An Investigation into EFL Learners' Evaluation of the On-Line Diagnostic Assessment: The Case of UDig Reading Test

Yan Gao[✉]

English School, Jilin International Studies University, Changchun 130117, China
gaoyan@jisu.edu.cn

Abstract. With the support of technology, on-line diagnostic assessment systems are developing all over the world. In Chines mainland, UDig diagnostic assessment system is a representative one, but it is still to be tested by EFL learners. Therefore, the present work is conducted to investigate EFL learners' evaluation of using the on-line diagnostic assessment, i.e., the UDig reading test. 225 English majors attended the UDig reading test on computers and 85 of them participated in the investigation by questionnaires. The results show that 73% of the students think highly of UDig. They find it convenient because the system immediately provides scores and diagnostic reports; and useful because they consider the report comprehensive, accurate, targeted, objective and reliable. However, the negative evaluations are mainly due to the eye fatigue caused by using computer or long screen time. With these results, it is hoped that the study could provide useful information for improving UDig system and make a better use of it in EFL learning and teaching.

Keywords: on-line diagnostic assessment · UDig · EFL learners · evaluation

1 Introduction

In the information age, the application of network and data mining technology has promoted the deep integration of information technology and foreign language education. Big data has been widely used in the innovation of language learning, teaching and assessing (Wang & Zhang, 2015). Multi-modal interactive teaching, personalized autonomous learning, customized teaching materials, multi-dimensional diagnostic assessment, etc. have become the trend of foreign language teaching and evaluation (Cheng & Chen, 2019). Therefore, this study places the emphasis on diagnostic assessment, especially the on-line diagnostic assessment, and intends to investigate into EFL learners' evaluation of its operation and effects.

2 Diagnostic Assessment and UDig

Diagnostic assessment is a set of diagnostic procedures aiming to identify learners' strengths and weaknesses related to their knowledge and skill in the target domain

(Lee & Sawaki, 2009). In recent years, diagnostic assessment has been widely used in the field of language assessment (Lee & Sawaki, 2009; Doe, 2015). It aims at measuring students' specific knowledge structure and information processing skills, and diagnosing and discovering students' strengths and weaknesses in cognitive ability (Leighton & Gierl, 2007: 3).

With the advancement of technology, targeted score report, automated grading and feedback become available in EFL language assessment; therefore, on-line diagnostic language assessment systems are developed around the world. The famous ones are DIALANG, Delta (diagnostic English language tracking) of Hong Kong Polytechnic University and DELNA (Diagnostic English Language Needs Assessment) of Auckland University.

In Chinese mainland, from the late 1990s to the early 21st century, scholars used experimental methods to discuss the framework, process, and technical requirements of computerized diagnostic assessment from the perspective of reading, and believed that it was necessary to strengthen the research on the overall model of computerized diagnostic assessment in the field of language assessment according to new theories and technologies (Du, 1999; Du & Gui, 2000). At present, Chinese scholars have begun to apply cognitive diagnosis technology to language assessment, and made remarkable progress (such as Chen & Zhao, 2013; Du & Ma, 2018). In terms of on-line diagnostic language assessment system in Chinese mainland, UDig is a promising one.

UDig (college edition) is an English ability diagnosis system, independently developed by Foreign Language Teaching and Research Press. The goal of the system is to help college students understand their strengths and weaknesses in English ability, determine the starting point and goal of learning, explore the role of diagnostic assessment in college English teaching, and promote the construction of an integrated mode of college English teaching, learning and evaluation. The system is designed on the basis of the China English Proficiency Scale, and provides online diagnosis and assessment for college students with scales 4 to 7, as well as individual and class feedback reports and suggestions on learning and teaching.

The system has been in use for 4 years, but it is still to be tested by EFL learners. Therefore, this study aims to investigating into EFL learners' evaluation of using UDig diagnostic assessment system.

3 Methodology

3.1 Participants

The participants are the freshmen of English majors in a private foreign language university. They have passed National College Entrance Examination (short for NCEE), with English scores ranging from 105–136, among which the grade of 110–119 occupy the largest proportion of 45.5%. 225 students attended the UDig reading test. They had never used the system, but received a training before the test. 85 of them participated in the investigation.

3.2   Questionnaire Design

This investigation was conducted by open-question questionnaires, and the questions are listed as follows.

1. How do you evaluate your experience of using UDig? Why?
2. What do you think are the advantages of this system in terms of examination content, examination form, diagnosis report, diagnosis reliability and other aspects you'd like to comment on?
3. What do you think are the shortcomings of this system in terms of examination content, examination form, diagnosis report, diagnosis reliability and other aspects you'd like to comment on?
4. What suggestions would you like to provide for UDig to improve itself?

3.3   Data Collection

The investigation was conducted by open-question questionnaires through the website "Wenjuangxing" after UDig reading test.

The UDig reading test was held in the second week of the first term in students' first academic year, when they had little college study and remained their high school English level. The test was arranged in computer rooms of the school and supervised by two teachers a room. Students did the reading test on the computers independently and were not allowed to refer to any reference materials or to other people. After the test, the UDig system immediately provided test scores and diagnostic reports on students' overall reading skills and micro-skills. Students were asked to read the reports before they left.

Immediately after the test, students were assigned the questionnaires through the website "Wenjuangxing" (https://www.wjx.cn/), 85 test takers answered the questions in the questionnaires.

3.4   Data Analysis

After collecting the data, with the help of the website "Wenjuanxing" (https://www.wjx.cn/) and Excel, the author analyzed students' evaluation of using UDig diagnostic system by the means of keyword frequency analysis.

4   Results and Discussion

4.1   EFL Learners' General Evaluation of UDig

Question 1 How do you evaluate your experience of using UDig? Why?
The author examined the answers and eliminated some answers concerning students themselves rather than the system, and finally, 69 answers were left for analysis. Through keyword frequency analysis, the study discovers the following results.

Table 1. EFL Learners' General Evaluation of UDig

Words	Frequency
good	35
not bad	15
bad	19
eye	5
computer	4
screen	3

From the Table 1 and figure, it can be seen that most of students (51%) gave a completely positive evaluation of the system, and only 27% of the students did not like the testing system. Moreover, students provided different reasons for their evaluations.

As for the "good" evaluations, 11 students provided reasons from three aspects. First, 5 students commented on the design of system. They mentioned that "I can operate the system fluently"; "It is nice to have a time reminder"; "It has a clear overall appearance, and there is no advertisement"; "I feel good when using the system"; and "The system is quite humanized". Secondly, 3 students commented on examination content. They wrote that "The questions are at high level"; "The questions are of high quality"; "The questions are new, and reading materials contain a lot of knowledge and a large vocabulary". Thirdly, 3 students commented on diagnostic report, saying that "The diagnostic report let me know what I lack of"; "It can reflect my own level"; "The diagnosis is in great detail".

As for "not-bad" evaluations, students did not provide too many reasons, just the evaluations.

As for "bad" evaluations, 10 students provided reasons from different perspectives. First, 50% of the students complained about eye fatigue due to long screen time. They wrote "Since I stare at the screen for a long time, my eyes are blurred"; "My eyes are uncomfortable because of electronic screen"; "My eyes hurt"; "I am tired; my eyes hurt", and "My screen is dazzling and my eyes are uncomfortable". Secondly, 30% of the students were not used to doing tests on computers. They wrote that "I'm not used to answering questions on computers"; "The answer sheet on computer is quite different from the paper ones, and I'm not used to it"; "Computer is not suitable for me to do the test". Thirdly, 20% of the students were dissatisfied with the quantity of questions in the test. They believed that "There are too many questions" and "Since there are too many questions, I have no patience to finish them".

To summarize, generally 73% of the students had a high opinion of UDig system in terms of system design, examination content and diagnostic report; while 27% of the students believed the system was bad because of eye fatigue, inadaptation to answering questions on computers and the length of test.

In the following questions, detailed the evaluation of the system will be discussed.

4.2 EFL Learners' Favorable Comments on UDig

Question 2 What do you think are the advantages of this system in terms of examination content, examination form, diagnosis report, diagnosis reliability and other aspects you'd like to comment on?

In analyzing this question, the author examined students' favorable aspects and the comments they had on the system. During the analysis, the author deleted 17 ineffective answers, that is, answers without comments on the system, and 68 of the answers were left for further analysis.

First, the author summarized and calculated the aspects that students favored. The results are shown in Table 2.

As is shown in the results, students mainly favored the design of system and diagnostic report, which are also the key features of UDig system.

For these aspects, students provided different comments and the author presents the frequent ones in Table 3.

Table 2. EFL Learners' Favorable Aspects of UDig

Favorable aspects	Number of supporting students
The design of the system	44
Diagnostic report	21
The design of questions	3

Table 3. EFL Learners' Favorable Comments of UDig

Favorable aspects	Favorable comments	Frequency
The design of the system	convenient	16
	quick	9
	clear	7
	easy	3
Diagnostic report	comprehensive	5
	accurate	5
	targeted	3
	objective	3
	reliable	2
The design of questions	creative	2
	moderate (difficulty)	2

The results show that, in terms of the design of system, students are most satisfied with its convenience, i.e., the system provides the score and diagnostic analysis, which is convenient for students to judge their English level by themselves. Moreover, students are also satisfied with the fast speed of the system because the score and diagnostic report are provided immediately after they finish the test. Furthermore, the clarity and easy operation also gained students' favor.

As for the diagnostic reports, students praised its comprehensiveness and accuracy most, targeted recommendation and objective judgments less, and reliability the least.

For the design of questions, students mentioned two favorable features, that is, creative design and moderate difficulty.

To sum up, in this question, students favored the design of the system most, especially its convenience. Next, students think highly of the diagnostic report because it provides comprehensive, accurate, targeted, objective and reliable analysis of students' strengths and weaknesses in reading. Finally, students mention that the test has a creative design and moderate difficulty.

In the following question, the negative evaluation will be discussed.

4.3 EFL Learners' Unfavorable Comments on UDig

Question 3 What do you think are the shortcomings of this system in terms of examination content, examination form, diagnosis report, diagnosis reliability and other aspects you'd like to comment on?

Through the examination of the answers, the author discovered that students mainly complained about the operation of the system, and their comments were also analyzed by word frequency as shown in Table 4.

From students' unfavorable comments, it reveals that most problems lie in the eye fatigue caused by using a computer or long screen time. Next, students also mention some problems with doing the text on computers. For example, they cannot find the wrong answers and the reasons; they cannot mark the key words or sentences in the reading materials like they did in the paper test; it is not easy to drag the word into the blank.

Table 4. EFL Learners' Unfavorable Comments of UDig

Word	Frequency
eye	20
computer	11
long time	4
errors	3
screen	2
filling blanks	2
to mark	2

However, it is still worth noticing that among 85 participants, 50 students believed there were no weaknesses in the system. It shows that students generally believe UDig test system is a useful and convenient test system.

In the following question, students' suggestions will be analyzed.

4.4   EFL Learners' Suggestions for UDig

Question 4 What suggestions would you like to provide for UDig to improve itself?

In the questionnaire, students offered a few suggestions. They mentioned that "It would be better to indicate the wrong answers and provide reasons"; "I hope the designers would enrich the type of questions"; "I hope the system could distinguish the question I finished and the one I didn't; otherwise, I often miss some questions"; "I suggest adding pre-use instructions".

5   Conclusion

In recent years, the importance of diagnostic assessment in language learning has been widely recognized (von Davier & Lee, 2019), and with the support of the technology, on-line diagnostic assessment systems are springing up. In Chinese mainland, UDig is the representative one. With the aim of obtaining EFL learners' evaluations of using the system, the author organized 225 students to participate in the UDig reading test and investigate their comments on the system.

The results show that 73% of the students hold a positive opinion on UDig. They believe that it is convenient for them that the system could immediately provide the score and diagnostic report; especially, the report is comprehensive, accurate, targeted, objective and reliable. Moreover, students consider that the system is fast at providing the test score and diagnostic report. Furthermore, students like its clarity and easy operation. However, 41% of the students point out the drawbacks of the test. More than half of them complain about eye fatigue related to using computer or long screen time. Some mention problems with the operation of the computers. For example, they cannot find the wrong answers and the reasons; they cannot mark the key words or sentences in the reading materials like they do in the paper test; it is not easy to drag the word to the blank, and so on.

To conclude, EFL learners generally consider that UDig test system is useful and convenient to help student to recognize their strengths and weaknesses in EFL learning. Further studies could lay emphasis on empirical studies of employing UDig in EFL teaching and learning to make a better use of it.

References

Cheng X., & Chen, P. (2019). Conception of English reading ability training and evaluation system based on big data. Foreign Language Audio-visual Teaching, (02),40–44+60.

Chen H., & Zhao G. (2013). Application of cognitive diagnosis-A new stage of language testing research. Foreign Language Testing and Teaching (2), 1-9.

Doe, C. (2015). One teacher's take on using a 'test' for diagnostic purposes in the classroom. Journal of English for Academic Purposes, (18), 40–50.

Du J., (1999). Diagnostic test in foreign language teaching. Foreign Language Teaching and Research, (4), 40-43.

Du J., & Gui S. (2000). Experimental research on computerized reading diagnostic test. Foreign Language Teaching and Research, (5), 345-351.

Du W., & Ma X. (2018). Construction of English reading diagnostic model based on cognitive diagnostic assessment. Foreign Language Teaching and Research (1), 74-88.

Lee, Y. W., & Sawaki, Y. (2009). Cognitive diagnosis approaches to language assessment: an overview. Language Assessment Quarterly, 6(3), 172–189.

Leighton, J. & M. Gierl (eds.). (2007). Cognitive Diagnostic Assessment for Education: Theory and Applications [C]. Cambridge: Cambridge University Press.

Von Davier, M., & Lee, Y. S. (2019). Handbook of diagnostic classification models: Models and model extensions, applications, software packages. New York: Springer.

Wang Z. & Zhang Y. (2015). Research on big data methods and foreign language teaching innovation-take "2015 million English writing research on the same topic" as an example. Foreign Language Audio-visual Teaching (05),3-8.