

COVID-19 is Still Here: Discussing Home-based Online Assessments

Asiah Kassim

Centre for Modern Languages, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.

ABSTRACT – At present, assessments which are part and parcel of the teaching and learning process, have gone through leaps and bounds transiting from the conventional face-to-face setting to home-based online platforms as a consequence of the COVID-19 pandemic that is still looming worldwide after more than a year from its first hit. This move is made possible with technology support that has been integrated in all language assessments of different skills. This editorial section previews some issues discussed in literature in relation to the assessments conducted online on four language skills – listening, reading, speaking and writing. The discussion attempts to relate the use of varied digital means and technological aids to the efficacy of the assessments' constructs in light of their benefits and challenges faced in the transition and administration phases. The section concludes with the call for collaboration between language practitioners and researchers on a platform where ideas and knowledge can be exchanged and eventually implemented in language classes and assessments to facilitate the enhancement of learners learning experience and achievements.

ARTICLE HISTORY

Revised: 25 July 2021

Accepted: 28 July 2021

KEYWORDS

COVID-19

Language assessment

Language skill

Home-based online assessment

INTRODUCTION

Language assessments, defined as a continuous process of measuring learners' performance in a larger realm of learning that can take a form of self, teacher or even peer evaluation (Brown, 2004) has been greatly affected by the current unpredictable conditions caused by the spread of COVID-19 worldwide. The pandemic that first hit in the early 2020 has forced practitioners in the teaching and learning field to muster ways and approaches of more digitally-inclined assessments in order to brave through the uncertainty of this worldwide predicament. A more focused form of an assessment, the language testing, has been devised to cater for the different needs of the test-takers in this circumstance while attempting at its best to measure what was originally set to measure regardless of the settings or approaches of the test constructed.

In today's context where the pandemic's intention of stay is indefinite and our dependence on the use of technology has increased, there are a good number of issues that need to be addressed in the context of language assessments. The assessments that were originally designed for the conventional physical face-to-face classrooms, would they still be appropriate and effective in virtual classrooms as a form of measurement of learning outcomes set for a course. Do they still assess what was meant to be assessed in conventional classrooms? Other aspects that naturally come together with the move from the conventional face-to-face assessments to the remote online mode, like security, internet access, related resources, and additional, unpredictable costs, have to also be taken into consideration to ensure the efficiency of conducting the assessments. This paper discusses literature on the use of online platform assessments considering some pertinent issues reported in previous studies in an attempt to infer the efficacy of the move from the conventional mode of assessments to the fully online mode in attaining the learning outcomes set for each language skill assessed.

LISTENING ASSESSMENT

Carrying out listening assessment that was originally meant for conventional face-to-face setting, may be a challenge for virtual classes like ones that are administered on online platforms such as Google Meet, Zoom or on Teams. With the learners' level of familiarity of various functions of these meeting apps and differences in adhering to the integrity and honesty in test taking, it is a great feat to ensure fairness and equality is achieved, for instance, when some learners turn on the captions while taking the listening test, will somehow render some other learners who do not turn on this function as disadvantageous. Certain functions that come together with available online platforms, such as that of captions or transcriptions, may be a boon or a bane depending on the test constructs and the learning outcome that each assessment is attempting to achieve.

Almost a decade before the pandemic hit the globe, Ghasemolani and Nafissi (2012) conducted a study on the use of caption in listening comprehension test, suggesting facilitative effect this function has on enhancing learners' listening comprehension. A more recent study by Chen and Chen (2021) made use of video-annotation function comparing its efficacy to that of self-determined review mechanism used for videos on Youtube. Findings suggest that the more advanced use of video annotated mechanism has significantly helped learners to enhance their listening comprehension. Even though the two studies are incomparable in terms of their aims and variables focused, in both studies, enhancement

to listening comprehension is achieved with the facilitation of the available functions of the online video that the participants were required to watch, such as caption and with added advanced review mechanism in the Chen and Chen study, facilitating the learners' performance. It can be inferred that the studies utilized digital tools to assess their efficacy in enhancing learners' listening comprehension achievement and both studies yielded positive effects of the tools utilized in listening assessment.

So, this begs a question of whether these technological functions should be allowed in course assessments. Should they be allowed to be used, other questions entail, for example, what question types should be constructed or if the different question types are developed, should the learning outcome set for the course be revisited? What this circumstance is indicating is that all aspects of a course should be taken into account when consideration is made to use certain technological assistance in an assessment.

READING ASSESSMENT

The use of digital support in reading assessment has been in the picture for decades. Bennet (2003) for example, suggested a number of benefits computer-based reading assessment can have on both the test administrators and test-takers, for instance, time and cost effectiveness and immediate feedback. On a similar note, Higgins et al. (2005) and Kim and Kim (2013) reported in their studies on learners' perspectives in using digital tools to complete assessments of which familiarity and comfort are the main aspects that made computer-based assessments to increase in their use. High-stake examinations are also turning to this mode of assessment, such as TOEFL iBT and more recently, IELTS (for listening, reading and writing). When the pandemic forced assessments to be carried out completely online, all levels of testing, from the in-house assessments to high-stake examinations moved from test-centre computer-based testing to remote home-based online assessments.

Reading assessment that was already conducted on computer-based mode, like TOEFL did not seem to go through much a hassle in this transition and was able to make a "swift and effective" transition to a home-based platform taking into great considerations the security aspects of the test administrations (Papageorgiou & Manna, 2021). The move that took place led to very small changes, if any, to the testing of this language skill in terms of its test-item constructions, leaving room for concern to more external issues which mainly focused on the security of test administrations. In other words, converting the questions that were once meant for physical face-to-face assessments to any of the online platforms, such as Qualtrics (Purpura et al., 2021) and Canvas (Green & Lung, 2021) did not affect the intended language ability the tests were meant to measure, which is nevertheless, a still debatable issue (Isbell & Krammer, 2020).

SPEAKING ASSESSMENT

Assessing the micro- and macro-skills of learners' speaking competence will need different types of digital platforms due to the differences in the focus of the speech productions. Brown (2004) categorised micro skills as the production of "smaller chunks of language such as phonemes, morphemes, words, collocations and phrasal units" (p. 142), which are mostly assessed in imitative and intensive taxonomies of speech production. "[F]luency, discourse, function, style, cohesion, nonverbal communication, and strategic options" are mainly the macro skills assessed in learners' speech (p. 142). While a number of existing tests measure the imitative and intensive ability of test-takers like Duolingo English Test (DET), others are more inclined to measure responsive and interactive skills, such as IELTS and TOEFL.

Assessing the ability to imitate a word, phrase or sentence is basically evaluating the learners' pronunciation, with no consideration to the learners' ability to understand or relate meaning in the production of speech. Where semantic is concerned, intensive ability is measured generally to assess learners' ability to relate grammatical, vocabulary and phonological elements to the utterance requiring the test takers to minimally interact with an interlocutor, such as in directed-response tasks and dialogue completion. For these two types of speaking assessments, computer generated scoring is possible as has been used in DET and Versant. Since these tests have been from their inception an online testing with no limitation of the confines of test-centres, provided that the test takers can fulfill all the security measures, there has been no pressing issue in the transfer process of the test to home-based as other tests, such as IELTS, a high stakes exam and many other in-house speaking assessments (e.g., Clark et al., 2021; Ockey et al., 2021). Some speaking tests, like IELTS, which measure mostly responsive and interactive elements, may have to go through an extensive work for the transfer from the conventional face-to-face setting to a remote online environment. In the context of IELTS, in an attempt to retain the interpersonal aspects of the testing, speaking test is conducted via Zoom with a real IELTS examiner.

The final type of speaking assessment, which is extensive oral production ability such as speeches and presentations, is mostly in-house speaking assessment conducted by many learning institutions (e.g., Green & Lung, 2021; Wagner & Krylova, 2020). Online speaking assessment can adopt two types of tests, categorized as semi-direct and fully automated speaking tests (Bernstein et al., 2010), with the former utilizing computers for test administrations involving test takers taking the test and collecting the responses, eventually rated by trained human examiners. The latter, on the other hand, does not involve human raters where a speaking assessment is fully computerized from the test development to the scoring process. A number of benefits of computerized speaking assessment have been reported in previous studies, which include among others cost effectiveness, tests' reliability and integrity as well catering to the demographics and geographical aspects of the test-takers (Jamieson et al., 2013; Qian, 2009). These benefits may suggest that assessing speaking on online platforms, after the move from the conventional mode of assessment, and after careful considerations of the security measures, does not compromise the learning outcomes the assessment is set to measure.

WRITING ASSESSMENT

Writing assessment is of no exceptions when it comes to being affected by the current setting. As with the other language skills assessments, the pandemic has forced writing assessment to also move to a fully remote online platform. Computerised writing assessment has actually been proposed and researched on for decades, emerging as early as the 1960s (Shermis et al., 2001), making use of a number of digital assistances as employed at present, in high stakes examinations like TOEFL and DET as well as in-house assessments and classroom tasks mostly utilizing automated writing evaluation (AWE) software such as Write and Improve, e-rater and Criterion (Keith, 2003). Results of previous studies suggest varied findings indicating facilitative aspects of the software as well as motions to further work on these digital tools to better assist learners in their writing tasks and assessments (e.g., Ranalli et al., 2017; Waer, 2021).

In addition to this notion, there are also predicaments that online writing assessment can bring to the table. Assessment conducted online give a huge space for the learners to plagiarise by simply browsing the internet while writing their essay or any written productions. The use of antiplagiarism software and the constant reminders given to the learners to always be honest and have integrity in completing an assessment may help deter this plagiarism activity, but for some learners, getting assistance from Mr. Google, for example, is just irresistible. The more enterprising test administrators make use of security software such as Secure Exam Browser (SEB) utilised in IELTS and TOEFL, that can inhibit test takers to use other applications on their computer since the software functions to disable the access to other applications throughout the test (Isbell & Krammer, 2020; Papageorgiou & Manna, 2021). In this sense, the security of an assessment is well guaranteed, giving almost no space for the learners to stray and use for instance, the internet on their device, to aid their writing test. In other words, with an investment into security applications and thorough proctoring system, writing assessment conducted remotely may not raise any issues in terms of intervening variables that may affect the intended learning outcomes set to measure learners' development and achievement on writing skills.

REFLECTION AND CONCLUSION

One of the main issues raised in administering remote online assessments for each language skill is the security of the tests. There are commercial apps equipped with proctoring security as mentioned above that can be used for a more efficient administration of the assessments, but the fact that the test takers are not within the same physical space as the instructors makes monitoring still very much a problem. Proctoring an assessment is a matter that has to some extent been addressed by employing services to aid monitor testing either with the presence of human or computer-based proctor, such as ProctorU used by TOEFL (Isbell & Krammer, 2020), and Honorlock, integrated into an in-house online placement test at Teachers College, Columbia University (Purpura et al., 2021). These services require the test-takers to have certain technical requirements on their computers and good access to the internet for the proctors to function efficiently. Artificial intelligence (AI) and data analytics are also utilized by many security service providers to ensure secured and reliable tests are conducted which include among others tackling issues of the test items exposure, demographic elements and time frame of the test takers to minimize possibilities of test fraud or cheating to occur. There are also tests that are administered depending solely on test takers' honesty, for example, TEF Express (Test d'évaluation de français) which only requires the test takers to sign a non-disclosure agreement and based on trust that they will uphold the integrity and not commit any actions that can compromise the security of the test (Isbell & Krammer, 2020; Wagner & Krylova, 2021).

In relation to these issues, considerations should also be given to instances where access to these commercial apps is financially a challenge, forcing both educators and learners to be more 'resourceful' in the selection of technological assistance to use in language classes and assessments. According to a UNICEF (2020) press release, 463 million children, which make up approximately a third of children in their schooling years worldwide do not have access to learning remotely, indicating huge gap in the access to online learning comparing children from a more privileged environment to those with very limited or no means to this mode of learning. In Malaysia, a survey conducted by the Ministry of Education indicates that 1.85 million learners neither have access to the internet nor possess suitable gadgets for learning, denying their chances to be part of the online home-based learning (Abdul Rahman, 2021). In these dire conditions, conducting assessment is not just a matter of reconstructing test-items or security or any of the issues discussed above, but it is fundamentally a matter of making it possible to even conduct an assessment. Teaching and learning must go on and closure of schools should not be a hindrance to the learning process. However, such pressing issue that is out of the hands of educators should be carefully threaded and making use of available resources if any, should be prioritised and customizing them as best as we can to ensure learning outcome is at least partially achieved to minimise what is predicted by UNICEF, the "repercussions" of this learners' attrition during their critical years on the "economies and societies for decades to come" (para. 2).

Empirical studies and evidence are very much needed for the language practitioners to make use of what is suitable and efficient in virtual classes and remote home-based assessments. With the pandemic unexpected visit and unpredictable intention of stay leading to the physical distance that separates us educators and our learners, it is acceptable to claim that the incessant ubiquity of technology use is inevitable. Inference can be made that technology use aids the process of carrying out assessments to some extent, but there are still a number of issues that need to be addressed on a continuous basis as discussed in the earlier sections of this paper. It is also noteworthy to point out that discussion in this article is focused mainly on technology use related to the availability of the internet. Other technological or any learning aids should also be considered in order to cater for varied learners' access to learning. There should be a channel where language practitioners can share first-hand experience integrating technology and other learning aids, the trials and tribulations faced, the ups and downs and of course practical and feasible solutions not just for language assessments but

for other aspects of language teaching and learning, while considering the means and access of learners from diverse backgrounds. This bilateral collaboration between researchers and practitioners is necessary for all of us to brave through these turbulent times and putting our learners' learning experience and achievements our utmost priority.

REFERENCES

- Abdul Rahman, M. N. (2021, June 19). Perbaiki pelaksanaan PdPR 2.0 atasi isu keciciran pelajar. *Berita Harian Online*. <https://www.bharian.com.my/kolumnis/2021/06/829593/perbaiki-pelaksanaan-pdpr-20-atasi-isu-keciciran-pelajar>.
- Bennett, R. E. (2003). *Online assessment and the comparability of score meaning*. Educational Testing Service.
- Bernstein, J., Van Moere, A., & Cheng, J. (2010). Validating automated speaking tests. *Language Testing*, 27, 355–377. <https://doi.org/10.1177/0265532210364404>.
- Brown, D. H. (2004). *Language Assessment: Principles and Classroom Practices*. Pearson Education.
- Chen, C., & Chen, I. (2021). The effects of video-annotated listening review mechanism on promoting EFL listening comprehension. *Interactive Learning Environments*, 29(1), 83-97. <https://doi.org/10.1080/10494820.2019.1579232>.
- Clark, T., Spiby, R., & Tasviri, R. (2021). Crisis, collaboration, recovery: IELTS and COVID-19. *Language Assessment Quarterly*, 18(1), 17–25. <https://doi.org/10.1080/15434303.2020.1866575>.
- Ghasemband, F., & Nafissi, Z. (2012). The effects of using English captions on Iranian EFL students' listening comprehension. *Procedia - Social and Behavioral Sciences*, 64, 105 – 112. <https://doi.org/10.1016/j.sbspro.2012.11.013>.
- Green, B. A., & Lung, Y. S. M. (2021). English language placement testing at BYU-Hawaii in the time of COVID-19. *Language Assessment Quarterly*, 18(1), 6-11. <https://doi.org/10.1080/15434303.2020.1863966>.
- Higgins, J., Russell, M., & Hoffmann, T. (2005). Examining the effect of computer-based passage presentation of reading test performance. *The Journal of Technology, Learning and Assessment*, 3(4), 3-35.
- Isbell, D. R., & Krammer, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*, 37(4), 600-619. <https://doi.org/10.1177/0265532220943483>.
- Jamieson, J., Wang, L., & Church, J. (2013). In-house or commercial speaking tests: Evaluating strengths for EAP placement. *Journal of English for Academic Purposes*, 12(4), 288–298. <https://doi.org/10.1016/j.jeap.2013.09.003>.
- Keith, T. Z. (2003). Validity and automated essay scoring systems. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective*. (pp. 147–167). Lawrence Erlbaum Associates Inc.
- Kim, H., & Kim, J. (2013). Reading from an LCD monitor versus paper: Teenagers' reading performance. *International Journal of Research Studies in Educational Technology*, 2(1), 1-10.
- Ockey, G. J., Muhammad, A. A., Prasetyo, A. H., Elnegahy, S., Kochem, T., Neiriz, R., Kim, H., & Beck, J. (2021). Iowa State University's English placement test of oral communication in times of COVID-19. *Language Assessment Quarterly*, 18(1), 26-35. <https://doi.org/10.1080/15434303.2020.1862122>.
- Papageorgiou, S. & Manna, V. F. (2021). Maintaining access to a large-scale test of academic language proficiency during the pandemic: The launch of TOEFL iBT Home Edition. *Language Assessment Quarterly*, 18(1), 36-41. <https://doi.org/10.1080/15434303.2020.1864376>.
- Purpura, J. E., Davoodifard, M., & Voss, E. (2021). Conversion to Remote Proctoring of the Community English Language Program Online Placement Exam at Teachers College, Columbia University. *Language Assessment Quarterly*, 18(1), 42-50. <https://doi.org/10.1080/15434303.2020.1867145>.
- Qian, D. D. (2009). Comparing direct and semi-direct modes for speaking assessment: Affective effects on test takers. *Language Assessment Quarterly*, 6(2), 113–125. <https://doi.org/10.1080/15434300902800059>.
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8-25. <https://doi.org/10.1080/01443410.2015.1136407>.
- Shermis, M. D., Mzumara, H. R., Olson, J., & Harrington S. (2001). On-line grading of student essays: PEG goes on the World Wide Web. *Assessment & Evaluation in Higher Education*, 26(3), 247–259. <https://doi.org/10.1080/02602930120052404>
- UNICEF. (2020, August 26). *COVID-19: At least a third of the world's schoolchildren unable to access remote learning during school closures*. UNICEF. <https://www.unicef.org/press-releases/covid-19-least-third-worlds-schoolchildren-unable-access-remote-learning-during>.
- Waer, H. (2021). The effect of integrating automated writing evaluation on EFL writing apprehension and grammatical knowledge. *Innovation in Language Learning and Teaching*, 1-25. <https://doi.org/10.1080/17501229.2021.1914062>.
- Wagner, E. & Krylova, A. (2021). Temple University's ITA Placement Test in Times of COVID-19. *Language Assessment Quarterly*, 18(1), 12-16. <https://doi.org/10.1080/15434303.2020.1862849>.