Sadeghi references an array of studies exploring the advantages of computer-assisted language testing, the author recognizes the dearth of research surrounding security and fairness. Sadeghi explains that the main aim of the book is to investigate the ways in which the pandemic was potentially “a blessing in disguise” (p. 3) for the field of L2 assessment, as it challenged the premise that technology is merely an addition to language testing, when in fact technology and assessment are inseparable.

In Chapter 2, Clark et al. argue for the validity and security of the IELTS Indicator—a high-stakes, at-home test that functions as a substitute for the traditional IELTS—whose release was expedited after COVID-19. The authors report on test-takers’ and examiners’ views and experiences with IELTS Indicator gathered via surveys and focus groups. Their findings indicate that despite unresolved technical issues at the outset of

the pandemic, most participants expressed satisfaction with their at-home test experience, mainly due to reduced stress attributed to familiar surroundings. At the same time, when discussing the lessons learned from the study, the authors acknowledge that variable home environments pose a challenge to the reliability of online assessments.

In Chapter 3, Bruce and Stakounis discuss the transition from in-person to online assessments during the pandemic, and the validity and longevity of those assessments in a UK English for Academic Purposes (EAP) context. The authors mention that since test-takers are likely unfamiliar with the online platform, delivering a test online when it was intended to be paper-based raises validity concerns. Therefore, the authors recommend using validity frameworks as guidelines for developing and delivering online tests. In addition, the authors identified four ways assessments moved online: (a) in-person tests remained the same but were delivered online, (b) in-person tests were modified for online use, (c) new tests were developed for the sole purpose of online assessment, and (d) some assessments were canceled. Ultimately, issues around fairness, integrity, literacy, and technology led to the decision to adapt, develop, or cancel certain assessments.

In Chapter 4, Farhady focuses on learner-oriented assessment (LOA). He evaluates the attitudes of a group of learners who took a course that incorporated LOA during the pandemic. According to Farhady, students welcomed this form of assessment since most LOA tasks took place outside the classroom, thereby lessening the pressure of being assessed. Farhady reports that four lessons were learned from this examination of LOA: (a) teachers must create a channel for communication with learners by asking them to write reflective essays after each session, (b) there needs to be more encouragement of student-student collaboration as peer- and self-assessment are key to improving students' involvement in learning, (c) teachers must guide learners through giving one another feedback before they receive teacher feedback, and (d) teachers should focus more on validity and reliability, which LOA can somewhat guarantee as it minimizes academic dishonesty.

In Chapter 5, Voss investigates argument-based validation of online assessments, which relies on a series of inferences. Within the argument-based validation framework, each inference is associated with claims based on warrants and assumptions supported (or refuted) by evidence (Chapelle & Voss, 2013). From the author's conceptual discussion, three main technology-related lessons emerge. First, available technology that supports the intended purpose of assessment must be identified. Second, as online testing becomes more commonplace, technology must be acknowledged as an integral part of assessment. Finally, technological knowledge must be developed among all stakeholders, including teachers, administrators, and test-takers.

In Chapter 6, Amer and Cabrera-Puche discuss the development of a mobile application and complimentary software for a placement test. The aims of their project were to “think about *assessment without borders*” and to innovate utilizing available technology (p. 96). A unique aspect of the software is that the placement test can be used with heritage language learners, and it provides information on issues related to distinguishing these students from other students. The authors' software and adjunct mobile application offer placement exams in six languages, including Chinese, French, Italian, German, Russian, and Spanish. The authors report that their experience with the placement tests during the pandemic encouraged them to be flexible and adaptable, particularly when working with learners in diverse contexts.

In the first chapter of Section II, East et al. examine the assessments that were administered in two L2 French courses at the University of Auckland over two semesters in 2020. During that time, the delivery of classes switched back and forth between on-site and online modes of instruction. The authors explain that for the online assessments to be fair, the assessment criteria and components underwent a few modifications (i.e., some components were removed, and the weighting of other components was adjusted). In the second semester, the course exams remained “as is” but moved online, and the instructions for the assessment were modified to reduce cheating. According to East et al., the lesson of this study is that there is no one-size-fits-all technology-mediated assessment in terms of validity, reliability, and fairness, as modifications will be needed to meet the intended purpose(s) of assessment.

In Chapter 8, Balteiro investigates students' perceptions of pandemic-induced changes to university education in Spain, such as moving classes from in-person to online. Despite a few negative perceptions, the authors report that students had an overall positive view of the changes in teaching and assessment necessitated by the transition to online instruction. Some negative views included a lack of attention to the assessment of oral skills, inadequate time provisions for taking online exams (since time limits were set to minimize the likelihood of cheating), and nervousness surrounding Internet connection problems. According to Balteiro, a key finding from the research is that there needs to be reliable platforms, which allow students to take their exams without worrying about a loss of connection or an administrator assuming that they are cheating.

In Chapter 9, Hardacre focuses on the assessment of ESL writing at a public university in Southern California. Because of COVID-19, writing assessments in an upper-level academic writing course had to be moved from an on-site, fully proctored format to a remote, unsupervised environment. The author determined that the change was effective since all students were able to meet the learning objectives, supported by Internet-based resources such as dictionaries and thesauri. When asked about their views, students expressed mostly positive attitudes. They believed their confidence had improved, and they enjoyed seeing their progress. Ultimately, Hardacre concluded that using technology facilitated the achievement of learning objectives as long as students had the chance to collaborate with peers and to use Internet-based resources to work on their writing skills.

In Chapter 10, Neiriz et al. report modifications to an English placement test at Iowa State University. The placement test already had an online equivalent, in addition to an in-person version prior to the pandemic, but during the surge in COVID-19 cases, the writing skills and oral communication sections of the test had to be administered online only. The main challenges with the writing skills part of the test were security and cheating, along with testing conditions and test access. The authors argue that the online test taken at home is authentic, but it lacks reliability, as at-home conditions differ across examinees. Interestingly, the oral communication component of the test had originally been done outdoors, which occasioned issues such as a noisy outdoor environment and no waiting area. Upon moving the oral component online, these issues were resolved, but new issues arose surrounding test security and fairness, with certain modifications benefitting students unequally.

In Chapter 11, Carrió-Pastor explores the assessment of data-driven learning (DDL) through digital learning objects, defined as 10-minute recorded presentations on a specific topic. For this project, Spanish university-level learners of English were asked to do a multimodal presentation of a DDL activity involving the use of corpora for language learning. The learners were given instructions and rubrics and taught about implementation and collaboration through a series of meetings on Microsoft Teams. After their presentations, the learners had to compose a reflection. Having analyzed these reflections, the author reported that the assessment engaged the learners and that the multimodal assessment of digital learning objects was effective.

Marking the beginning of Section III, Chapter 12 features Lam's analysis of the usefulness of an e-portfolio writing assessment in terms of validity, reliability, washback, and practicality. Lam identifies two types of e-portfolios: a strong and a weak version, where the weak version takes place online as part of the general online classroom assessment scheme, and the strong version encompasses assessment outside of the classroom scheme. To create e-portfolios, all students had to draft and plan each step of the writing process in Google Docs so the teacher and other students could keep track of their work. For the author, the opportunity to monitor students' progress over time shows that e-portfolios can be a viable alternative to traditional forms of assessment.

In Chapter 13, Zhang and Isaacs examine the use of videoconferencing for assessing oral pragmatic competence from perspectives of practicality, authenticity, reliability, impact, validity, interactivity, and accessibility. They discuss the benefits of using video-conferencing tools for assessing speaking, such as remote test delivery and increased cost-effectiveness, as well as the constraints, including threats to reliability caused by technical issues and a lack of extensive access to non-verbal cues.

In Chapter 14, Ducasse presents a case study of several Spanish language assessment tasks and the adjustments that had to be made to the tasks, the rubrics, and the feedback due to the pandemic. Carried out in the Australian higher education context, this case study, which documented the experiences of the researcher as a teacher, explores the stages of designing, developing, and administering assessment tasks. The author explains that she had to improve her digital literacy to smooth the transition from on-site to online assessment for herself and her students. In the end, the assessment tasks required a significant amount of feedback, and some students were demotivated by the online aspect due to technical difficulties with submitting or opening files.

In Chapter 15, Imamura et al. describe the development of a Japanese placement test comprising an asynchronous online test and an oral interview at a large public university in the U.S. The four sections of the test (i.e., grammar, reading, listening, and writing) contained questions of varying difficulty levels, which corresponded to the four levels of Japanese language courses available at the university. The questions used for each level were items taken from the final exam of the corresponding course. Although this adaptive online placement test was flexible and efficient, the Qualtrics tool utilized for test delivery and automatic scoring was found to be inadequate since the platform was designed primarily as a tool for creating and delivering surveys.

In the concluding chapter, Sadeghi reflects on the lessons learned from the studies included in the volume, highlighting the value of the authors' findings for shaping online assessments after the pandemic. The author concludes by expressing regret about lessons that might not have been learned during COVID-19, commenting on how the pandemic seemed to fuel new forms of war instead of bringing humanity together.

Overall, this edited volume offers a collection of cutting-edge studies on L2 assessment during the COVID-19 pandemic. The empirical studies cover a wide range of contexts (e.g., high-stakes, low-stakes, placement, and EAP), languages (e.g., English, Spanish, and French), foci (e.g., writing, speaking, and alternative assessments), and aspects (e.g., validity, fairness, and student and teacher perceptions). The volume, however, is not without weaknesses. For example, with the exception of Chapter 3, the studies did not examine any changes made to the assessment of receptive skills. In sum, this volume offers a valuable contribution to the field of second language assessment and technology-mediated language testing as it provides insights not only for researchers but also for L2 teachers and test developers.

References

- Bond, M., Bedenlier, S., Marín, V. I., & Händel, M. (2021). Emergency remote teaching in higher education: Mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 18(1), p. 1–24. <https://doi.org/10.1186/s41239-021-00282-x>
- Chapelle, C. A., & Voss, E. (2013). Evaluation of language tests through validation research. In A. J. Kunnan (Ed.), *The companion to language assessment: Vol. III. Evaluation, methodology, and interdisciplinary themes* (pp. 1–17). Wiley-Blackwell. <https://doi.org/10.1002/9781118411360.wbcla110>

About the author

Mina Hashemi (Master of Professional Education, University of Western Ontario) is an EAP instructor at Western English Language Centre (WELC). Her research interests include L2 assessment and the assessment of writing.

E-mail: shashe32@uwo.ca