Test Takers' Perceptions of Design and Implementation of an Online Language Testing System at a Thai University during the COVID-19 Pandemic

Kornwipa Poonpon
English Department,
Faculty of Humanities and Social Sciences,
Khon Kaen University, Khon Kaen, Thailand
Email: korpul@kku.ac.th

Abstract

The COVID-19 pandemic caused challenges in language testing for many universities and institutions. The present study aimed to address some challenges and restrictions by designing and implementing an online testing system for a high-stakes proficiency test at a Thai university during the COVID-19 pandemic and exploring test takers' perceptions of the design and its implementation. The online language testing system, i.e., test interface, test delivery mode, and proctoring system, were designed and implemented. The online test interface was developed using Fulcher's (2003) computer-based language test development framework. The design and implementation of the synchronous at-home test delivery mode and the remote proctoring system were based on the university's test context and technological affordances. A total of 218 test takers completed a questionnaire after taking the remote online proficiency test. The test takers' perception of the test design and implementation was analyzed quantitatively and

qualitatively, using descriptive statistics and content analysis. The results revealed the strengths and weaknesses of the online test interface, synchronous at-home test delivery mode, and remote proctoring system. The study findings suggest that this online language testing system is applicable to universities in similar testing contexts.

Keywords: test takers' perceptions, proficiency test, online testing, Thai university, COVID-19

Introduction

The COVID-19 pandemic has resulted in rapid changes in language education worldwide. English language teaching and learning at all levels have been shifted from face-to-face to online classes. Teachers and students have had to adjust to both synchronous and asynchronous technology-based instruction. The COVID-19 pandemic has created challenges for language testing and assessment.

English language assessments in higher education settings have been particularly challenged by such sudden changes. Indeed, due to the shift from face-to-face to remote testing, there have been various emerging needs in language testing and assessment. For example, since the beginning of the pandemic in 2020, many universities and institutions have explored several issues related to the design and application of technology-based tests, with particular reference to technology familiarity, practicality, and innovation (e.g., Alghammas, 2020; Clark et al., 2020; Green & Lung, 2021; Muhammad & Ockey, 2021; Ockey, 2021; Ockey et al., 2021; Purpura, Davoodifard, & Voss, 2021; Wagner & Krylova, 2021).

As in many countries, the COVID-19 pandemic has affected language education in Thailand. After the first case of COVID-19 infection was reported in January 2020, the Thai government announced precautionary measures to control the situation and declared a state of emergency after finding a surge of infections in March 2020 (Rajatanavin et al., 2021). Lockdown measures in the first wave were implemented until mid-May 2020. The second and third

waves of the pandemic were from December 2020 to February 2021 (Rajatanavin et al., 2021) and from April 2021 to October 2021 (Reuters, 2021), respectively. Throughout this crisis, schools and universities were closed, and this brought significant disruption to teaching and testing across the country.

Many universities in Thailand have attempted to seek ways to manage language testing and assessment throughout these lockdowns. Language instructors actively employed online classroom-based assessments to evaluate their students' learning and determine their achievement at the end of the semester. However, it has been very challenging for some colleges and universities to administer high-stakes language tests during the COVID-19 pandemic, especially when this had to be done remotely online rather than face-to-face as usual. However, thanks to advances in technology, designing and implementing high-stakes language tests online for test takers has become feasible (Chalhoub-Deville, 2001; Chapelle & Douglas, 2006; Chapelle & Voss, 2016; Clark et al., 2020; Dooey, 2008; Green & Lung, 2021; Noijons, 2013; Ockey, 2021).

A case in point, at Khon Kaen University, a locally developed language proficiency test known as the Khon Kaen University Academic Language Test (KKU-AELT) is offered and used as a requirement for graduate student and academic staff admissions. When there was a need for graduate student recruitment during the countrywide lockdown at the beginning of 2020, a debate emerged on how the KKU-AELT could be converted into an online test and delivered to test takers across Thailand and abroad. To date, only a small number of studies in Thailand and elsewhere have explored how high-stakes proficiency tests have been prepared, implemented, and monitored at universities during pandemic lockdowns. During these challenging times, test centers must take care to ensure that any changes in language testing mode do not decrease test takers' motivation or affect their performance (Zhou & Yoshitomi, 2019). As test takers are the most important stakeholders (Bachman & Palmer, 2010; Gu & So, 2015; Rea-Dickens, 1997), this paper examined test takers' perceptions of the design and implementation of KKU-AELT during the COVID-19

pandemic. It hoped to shed light on the implementation of innovations in language testing and assessment in Thailand initiated in response to pandemic lockdowns which made necessary a move away from face-to-face to virtual modes of delivery.

Literature Review

    Computer-Assisted Language Testing (CALT)

In language assessment, technology is used for testing purposes, e.g., creation and modification of test tasks, test delivery, item selection, feedback provision, and scoring (Chapelle & Douglas, 2006). Computer-assisted language testing (CALT) is an integrated procedure in which language performance is elicited and assessed with the help of a computer (Noijons, 2013). This procedure integrates the use of a computer for generating tests automatically, interacting with test takers in the form of online interaction, and evaluating test takers' responses.

According to Chapelle & Douglas (2006), there are two main types of CALT: computer-based tests (CBT) and computer-adaptive tests (CAT). In CBT, data are transferred from paper-based tests to computer-based tests. Test structure, components, and content are identical; the delivery platform is different. This is considered a linear approach as the test takers take the same set of test items or tasks. On the other hand, CAT relies on the item response theory (IRT) to control adaptivity based on test takers' performance on each item on the test. The test taker is presented with an item or task, dependent on his or her responses to the preceding item or task. That is, the test is "adapted" to the test takers' level of ability.

CALT has advantages and disadvantages (Chapelle & Douglas, 2006; Chapelle & Voss, 2016, Noijons, 2013). Advantages of CALT include that it can help overcome administrative and logistic burdens, offer test consistency and uniformity, assure enhanced authenticity and more significant interaction, offer insights into test takers' routes and strategies, and guarantee immediate test results and feedback. However, CALT requires suitable or compatible hardware or software and test takers with computer skills and digital literacy. For those with

limited digital skills, technical preparation may take longer. During the test, technical problems can also distract or demotivate test takers. Moreover, these factors can lead to test anxiety.

Test Interface Design for Online Language Tests

Designing a test interface is one of the most important considerations for developing CBT and CALT. Fulcher (2003) described considerations to be addressed by test developers designing an interface: planning and initial design, usability testing, and field trial and fine-tuning (Figure 1).

Phase 1 Planning and Initial Design	Phase 2 Usability Testing	Phrase 3 Field Testing & Finetuning
1. Designing prototypes • Hardware considerations • Software considerations 2. Good interface design • Navigation • Page layout • Text & text color • Toolbars & controls • Icons & graphics • Help facilities • Item types • Multimedia • Forms for writing/short-answer tasks	• Searching for problems and solutions • Selecting test takers from usability studies	• Verifying that the interface works across sites and platforms • Seeking interface-related resources of variance
Concurrent activities: • Delivery systems • Score retrieval & storage • Distribution sections score by human raters	**Concurrent activities:** • Item writing and banking • Pre-testing • Trialing score rubrics • Constructing structural construct studies	**Concurrent activities:** • Developing tutorials • Producing practice/example tests • Developing rater training packages and conducting rater training • Considering scaling studies and score reporting • Planning further validation studies

Figure 1. Computer-based language test design considerations (adapted from Fulcher, 2003)

Phase 1 involves designing prototypes which relates to hardware and software considerations (e.g., suitable computer specs or testing software), and interface design which involves navigation, page layout, text and text color, toolbars and controls, icons and graphics, help facilities, item types, multimedia, and forms for writing or short-answer tasks. This first phase also includes concurrent activities, including delivery systems, score retrieval, and storage, and distribution sections scored by human raters. Phase 2 describes tasks after the prototypes and interface are designed. It aims at searching for usability-related problems and solutions that the designed interface may cause. The concurrent activities for this phase include item writing and banking, pre-testing, trialing score rubrics, and constructing structural construct studies. After piloting an interface in Phase 2, it is necessary to verify that the interface works across sites and platforms in Phase 3. This may involve seeking other interface-related resources. The concurrent activities that facilitate the final phase are, for example, developing tutorials, producing practice or example tests, developing rater training packages, and conducting rater training, as well as considering score reporting. The test interface should be designed carefully to ensure that it does not interfere with assessment because this can yield contaminated scores that may threaten valid score interpretation (Fulcher, 2003).

Online Language Testing Delivery Modes

The impact of the COVID-19 pandemic has put in focus how test centers worldwide deliver their language tests. Ockey (2021) invited famous test centers in the US to share their adaptations in *Language Assessment Quarterly Journal*. Based on what these test centers reported, test delivery modes can be categorized into three types: synchronous at-home test delivery, synchronous outdoor face-to-face test delivery, and asynchronous at-home test delivery.

Asynchronous at-home test delivery is when test takers are in their homes while test administration happens on campus or in test centers. For instance, a placement test at Brigham Young University, Hawaii, was delivered to new undergraduate students as a virtual test (Green & Lung, 2021). The online delivery test still included four skills:

listening, reading, speaking, and writing. The test center used Canvas as a platform to deliver the test and Proctorio as a proctoring program. At Temple University, they also gave their Internal Teaching Assistant Test online (Wagner & Krylova, 2021). Taking advantage of the Covid-19 pandemic, the test developers amended the assessment process by adding interactional competence instead of prerecorded questions and using video-mediated technology to proctor test takers. IELTS is another example of the synchronous at-home test delivery mode (Clark et al., 2020). With IELTS Indicator, specially developed as online IELTS, the test introduced *Inspera* as an online assessment platform. The proctoring system includes human monitoring and controls over the browser. A video-call speaking test is used instead of a face-to-face interview. IELTS has also conducted research to back up their changed test conventions.

Synchronous outdoor face-to-face test delivery describes tests made available face-to-face without changes in testing procedures but with COVID-19 safety measures. For example, the Iowa State University English Language Placement Test of Oral Communication (EPT OC) used this delivery mode (Ockey et al., 2021). Before the pandemic, the test was given in face-to-face indoor environments at the test center. However, during the pandemic, the test center used a large university building and arranged outdoor assessment stations at the corners of the building. Test takers and test administrators were required to wear face shields and face masks. However, during the actual speaking test, test takers and interlocutors removed their masks when speaking.

Asynchronous at-home test delivery means tests are delivered anytime using high-quality proctoring software. TOEFL iBT Home Edition and the Placement Test for the Community English Language Program (CLP) at Teachers College, Columbia University, are examples of this type (Muhammad & Ockey, 2021). These two tests use both human proctors and Artificial Intelligence (AI) to help monitor and detect the test takers' suspicious behaviors. While TOEFL iBT uses ProctorU, which uses both human proctors and AI, the placement test for CLP uses Canvas as an LMS and Quartics test delivery software

together with Honorlock, the AI remote proctoring software (Purpura, Davoodifard, & Voss, 2021).

Test Takers' Perceptions of Online Language Testing

Despite the benefits of moving face-to-face testing to online modes, some challenges in terms of accessibility, fairness, security, and cheating seem to be unavoidable. Previous studies suggested that test developers and administrators should be aware of these challenges as they may cause test anxiety and demotivation, influence test takers' performance, and eventually affect test validity and reliability (e.g., Bachman & Palmer, 2010; Gu & So, 2015; Zhou & Yoshitomi, 2019). Because test takers are the most important stakeholders who directly experience these adverse effects (Bachman & Palmer, 2010; Rea-Dickens, 1997), several studies focused on exploring test takers' perceptions of and feedback on changes from face-to-face to online or computer-based tests (e.g., Brunfaut, Harding, Batty, 2018; Gu & So, 2015; Zhou & Yoshitomi, 2019). These studies have reported the challenges and solutions faced by the test takers, and also revealed issues that test developers or test organizations should consider when making decisions about moving language tests online.

A review of relevant literature has guided the present study's test interface design and selection of an appropriate test delivery and proctoring systems for the Thai university setting and shaped the preparation and implementation of the design of the online language testing system. The study explored how an online language testing system was designed and implemented at a Thai university and how the test takers gave feedback on the design and implementation. It was expected that the study would contribute to the growing research on online language testing during the pandemic in Thailand.

Context and Methodology

Khon Kaen University Academic English Language Test (KKU-AELT)

KKU-AELT is a language proficiency test developed by the Center for English Language Excellence, established in 2014, within the Faculty of Humanities and Social Sciences at Khon Kaen University,

Thailand. The test aims to measure test takers' academic English proficiency, focusing on academic reading and writing. The Common European Framework of Reference for Languages (CEFR) B level group (Independent User) (Council of Europe, 2001) was used as a framework when developing this test's specifications. The test lasts three hours: two hours for reading and one for writing. The reading test includes six passages with 69 multiple-choice and gap-filling test items. The writing test has one prompt (i.e., an agree or disagree question), asking the examinees to express their opinion on a given topic for at least 250 words. The prompt example is, *Do you agree or disagree with the following statement? "University should be open to all ages of learners." Use specific details and examples to support your answer.* The scoring criteria are adapted from the TOEFL independent writing scoring rubrics related to topic development, organization, and language use.

The KKU-AELT test score for each skill is one hundred. (The scores are still in the process of mapping to the CEFR.) The test scores for both skills are reported separately on a score report form, downloaded from the CELEx website two weeks after the test date. The scores are used to recruit graduate students and university staff. Before COVID-19, the test was administered monthly at the Faculty of Humanities and Social Sciences test center. However, after the COVID-19 first attacked Thailand at the beginning of 2020, face-to-face testing at the university was suspended.

Due to the need to recruit both Thai and international graduate students, the Graduate School requested that the KKU-AELT be made available virtually as an online test that student applicants could take it from anywhere in the world. Therefore, since June 2020, the CELEx has been administering and fine-tuning the online version of the test. It was first offered online in June 2020, with 449 test takers, using a specially designed KKU-AELT test platform and a Zoom and human proctoring system. Unfortunately, this first attempt was unsuccessful since the webserver was unable to cope with the required volume of data. In the second attempt, Google Forms were used as the online test platform in conjunction with a Zoom-and-human proctoring system. This model was used in January and February 2020. It was found that

the Google Forms had certain limitations even though they were quite flexible for receiving test responses from a large number of test takers at a given time. The test interface and display of the Google Forms were not appropriate for the KKU-AELT reading part, and test takers' responses could not be saved automatically if their Internet connection was broken, thus leading to test takers' anxiety. More importantly, test takers were able to visit other websites to search for the meaning of unknown vocabulary and use automatic correction while working on their writing test. From March 2020, the local pandemic situation improved, and the KKU-AELTS exams were administered onsite again until February 2021. The third model, which has been used in the present study, was designed and used from March 2021 until September 2021.

Participants

Test takers of the KKU-AELT exam administered in August 2021 were asked to participate in the study. The August exam was selected as the focus of this study because this administration of the online test version achieved the best internal consistency coefficient (Cronbach's alpha at .87). Among 305 test takers, 218 participants (71.7%) responded to a questionnaire about their perceptions towards the design and implementation of the online language testing system. About two-thirds (72.9%) of the respondents were female, and one-third (27.1%) were male.

Materials and Instruments

A consent form

The consent form asked test takers to accept conditions of the online test (i.e., their responsibilities to prepare two devices and a strong Internet connection for the online test) and to give permission for use of their test scores in research.

A set of e-manuals for test takers

These were provided to the test takers after test registration confirmation and included a manual for using the online KKU-AELT testing system (KKU Exam), a Zoom user guide, an instructional video for the online KKU-AELT testing system, the Zoom program, and instructions and guidelines for online KKU-AELT test takers. Test

takers were required to study these manuals before a test orientation, held three days before their test date.

Questionnaire

This questionnaire was used to gather feedback from test takers on the design and implementation of the online language testing system. The questionnaire was divided into three parts: test takers' demographic information, experience of online learning and testing, and perceptions of online testing system. The first two parts were designed as a five-point Likert scale and the final part as an open-ended response.

Design of Online Language Testing System in the present study

Since the KKU-AELT test is used for university admission, test takers, both Thai and international, are required to be alone during its administration. Therefore, the challenge was administering the KKU-AELT exam without risks to test validity, reliability, and practicality. The design of the online testing system can be divided into three aspects: a test interface, a test delivery mode, and proctoring system. Decisions on the design were made based on the university's technological affordances.

First of all, the KKU-AELT paper-based test had to be converted into a web-based test. It was then necessary to consider an appropriate test interface. This study transferred the KKU-AELT paper-based test into an online test using the KKU exam, and the Moodle Language Management System (LMS), as the university already has this test platform. Throughout this process, the test construct and structure remained unchanged. That is, similar to the paper-based test, the online version of the test included academic reading, multiple-choice test items, and academic essay writing. To ensure test security, the KKU exam shuffled questions and choices as well as set automatic test starting and ending times. After each test taker completed the test, the system automatically scored all multiple-choice items in the reading part. The writing responses from the writing test were downloaded and scored by at least two human raters. If the two writing scores did not

match, a third rater was assigned and the two most closely aligned scores were used. Inter-rater reliability was considered high (r = .91).

The synchronous at-home test delivery mode was used for the online KKU-AELT. Test takers were at their homes while the test was administered from the test center on campus. The KKU-AELT delivery system made use of the KKU Exam as its online test platform to deliver the test to the test takers. On the test day, test takers logged in to the KKU Exam platform with their assigned username and password before taking the test.

The proctoring system of the KKU-AELT online included the use of Safe Exam Browser (SEB) and human monitoring via Zoom. During the exam, SEB was used to control test takers' access to other resources, e.g., system functions, websites, and applications. In other words, test takers could not visit other websites or screens while taking the test. Test takers were required to install the SEB beforehand as the software needed to be activated when logging in to the KKU Exam platform. To monitor suspicious behaviors, human invigilators remotely proctored using a Zoom-mediated channel. Test takers were required to use two devices, i.e., a desktop or notebook and a smartphone or iPad, at their testing station. The first one was aimed to function as a test computer, and the other was to monitor the test taker's behavior during the exam.

Figure 2 summarizes the model of the online KKU-AELT testing system used at the university. Test takers took the online KKU-AELT from their homes while Zoom-mediated human proctors worked from the test center on campus. The KKU Exam platform and SEB were used together to create a test delivery platform with a secured computer screen.
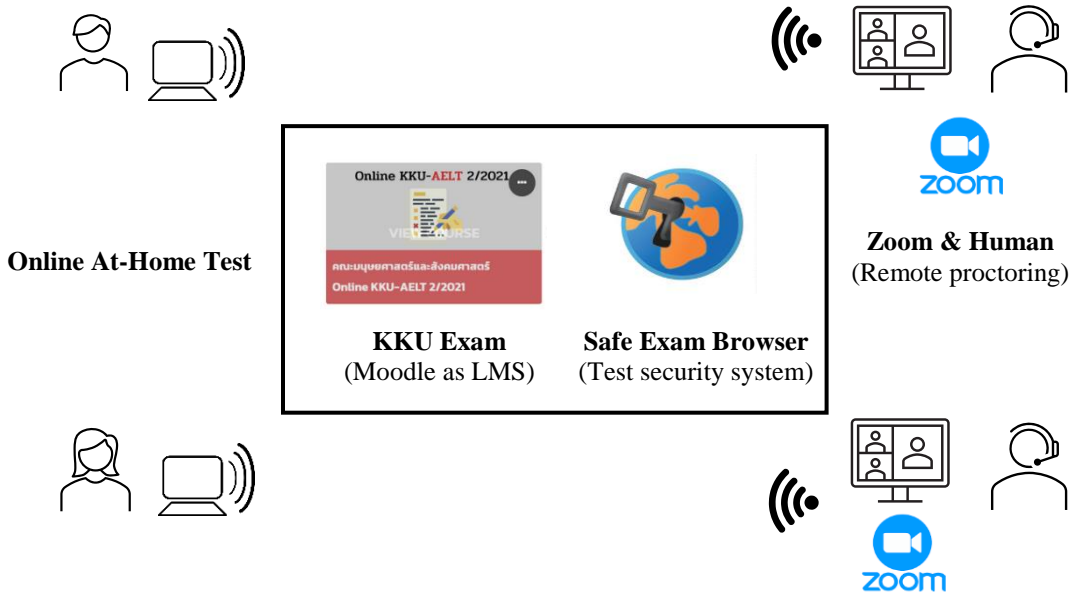
Figure 2. A model of the online KKU-AELT testing system

The Implementation of the Online Language Testing System in the present study

Figure 3 depicts the KKU-AELT online test procedures. It started with test takers' online registration two weeks before their test date. After that, the registered test takers received an **email** to confirm their successful registration and were sent a consent form, a set of manuals with an instructional video, and a link to an orientation. They were also asked to read, sign, and email the consent form back to the test center before the test date. The test takers were also required to study the manuals before a test orientation day. At the virtual orientation, the test takers were introduced to the online KKU-AELT exam, i.e., test structure, test score report, and how to access and start the online test. A mock online KKU-AELT was provided for the test takers to familiarize themselves with the test and the online testing system and to help them with any technological problems. On the test day, the test takers were at their homes using their computers to access the online testing system and take the test. They were required to be in a quiet place and were not allowed to have anybody around them. Stationery and other electronic devices were not allowed during the test. The test takers were

divided into small groups and put in different Zoom rooms, each of which included about 20 test takers and at least one proctor. Each test taker was required to enter the test room with two devices: a computer or notebook as a test station with a front camera and a smartphone or an iPad as a monitoring camera (Figure 4). Each proctor began with a test takers' identity check, followed by helping them adjust their monitoring camera to ensure the correct position before formal commencement of the test. During the test time, each proctor monitored any suspicious behaviors and provided help to test takers when needed. To confirm cheating behaviors, at least two proctors and a head proctor had to concur. After taking the test, the test takers could check and print out their online test scores from the KKU-AELT website.

All proctors attended a two-hour training session before the test date. The proctor training aimed to provide an overview of the online KKU-AELT, train proctors in how to use the KKU-Exam platform, how to set an appropriate test station for test takers, how to use Zoom to control test takers behaviors, how to detect suspicious behaviors, and how to facilitate the test takers during the exam.
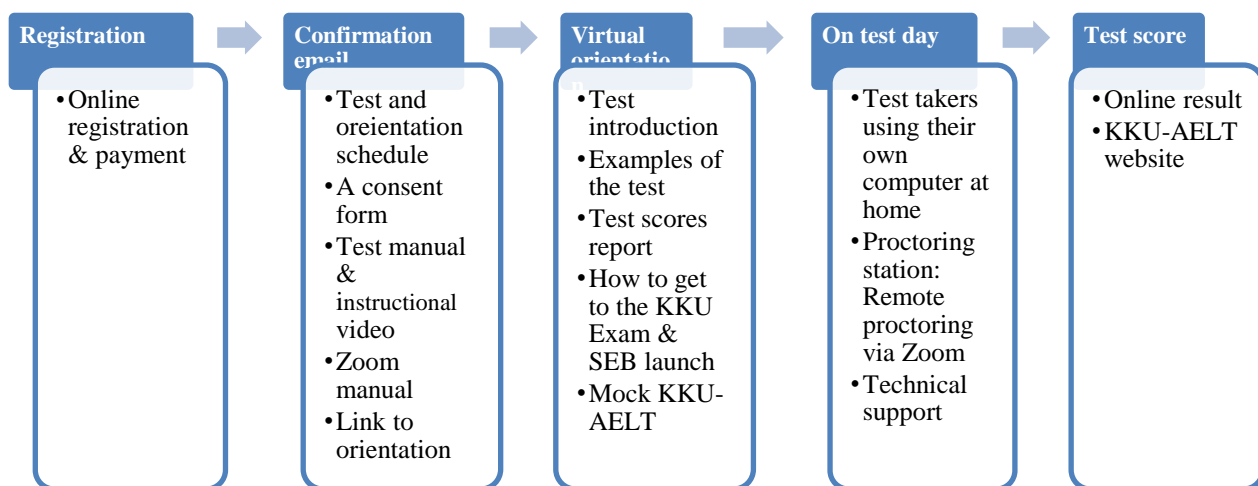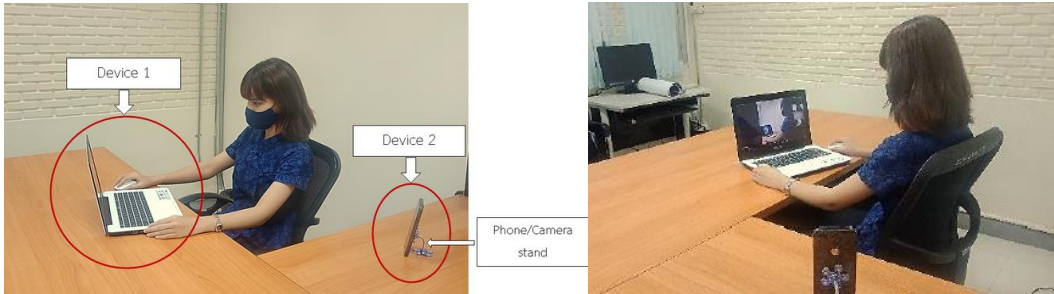
Registration	Confirmation email	Virtual orientation	On test day	Test score
• Online registration & payment	• Test and oreientation schedule • A consent form • Test manual & instructional video • Zoom manual • Link to orientation	• Test introduction • Examples of the test • Test scores report • How to get to the KKU Exam & SEB launch • Mock KKU-AELT	• Test takers using their own computer at home • Proctoring station: Remote proctoring via Zoom • Technical support	• Online result • KKU-AELT website

Figure 3. KKU-AELT online testing procedures

Test taker's test devices and positions

Proctor's monitoring screen   Proctoring via Zoom

Figure 4. KKU-AELT Proctoring

Data Collection

The online testing system in this study has been in use since March 2021. However, the test developers trialed, evaluated, and fine-tuned the system in the first few months of its implementation. Feedback from test takers of the August 2021 exam was collected via an online questionnaire as this exam showed the highest coefficient of internal consistency (Cronbach's alpha at .87). Before the test day, the test takers signed a consent form to allow the test center to use their scores and information and volunteer to answer the questionnaire after they completed the test. They were informed that their responses to the questionnaire did not affect their test scores.

Data Analysis

The quantitative data from the questionnaire were analyzed using descriptive statistics of mean, standard deviation, and percentage. The criteria for the mean interpretation were as follows.

Strongly Agree	4.21 – 5.00
Agree	3.41 – 4.20
Neutral	2.61 – 3.40
Disagree	1.81 – 2.60
Strongly Disagree	1.00 – 1.80

Summative content analysis was also used to analyze the qualitative data from open-ended questions. Two coders read all the data and did systematic coding of emerging themes. Then, the two coders had an online meeting and discussed their coding of the data, reaching 90% consensus on their coding.

Results

Test Takers' Background Information

The questionnaire data revealed test takers' background information including their ages, gender, test purposes, digital skills, and experiences of online learning, online testing, and the online KKU-AELT.

Out of 218 respondents, more than half (54.2%) were 21 to 30 years old. Half of them had taken the paper-based version of the KKU-AELT previously. More than half of them, 55.9 percent, had taken the online KKU-AELT before. Moreover, about 62 percent had experienced online learning, and 60.2 percent had experienced online testing before.

Test Takers' Perceptions towards the Design of the Online Language Testing System

The results from the questionnaire showed test takers' perceptions towards the design and implementation of the online

testing system as it pertained to tools, the Internet, and the online testing interface (Table 1).

Table 1. Test takers' agreement level regarding the online test interface design (N = 218)

Statement	Mean	SD	Interpretation
Tools and the Internet			
• Your equipment was ready for the test.	4.43	0.75	Strongly Agree
• Your Internet connection remained strong throughout the test.	4.42	0.75	Strongly Agree
• You were ready for the online test.	4.17	0.92	Agree
Online testing interface			
• Seeing time left on screen made me progress better.	4.62	0.56	Strongly Agree
• You could easily follow the online test direction without any problem.	4.52	0.76	Strongly Agree
• Logging-in and logging-out stages were clear and appropriate.	4.51	0.75	Strongly Agree
• The exam interface was appropriate.	4.46	0.79	Strongly Agree
• KKU Exam screen layout was appropriate.	4.46	0.79	Strongly Agree
• The online test platform was convenient.	4.44	0.72	Strongly Agree
• Using two devices (i.e., a desktop and a smartphone/iPad) was appropriate.	4.45	0.68	Strongly Agree
• Safe Exam Browser was suitable for the test.	4.43	0.83	Strongly Agree
• Using Zoom was appropriate for the online exam.	4.38	0.65	Strongly Agree

In terms of tools and the Internet needed for the online testing, the respondents strongly agreed that their computer and smartphone/iPad were ready for the online test ($\bar{x}$ = 4.43, SD = 0.75), and their Internet connection was good throughout the test ($\bar{x}$ = 4.42, SD = 0.75). They also agreed that they were ready for the online test ($\bar{x}$ = 4.17, SD = 0.92).

The KKU Exam system was considered appropriate as an online testing interface. The most strongly accepted procedures were the online test timer and test direction. While taking the exam, the respondents firmly believed that the KKU Exam timer helped them progress better ($\bar{x}$ = 4.62, SD = 0.56), and its instructions were clear

enough for them to follow without any problems ($\bar{x}$ = 4.52, SD = 0.76). The respondents also strongly agreed that the logging-in and logging-out stages were clear ($\bar{x}$ = 4.51, SD = 0.75) and that the KKU Exam interface and screen layout were appropriate ($\bar{x}$ = 4.46, SD = 0.79). The Moodle platform was also considered convenient ($\bar{x}$ = 4.44, SD = 0.72). In addition, other devices and programs used with the KKU Exam were considered appropriate. The respondents also strongly agreed that it was appropriate to use two monitoring devices during the online test ($\bar{x}$ = 4.45, SD = 0.68) together with the Safe Exam Browser ($\bar{x}$ = 4.43, SD = 0.83) and the Zoom program ($\bar{x}$ = 4.38, SD = 0.65) to facilitate proctoring.

Test Takers' Perceptions towards the Implementation of the Online Language Testing System

In addition to the respondents' views about the preparation and readiness of their test devices and the online testing interface, they also revealed their perceptions of the online testing procedures. The results showed that the respondents strongly agreed that all online testing procedures were reasonable. Before the test day, they greatly appreciated that the test orientation, which introduced them to the test structure, guided them on how to answer reading questions and respond to the writing prompt, as well as walking them through the online KKU Exam system, helping them prepare for the online test ($\bar{x}$ = 4.61, SD = 0.61). The online test manuals were considered very useful ($\bar{x}$ = 4.55, SD = 0.59). The respondents showed their strongest agreement that the proctor could handle problems during the test very well ($\bar{x}$ = 4.67, SD = 0.52). The respondents also appreciated the proctor's technical assistance before the test started ($\bar{x}$ = 4.62, SD = 0.61). However, they evaluated their ability to finish the allocated time with the least confidence ($\bar{x}$ = 4.19, SD = 0.98).

Table 2. Test takers' agreement level for the online testing implementation (N = 218)

Implementation stage	Mean	SD	Interpretation
Before the test day			
• Registration was convenient.	4.40	0.72	Strongly Agree
• Name list confirmation via email was appropriate.	4.37	0.64	Strongly Agree
• The test orientation was helpful.	4.61	0.61	Strongly Agree
• Online test manuals were useful.	4.55	0.59	Strongly Agree
• Communication channels (i.e., email, Facebook Page, website, phone) were convenient.	4.45	0.55	Strongly Agree
On the test day			
• Before the test started, the proctor's technical assistance was reasonable.	4.62	0.61	Strongly Agree
• The verification of test takers' identities was convenient.	4.56	0.65	Strongly Agree
• During the test, the proctor could handle problems very well.	4.67	0.52	Strongly Agree
• Time remaining reminder by the proctor was helpful.	4.37	0.69	Strongly Agree
• You could finish the test within the allocated time.	4.19	0.98	Agree

The qualitative data from the open-ended questionnaire revealed both positive and negative perceptions of the online testing design and implementation. Most respondents believed that the online KKU-AELT exam was excellent and convenient during COVID-19. They agreed that the online KKU-AELT interface and online testing system were helpful and straightforward. However, some technical glitches and problems with screen layout and the online testing environment were experienced while taking the online test. Some respondents had technology-related issues, e.g., incompatible computer specs or an outdated Windows program, so they could not activate the SEB and the KKU Exam system. A few of them did not realize that they had to install the SEB before the test day. KKU-AELT technical staff had to help these respondents to ensure they could take the test which wasted the test takers' exam time. These respondents were found to be those who did not attend the test orientation or read the test manuals.

Moreover, an unstable Internet signal interrupted the online test for some respondents.

Many respondents complained about the screen layout for long reading passages. As the test interface was designed to put a reading passage before reading questions, test takers had to scroll down to read the whole passage and later answer the following questions. Though it was fine with short reading passages, the respondents found it challenging to read very long passages and answer questions as they had to scroll up and down. This caused some of them to complain about sore eyes and anxiety while taking the test.

Some respondents also experienced an undesirable testing environment. As there were 20 test takers in each Zoom test room, every test taker was asked to mute their microphone during the test and communicate with the proctor by typing in a chat screen if they had any questions or problems. However, some test takers ignored these directions and switched on the microphone to communicate with the proctor. This disturbed other test takers in the same Zoom room. Suggestions from the respondents included creating a separate communication channel for those who needed help during the exam.

Discussion

This study aimed to design and implement an online testing system for a high-stakes proficiency test at a Thai university during the COVID-19 pandemic. The online testing system, i.e., the test interface, the test delivery mode, and the proctoring system, were designed and implemented. The test takers' perceptions of the design and implementation of the online test were investigated. The results from the investigation can be discussed in several aspects as follows.

The online testing interface and programs used in the present study were those made available by the university. KKU Exam or Moodle as a test management system appeared appropriate and user-

friendly to the test takers who were satisfied with the range of functions in KKU Exam, e.g., the page layout, page navigation or instructions, forms for the writing part, and multiple choice item layout. This suggested that the online test interface designed for KKU-AELT followed Fulcher's (2003) considerations was somewhat effective as it did not elicit negative perceptions among test takers. This could also point to test validity as the interface did not interfere with the assessment's provision of valid test scores (Fulcher, 2003).

The synchronous at-home delivery mode and the proctoring system used in the present study were also found to be generally suitable for the university's context during the unexpected COVID-19 pandemic. Similar to the online language tests offered by Brigham Young University in Hawaii (Green & Lung, 2021) and Temple University (Wagner & Krylova, 2021), the systems were designed based on the university's test contexts. The whole monitoring procedure could be controlled using a combination of various technologies. Technological tools such as Moodle, SEB, and Zoom were used together to facilitate the high-stakes test delivery and proctoring system. Moodle was used as an online test platform with a user-friendly interface. SEB was attached to Moodle as a test security program that could control other websites and applications and prevent any resources from being used during an exam. Zoom was also used as video-based proctoring platform to manage the synchronous sessions and increase test security. Apart from the technology, human proctors were able to help the test takers with any difficulties and catch test takers' suspicious behaviors during the exam. Test takers felt the proctors helped facilitate their examination and somewhat helped them reduce their technological-related anxiety. It seemed therefore the proctors were generally successful in preforming their duties remotely which is in line with previous research (e.g., Clark et al., 2020; Purpura, Davoodifard & Voss, 2021) in that virtual proctoring by video-mediated technology

and human proctoring can increase test security and control synchronous sessions in a context with limited technological affordances (Güngö & Güngö, 2021).

The implementation procedures used in the present study were found to be practical for the university's situation during the COVID-19 pandemic and supported test takers. All technical support provided for the test takers (as suggested in Fulcher's phrase 3)—i.e., test takers' manuals, test orientation, instructional video, technical support, and proctor's support during the exam—were found to be very useful (Fulcher, 2003; Güngö & Güngö, 2021; Purpura, Davoodifard & Voss, 2021). However, despite the well-prepared design and implementation of the online KKU-AELT exam, some problems were still found. The most common problems were related to computer program capability and Internet connection. As the test takers had to use their own devices and Internet connections for the online examination, this factor was somewhat uncontrollable. However, the technical support provided for the test takers in the test orientation and during the exam helped them cope with such difficulties. The present study confirms previous relevant research (e.g., Fulcher, 2003; Green & Lung, 2021; Wagner & Krylova, 2021) that a necessity for an online testing system is an appropriate computer device provided with a smooth Internet connection.

Other problems identified were related to test takers' digital literacy, experience of online testing, and their preparation for the test. As the study showed, many test takers had limited or no experience or knowledge of taking online exams and using technology in daily life. This caused them some anxiety and pressure. This is why test takers were offered a KKU-AELT orientation and test takers' manuals and other support. It can be said that the support provided was quite useful for the test takers and helped reduce many of the problems and concerns.

The design and implementation of online exam system raises threats to quality (Muhammad & Ockey, 2021). However, the online testing system designed for KKU-AELT could generally ensure test validity as the test content was transferred from paper-based to computer-based mode successfully. The multiple-choice format could guarantee reliability, authenticity, and practicality for the online language test (Purpura, Davoodifard, & Voss, 2021).

From the design and implementation procedures used in the present study, it is possible to recommend six essential stages for the online testing system development: 1) online test platform preparation, 2) online exam preparation, 3) online proctoring system planning, 4) test takers' familiarization with the online test system, 5) test administration, and 6) test system evaluation. In the first stage, a test center needs to check and prepare a test platform, software programs, and a proctoring mode that is applicable and appropriate for the test purposes. Second, an online exam version should be created on the chosen platform. This can be either a linear test or a computer-adaptive test. Third, an online proctoring system should be designed. This can depend on various technological factors and educational contexts. After that, the test takers should be informed and trained in how to access and use the online test system. This is to familiarize the test takers with the online test platform and proctoring system. At this stage, a test taker orientation is recommended to reduce problems related to the online test system on the test day. Finally, the online testing system should be evaluated by test takers, proctors, and other stakeholders. The feedback from these groups can be used to improve the design and implementation of the online testing system (Fulcher, 2003; Noijons, 2003).

The present study had some limitations. First, the study used only a questionnaire to examine the test takers' perceptions. Future

research should consider an interview as an additional research tool to gain a more detailed insight into online testing. Second, the online KKU-AELT interface, which was based on the technological resources available at the university, may not serve as well as designed (e.g., the reading screen layout in the online KKU-AELT test). Therefore, a more user-friendly test platform may need to be found or designed and developed. In addition to the test interface, the online proctoring system used in the present study may not be able to fully eliminate cheating. Weaker test takers may be tempted to cheat as their suspicious behaviors may not be caught by human proctors (Suryadi Rahmawati, 2017). If possible, in the future, more advanced technology, e.g., AI-based proctoring (Nigam et al., 2021; Purpura, Davoodifard & Voss, 2021), that can automatically identify suspicious behaviors can be used to ensure a more effective way to prevent cheating and secure the test environment. Finally, since language proficiency test formats may vary at different universities or test centers, caution should be taken when interpreting and applying the findings of the study. Future research could also focus on ongoing improvements in the design of online test interfaces and systems and their implementation in high-stakes language tests as well as classroom-based tests. It is expected that a specifically designed and developed testing interface and platform will help in conducting controlled and organized exams, thus increasing the validity and practicality of conducting high-stakes language proficiency exams online.

Conclusion

The present study demonstrates how the design and implementation of a remote high-stakes language proficiency test was possible at a Thai university using available infrastructure during the COVID-19 pandemic. In particular, the design of the web-based test

interface was designed and implemented based on Fulcher's three phases of computer-based test interface design. The synchronous at-home test delivery mode and Zoom-and-human-based proctoring system were employed in the present study. The test takers' perceptions towards the design and implementation of the online KKU-AELT exam revealed that the KKU Exam system (Moodle), in conjunction with the Safe Exam Browser, was a suitable system to help deliver an online test, where test takers were in different parts of the world. The use of technology and human proctors for invigilation was an effective combination for monitoring test takers' behavior during the exam. In addition, since the test takers had different levels of digital literacy and varied access to technology-related resources, various kinds of support (i.e., online test manuals, communication channels, an orientation, technical assistances) were necessary to prepare the test takers for the online test. The online testing system design and its implementation enabled the high-stakes test administration and facilitated the graduate-student recruitment process at the university. It was also doable within budget constraints and convenient for the test center and test takers. The processes described in this study may be useful at other universities or institutions in similar contexts for adapting the tests and exams for online administration.

About the Author

Kornwipa Poonpon is the head of the English Department and the director of the Center for English Language Excellence, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand. Her research interests include language assessment, corpus linguistics, and English language teaching.

ORCid ID: https://orcid.org/0000-0002-8349-2834

References

Alghammas, A. (2020). Online language assessment during the COVID-19 pandemic: University faculty members' perceptions and practices. *Asian EFL Journal, 27*(44), 169–195.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice.* Oxford University Press.

Brunfaut, T., Harding, L., Batty, A. O. (2018). Going online: The effect of mode of delivery on performances and perceptions on an English L2 writing test suite. *Assessing Writing, 36*, 3-18. https://doi.org/10.1016/j.asw.2018.02.003

Chalhoub-Deville, M. (2001). Language testing and technology: Past and Future. *Language Learning & Technology, 5*(2), 95–98.

Chapelle, C. A., & Douglas, D. (2006). *Assessing language through computer technology.* Cambridge University Press.

Chapelle, C. A., & Voss, K. (2016). 20 years of technology and language assessment in Language Learning & Technology. *Language Learning & Technology, 20*(2), 116–128. http://llt.msu.edu/issues/june2016/chapellevos.pdf

Council of Europe. (2001). *The Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge University Press.

Dooey, P. M. (2008). Language testing and technology: Problems of transition to a new era. https://doi.org/10.1017/S0958344008000311

Fulcher, G. (2003). Interface design in computer-based testing, *Language Testing, 20*(4), 384–408.

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

Gu, L., & So, Y. (2015). Voices from stakeholders: What makes and academic English test 'international'? *Journal of English for Academic Purposes*, *18*, 9–24. https://doi.org/10.1016/j.jeap.2014.10.002

Güngö, M. N., & Güngö, M. A. (2021). Reconceptualizing testing in times of COVID-19 from an activity theory lens: A narrative inquiry of a test designer's experiences. *The Asia-Pacific Education Researcher*, *30*(6), 531–539. https://doi.org/10.1007/s40299-021-00627-0

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A systematic review on AI-based proctoring system: Past, present and future. *Educ Inf Technol*, *26*, 6421–6445. https://doi.org/10.1007/s10639-021-10597-x

Noijons, J. (2013). Testing computer-assisted language testing: Towards a checklist for CALT. *CALICO Journal*, *12*(1), 37–58.

Ockey, G. J. (2021) An Overview of COVID-19's Impact on English Language University Admissions and Placement Tests. *Language Assessment Quarterly*, *18*(1), 1–5. https://doi.org/10.1080/15434303.2020.1866576

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

Purpura, J. E., Davoodifard, M., & Voss, E. (2021). Conversion to remote proctoring of an online language placement exam. *Language Assessment Quarterly*, *18*(1), 42–51. https://doi.org/10.1080/15434303.2020.1867145

Rajatanavin, N., Tuangratananon, T., Suphanchaimat, R., & Tangcharoensathien, V. (2021). Responding to the COVID-19

second wave in Thailand by diversifying and adapting lessons from the first wave. *BMJ Global Health*, *6*, 1–9. https://doi.org/10.1136/ bmjgh-2021-006178

Rea-Dickins, P. (1997). So, why do we need relationships with stakeholders in language testing? A view from the UK. *Language Testing*, 14(3), 304–314.

Reuters. (2021). UPDATE3-Thailand mulls more restrictions amid second wave of coronavirus. https://www.reuters.com/article/idUSL1N2JE018

Suryadi, B., & Rahmawati, Y. (2017). Integrated moral values in standard-based assessment: Opportunities and challenges of computer-based test in Indonesian national assessment. *Turkish Online Journal of Educational Technology*, 187–199.

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

Zhou, Y., & Yoshitomi, A. (2019). Test-taker perception of and test performance on computer-delivered speaking tests: The mediational role of test-taking motivation. *Language Testing in Asia, 9*(10), 1–19. https://doi.org/10.1186/s40468-019-0086-7