

# **The impact of test elements on students' performance in EFL assessment**

Joma Alrzini, Diane Rasmussen Pennington and Mark D Dunlop

*University of Strathclyde, Glasgow-UK*

## **ABSTRACT:**

Students applying to study at a university in a country that does not use their native language typically have to sit a formal language assessment exam. This exam is now most commonly done electronically. This can be challenging for students if they do not have previous experience with online tools. While previous studies have shown that practice of some skills exam, such as writing tasks, help improve students' performance, while other test skills has to be investigated. Still, there are few scientific and published studies on the usability and the effect of the design of test elements and their effect on test accuracy and effectiveness. The purpose of this study is to investigate factors of online English as a foreign language (EFL) tests and how they can adversely affect students' performance during the test. To meet this aim, we investigated the correlation between test structure and some test-taker characteristics. The study's significance is considered a guideline for test designers to improve the test design, which should lead to a more accurate assessment of the student's performance. This research was conducted using an online survey of new University entrance students. The key findings of this study are that less confidence during the test may result in an artificially reduced test score, with a potentially significant impact on a student's career or need to retake repeatedly. Moreover, other facts have been measured about an English online test such as first language, subject area and previous training on such examination. These contribute to challenges during the exam, which were raised in participants' responses on having difficulty interacting with the test.

## **Keywords:**

English language testing, test impact, students' performance, test elements.

## **1 INTRODUCTION**

With the increased use of digital technologies in the education sector, students are faced with challenges to adapt to diverse online learning systems. Such a system is utilized for learning the English language as a second language, and conducting the assessment online. Many factors influence the performance of students through an English online test such as their perception, motivation, knowledge, and desire as well as the delivery mode of the test (Budin 2014; Brunfaut et al 2018). Furthermore, some

studies revealed there is an impact of the test method on writing results, by looking at the average obtained result of writing tasks on a computer-based (Breland, Lee & Muraki 2005). This may be due to the familiarity, prior knowledge and usability of the test which assist test-takers to obtain better scores (Breland, Lee & Muraki 2005).

Zheng et al (2018) stated that the integration of technology in language learning has become an essential trend in many language institutes, due to the renewal development and expanded applications of technology in learning tools. Particularly in an online learning environment, students are engaging with a new dynamic environment, so their interaction with technology becomes an effective factor in learning success. Moreover, the rehearsal activities may help in familiarising with the test structure which in turn leads to confidence during the test.

**Objectives of the study:**

- To define the level of confidence of students during English online tests and the difficulty they have.
- To investigate the relationship between usability of the test and the most preferred features of the test.
- To identify other variables that can improve the test.

**Research Questions:**

- What challenges students may face during the test?
- To what extent the confidence feeling of a student helps them to perform better in the online test?
- What is the level of confidence of students during the test?

The rest of the paper is organised as follows: Section 2 discusses the methodology being utilized for this research, including a detailed literature review; Section 3 describes the obtained results; Sections 4 and 5 present the discussion and conclusion

## **2 METHODS**

This study is an initial part of the whole study investigating user experience with English online tests. A mixed-method is being utilized for the quantitative research approach. In this paper report, a scoping literature review and an online survey of non-native English speakers who are studying English courses or have done so recently have been recruited. This study is an initial part of the whole study investigating user experience with English online tests. A mixed-method is being utilized for the quantitative research approach.

### **2.1 Literature review**

We reviewed recently published studies following a scoping review format with searches on the following databases: ACM Digital Library, IEEE Digital Library, Google Scholar and Web of Science. In this literature, the used approach is identifying the relevant studies by the title of the paper and abstract. This approach is synthesizing

evidence from recent studies to identify the knowledge gap and clarify concepts of the research (Munn et al, 2018).

The review has been done with the following guide of scoping review methodology and Prisma diagram which is an extension of scoping review (Hamel et al, 2021). The strategy of this review is to use separate search queries for each database for the review period of 2010 to 2020 (inclusive). The use of search queries provided a variety of retrieved results, and these results are filtered based on relevance. In this theoretical method, the main focus is on the interaction between students and an English test done online, which means in the field of human-computer interaction. The impact of learning English online has been highlighted in this literature. Computer-assisted language learning is still in its development phase, recent studies presented the correlation between the practical usage of a digital English learning environment with the psychological aspects of it such as motivation, confidence and effectiveness (Lai, 2019). Studies indicated also that there is an improvement in students' performance in grammar and vocabulary knowledge with digital English learning (Cole & Vanderplank 2016; Lee 2019). At the same time, these studies revealed that EFL (English as a Foreign Language) learners interact more with virtual learning environments as a result of the frequent use of technology tools. However, these studies did not investigate the impact of the test and the approach of interacting with the online test.

Test format can impact the performance of students during the test, with the effect of gender bias as well. A comparison study was done for pre-intermediate level students on two types of tests multiple-choice (MC) and Constructed Response (CR) EFL vocabulary test. It indicated that students perform better in MC format than in CR test format, with variation between genders in the two formats, which means males scored relatively better in MC format than females (Masoumi & Sadeghi, 2020). Furthermore, language testing may be affected by a variety of factors, some from test format and methods and others from the test-takers perspective. The test itself has some formats such as open-end, gap-filling, ordering, multiple-choice...etc., while none of these formats of the test has an ideal test format, the test administrators have to employ the format based on the purpose and context of the test. Moreover, the familiarity with the test content can help reading comprehension test, these two influenced elements lead to influence scores of a test-taker (Rezaei, Barati & Youhanaee, 2012).

Test structure and test design are the most items of test usability, get used to testing is not an easy task but needs more practice. As Saiful et al (2020) pointed out that students need to be trained on the TOEFL test which is one of the official English language tests, particularly for the writing section. This is not only for acquiring knowledge but for time management as well (Saiful et al 2020). This study is only for the writing section, so other sections need to investigate. In addition to the test structure, student-centred design is an essential element of the test to offer interaction via technology. This means learners' comfortability, level of confidence and ability to achieve a better result from real interaction with an English learning environment (Wu et al, 2011). Furthermore, previous studies in this field confirmed that interaction tools build the confidence of students which leads to a direct correlation between confidence and motivation of students with their learning ability (Wu et al 2010).

In terms of usability design, more research needs to be conducted to investigate the different test designs and how they are usable for test-takers. This also includes the validity and effectiveness of the test design. The development of a web application for the IELTS reading test done by Rokhaniyah and Putra (2021) evaluated the performance among test-takers, this development carried out the feasibility and the effectiveness of the test as essential phases of the development process. The results of this study show an enhancement on obtained better scores and test-takers are motivated more.

To sum up, the scoping review for this study is summarised in the following Prisma diagram.

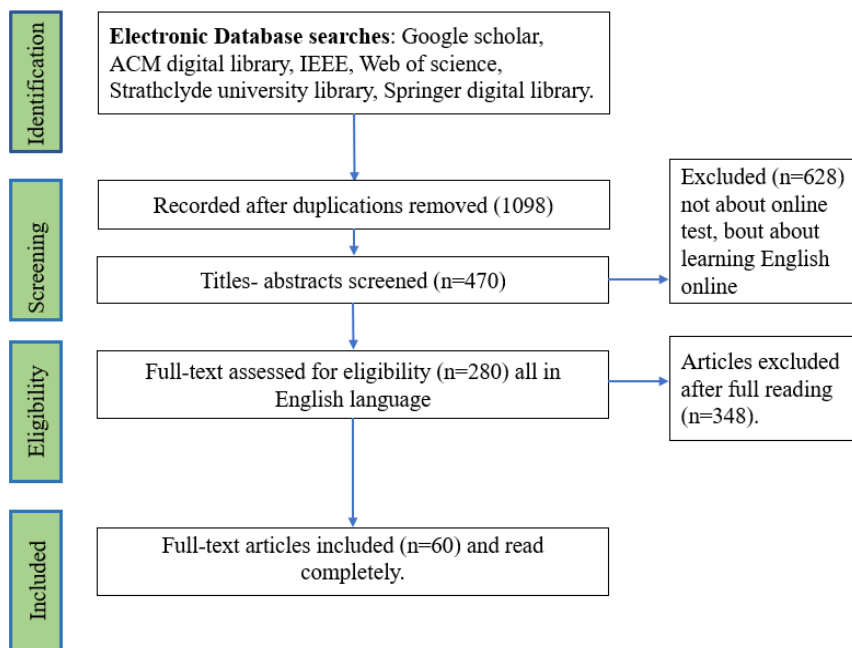


Figure1: Prisma Flow Diagram Of Reviewed Articles.

## 2.2 Online questionnaire

A survey questionnaire has been deployed to investigate the user experience with an English online test. A set of 21 questions were sent to language schools and first-year overseas students at universities in the UK for distribution through gatekeepers. The total number of submissions received was 132, mainly from university first-year students who studied English recently. After removing uncompleted responses we obtained 124 completed questionnaires. The survey was conducted under departmental ethical approval.

The questionnaire was composed of four sections:

1. An informed consent page

2. A demographic information page (age range, gender, subject study domain (e.g. science, engineering, humanities...))
3. Questions concerning language testing: background qualification levels, the experience of online testing, confidence level and comments on style and challenges in the use of test systems and training provided prior to testing.
4. A reduced version of the System Usability Scale (Brooke 1996) focussing on general impressions of online testing.

### 3 RESULTS

Some previous studies reported that there is an impact of the test types and question types as well. However, there are not many pieces of research that considered the impact from the user perspective of a test-taker performance. The survey questions are mainly to examine the user experience. Data is being analysed quantitatively using Qualtrics software in order to describe it statistically. This approach assists in classifying received responses from the participants based on the categories of the questions and builds a comprehensive picture of the observed results. Most of the participants at the undergraduate or postgraduate level of study are from the under-35 age group. They studied the English language as a second language and their top six first languages were Chinese, Arabic, French, Greek, and Hungarian. The background of the participant is mostly Engineering and science.

The 1st part of the questionnaire is about general information to classify the participants according to age, gender, first language and subject area. Moreover, most of the participants have done the test online, and only 14.7% of them have never done the online test before as shown in figure1.

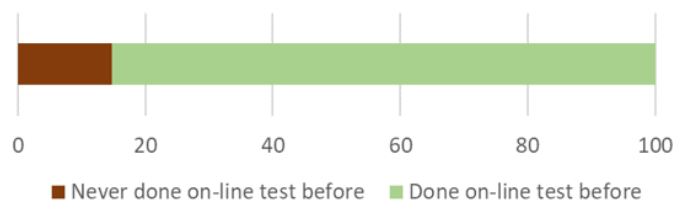
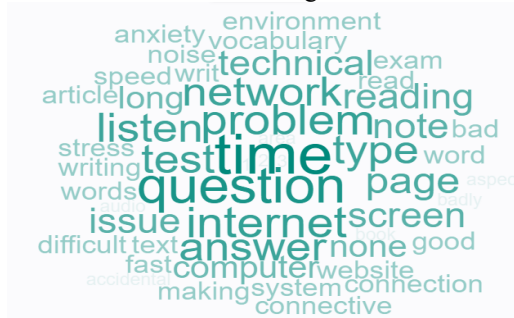


Figure1: Percentage Of Students Who Had An Online Test And Did Not.

The 2<sup>nd</sup> part of the questionnaire is about the challenges of taking a test online and how test-takers feel during the test. For this purpose, we used an open-ended question to find out the correlation between some elements such as interaction difficulty and challenges that may be counted during the test. From the responses to the question of challenges students may face some of them have a problem related to familiarity with test structure and others do not have, meanwhile, most of them have challenges. These challenges can affect their performance in the test and it is summarized by frequently

mentioned in responses (time, question type, technical problem, environment and vocabularies, etc.) as shown in the word cloud image.



**Figure2:** Word Cloud of Most Affected Challenges On The Performance.

The main focus is on the correlation between variables and the situation of test-takers during the test. For instance, the correlation between the first language of students and the level of confidence. As shown in the table most frequent mother language of the participants and level of confidence. It seems no significant effect from the first language side, but with the Chinese language, the shape of alphabetical characters may affect some of their stress and less confidence during the test.

**Table1:** Level Of Confidence Corresponding To The Language Of Students.

Students first language	Frequently	Level of confidence percentage
Chinese	36	27% unconfident
		21% less confident
		41% confident
Arabic	25	50% confident
		50% less confident
Spanish	3	All confident
Hungarian	4	All confident

In our investigation about the difficulty that test-takers may have, the results of the survey showed that 66% of the participants have difficulty with an online test and 34% do not have difficulty. This provides insight into figuring out reasons for that which will be for the rest of the project. Furthermore, training on how to use the online test systems plays a fundamental role in this part, 85% of participants have not been trained on the system before the test had difficulty during the test. In the last part of the questionnaire participants were asked about anything need be improved, and most of the answers were summarised as follows:

- Provide help tools such as highlighting or underlining and giving students note features so that they may put a mark on any question they did not understand or leave it to the end.
- Question and answer on the same page by making all questions and their answers on the same page.

- Test structure improvements to help students navigate through exam questions.
- Preferred questions some students prefer multiple-choice.
- Navigation icons and links should be obvious.

These improvements need to be done on the system side, however, from a user perspective, the system is in the middle of being easy to use as evaluated by the participants. Figure 3 shows.

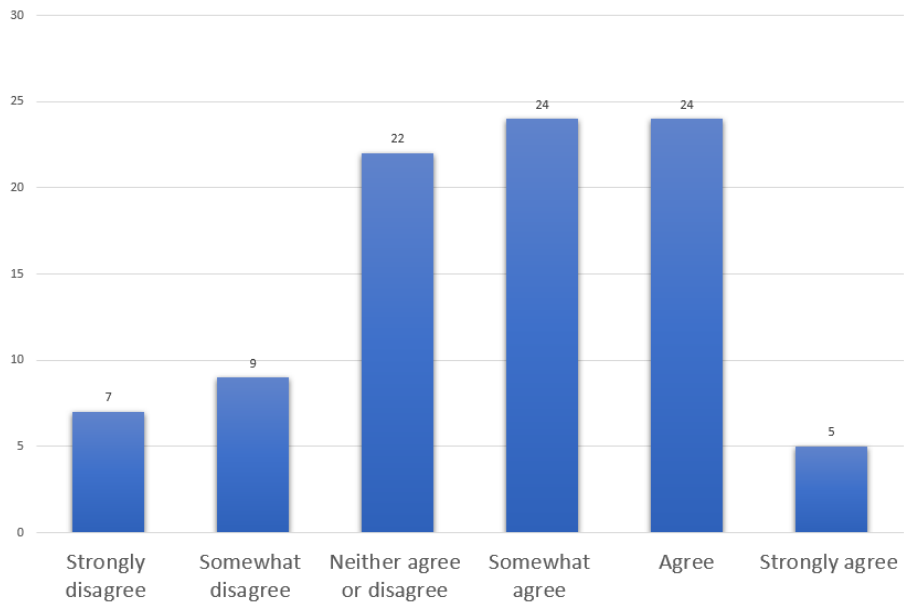


Figure 3: Participants' Opinions About Improve The Usability Of The Test.

#### 4. DISCUSSION

The result towards the end of this study which is based on two used methods the questionnaire and the literature of previous research demonstrated that the impact of the test can be from the test structure or a test-taker aspect. This largely confirms the research questions as follows: the level of confidence is differentiated between students based on other factors such as the first language which is the most noticeable specifier. In addition, in writing tests students would prefer the online method rather than speaking tests ( Wu et al 2010). In the questionnaire we have not measured the performance of students during the test but the correlation between confidence level and perform better can be extracted from the challenges mentioned by participants. This gives insight into affected elements of the test to help students become more confident during the test in order to get a better score.

Furthermore, as shown in the result section most of the students are at the normal confidence level. Firstly, test design has a limitation that can affect the performance of students such as page layout, question type, navigation icons and others. These factors are taken from test-takers experience and previous studies. However, it cannot be generalized due to the limited samples of participants and the variety of test designs. Secondly, from the user perspective, some factors can affect user performance. For instance, user background and subject area which is meant adapting to a computer for students who utilize computers mostly. From this side, we identify that prior experience with a computer helps students to perform better on online tests and improve user satisfaction. Additionally, training was raised as the most important issue, particularly for students who are not familiar with a test on a technology device. The training leads to answering the test easily and saves time as well. As a result, this would improve a student's score on the test.

## **5. LIMITATION**

In addition to the literature review, this study was conducted with a relatively small number of participants (n=124) primarily those already in University courses completing a questionnaire online, but it was an in-depth study and in line with the previous studies in the area called the impact of online English testing. Although we do not believe there was a direct impact on responses, recruitment and study were carried out during international lockdowns due to the Covid-19 pandemic.

## **6. CONCLUSION AND FUTURE WORK**

This paper has investigated the impact factors that may affect the performance of students on English as a Foreign Language online testing. It has provided insight into how the challenges of the test environment impact the performance of students. The main contribution of this research is summarizing the main challenges faced during an English online test. As presented in the result section, evaluating the user experience with English online tests shows a disparity between students depending on background, their first language, test design and others. The approach of training on how to do the test had been reported in a previous study about writing tasks only but in this study, we recommend prior test training should be taken for all English test skills. Preliminary findings of this study enable us to investigate the influence of additional factors from the test design and how can we overcome them. In particular, we are now investigating the impact of pre-test training to reduce the impact of tests on the accuracy of students' results. More participants and more personal studies may also report more factors and challenges of the English online test. These may play a fundamental role in improving the scores of a test-taker on an English online test. The recommendation of this continuing study also is to evaluate the accuracy and effectiveness of the test structure against the system usability scale.



**References:**

1. Budin, M., *Investigating the relationship between English language anxiety and the achievement of school-based oral English test among Malaysian Form Four students*. International Journal of Learning, Teaching and Educational Research, 2014. **2**(1).
2. Brunfaut, T., L. Harding, and A.O. Batty, *Going online: The effect of mode of delivery on performances and perceptions on an English L2 writing test suite*. Assessing Writing, 2018. **36**: p. 3-18.
3. Breland, H., Y.-W. Lee, and E. Muraki, *Comparability of TOEFL CBT essay prompts: response-mode analyses*. Educational and Psychological Measurement, 2005. **65**(4): p. 577-595.
4. Zheng, C., et al., *The relationship between English language learners' motivation and online self-regulation: A structural equation modelling approach*. System, 2018. **76**: p. 144-157.
5. Munn, Z., et al., *Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach*. BMC medical research methodology, 2018. **18**(1): p. 1-7.
6. Hamel, C., et al., *Defining Rapid Reviews: a systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews*. Journal of Clinical Epidemiology, 2021. **129**: p. 74-85.
7. Lai, C., *The influence of extramural access to mainstream culture social media on ethnic minority students' motivation for language learning*. British Journal of Educational Technology, 2019. **50**(4): p. 1929-1941.
8. Cole, J. and R. Vanderplank, *Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning*. System, 2016. **61**: p. 31-42.
9. Lee, J.S. and K. Lee, *Informal digital learning of English and English as an international language: The path less traveled*. British Journal of Educational Technology, 2019. **50**(3): p. 1447-1461.
10. Masoumi, G.A. and K. Sadeghi, *Impact of test format on vocabulary test performance of EFL learners: the role of gender*. Language Testing in Asia, 2020. **10**(1): p. 1-13.
11. Rezaei, O., H. Barati, and M. Youhanaee, *The Effect of Content Familiarity and Test Format on Iranian EFL Test Takers' Performance on Test of Reading Comprehension*. International Journal of Applied Linguistics and English Literature, 2012. **1**(4): p. 1-14.
12. Saiful, A., et al., *EFL LEARNERS' DIFFICULTIES IN THE STRUCTURE AND WRITTEN EXPRESSION SECTION OF TOEFL TEST IN AN INDONESIAN UNIVERSITY*. Englisia, 2020. **7**(2).
13. Wu, W.-c.V., L.L. Yen, and M. Marek, *Using online EFL interaction to increase confidence, motivation, and ability*. Journal of Educational Technology & Society, 2011. **14**(3): p. 118-129.
14. Wu, W.-c.V. and M. Marek, *Making English a "Habit": Increasing Confidence, Motivation, and Ability of EFL Students through Cross-Cultural, Computer-Assisted Interaction*. Turkish Online Journal of Educational Technology-TOJET, 2010. **9**(4): p. 101-112.

15. Rokhанийah, H. and O.V. Putra, *DEVELOPING WEB-BASED ONLINE TEST SYSTEM TO BOOST IELTS ACADEMIC READING SCORE*. English Review: Journal of English Education, 2021. **9**(2): p. 9-18.
16. Brooke, J. 1996. "SUS: A "quick and dirty" usability scale". In *Usability evaluation in industry*, Edited by: Jordan, P. W., Thomas, B. A. Weerdmeester and McClelland, I. L. 189–194. London: Taylor & Francis.