ЁШ ТАДҚИҚОТЧИ МИНБАРИ

УЎК (УДК, UDC): 004.94:378.12

ENHANCING ASSESSMENT SYSTEM VIA TECHNOLOGY-ADAPTED TESTING

Mahkliyo Hasan qizi KHAZRATKULOVA
2nd year student,
English language and literature faculty-1
Uzbekistan State World Languages University
Tashkent, Uzbekistan
xazratqulova15.05@gmail.com
Supervisor: Gulbakhor Abdukadirova

Abstract

This review article provides analysis of articles and researches conducted on the topic of technology-assisted testing in language learning during two last decades. Based on the works of scholars published in *Language Learning&Technology* 5 research articles, 6 review articles and 2 published books were analyzed. The potential of this means of assessment was studied and the works that should be done in order to improve its effectiveness were clarified.

Key words: Computerized-adaptive testing; assessment; validity; reliability; authenticity; efficiency; innovative.

ТЕХНОЛОГИЯЛАРГА МОСЛАШТИРИЛГАН ТЕСТЛАРДАН ФОЙДАЛАНИБ, БАҲОЛАШ ТИЗИМИНИ ТАКОМИЛЛАШТИРИШ

Махлиё Хасан қизи ХАЗРАТҚУЛОВА
2-курс талабаси
Хорижий тил ва адабиёти факультет №1
Ўзбекистон давлат жаҳон тиллари
Ташкент, Ўзбекистон
xazratqulova15.05@gmail.com

Ушбу мақолада сўнгги йигирма йилда таълим жараёнида тил ўрганувчилар билимини баҳолашда технологиялардан фойдаланиш бўйича олиб борилган тадқиқотлар таҳлил қилинди. Олимларнинг "Language Leaning & Technology"да чоп этилган ишлари – бешта илмий, олтита тавсифий мақола ва иккита китоб ўрганиб чиқилди. Ушбу баҳолаш воситаларининг салоҳияти ўрганилиб, уларнинг самарадорлигини ошириш учун бажарилиши зарур бўлган ишлар аниқланди.

Калит сўзлар: компьютер ёрдамида баҳолаш; баҳолаш; ишончлилик; ҳақиқийлик; самарадорлик; иновацион.

СИСТЕМА ОЦЕНИВАНИЯ С ПОМОЩЬЮ ТЕХНОЛОГИЧЕСКИ АДАПТИРОВАННОГО ТЕСТИРОВАНИЯ

Махлиё Хасан кизи ХАЗРАТКУЛОВА
Студент
Факультет английского языка и литературы
Узбекский государственный университет мировых языков
Ташкент, Узбекистан
xazratqulova15.05@gmail.com

Аннотация

В этой обзорной статье представлен анализ статей и исследований, проведенных по теме технологического тестирования в изучении языков за последние два десятилетия. На основе работ ученых, опубликованных в Language Learning & Technology, было проанализировано 5 научных статей, 6 обзорных статей и 2 опубликованных книги. Потенциал этого средства оценки был изучен, и были разъяснены работы, которые должны быть выполнены для повышения его эффективности.

Ключевые слова: Компьютерное-адаптированное тестирование; оценка; валидность; надежность; достоверность; эффективность; инновационний.

Introduction

Assessment of learners' language proficiency is an important part of language education, which has been influenced by computer technology at least as significantly as language acquisition has, as it was written in the book of Chapelle, C. A., & Voss, E. (1, 116). Indeed, technology plays an integral role in our lives, and advances in technology which will make possible the dimensions of an expanding array of constructs, it is clear that the usage of technology-adaptive testing for language assessment and other educational/occupational assessment purposes will become increasingly predominant in the immediate hereafter. Computer-adapted assessment refers to the research and practice that goes into elaboration and testament of tests that use technology to interactively monitor test takers' performance and branch, based on an algorithm specified by the test developer. As it is clear for professionals in language learning, branching can be developed to accomplish different purposes, and the same is true in language testing.

The efficiencies of computer-adapted assessment

According to many scientists in the sphere of assessment, a successful testing of different language skills needs to be highly authentic, valid, and reliable enough to be used. Taher Bahrani enlightens these features briefly in his article – "Technology as an Assessment Tool in Language Learning". Authenticity in general refers to real world activities. Accordingly, if a test is supposed to be authentic, it should reflect the real world activities (2, 5). As a matter of fact, the activities which are used in testing should not be artificial such as matching items. Communicative aspects of the language in real world communication should be reflected in them. Validity

means assessing according to what is supposed to assess. The content of the test is one kind of validity which should reflect the purpose of the test.

The last but not the least important factor in testing is reliability which generally refers to the consistency of the scores. In his regard, the traditional methods of testing such as multiple-choice, which are still used by many teachers, often lack authenticity and validity because they do not represent how we use the language for communication with other people in informal setting outside the classrooms. They are rather artificial than real.

It is mentioned in the review article of Chapelle, C. A., & Voss, E. (1,116) that

technology-usage in the process of language testing, like in teaching language, was presented in the 1960s with the motivation of making the process of testing more efficient.

García Laborda discussed the efficiency level of the very assessment type clearly in his 2007 review article in which he concluded - the merits of online assessment overcome any of its drawbacks, as it is faster, more effective and less costly compared to traditional paper-and-pencil judgement of language proficiency level (3, 3). In addition, multimedia prompts can help make tests feel more 'real.' I think this kind of trial can make the difficult assignment of rapid diagnosis easier, and self-correcting tests can improve the procedure of correction, feedback, and reporting.

I believe that computer-adapted testing allows greater accuracy and effectiveness in several types of assessment by adapting to it after having evaluated a test-taker's language proficiency level and offering only those points that are thought to provide with the most valuable data about that subjective.

The scientific article – "Technology as an Assessment Tool in Language Learning" written by Taher Bahrani investigates the implementation of various technology-based tools such as computer, podcast, and chat for evaluating language mastery in EFL classrooms (2, 5). The main objective of the very paper was to introduce some activities which can be incorporated with various technologies that can be utilized not only for language acquisition but also for language assessment. Accordingly, teachers can use computers, the internet and mobile phones to encourage the procedure of language learning, as well as evaluating it. He says that assessing by technology, which is related to as alternative assessment differs from inherited typical testing (2, 5). The technology-adapted assessment includes inventive ways and

strategies to measure language proficiency. The very kind of assessment will be efficient only when it provides the language learners an opportunity to use what they have learnt.

The benefits of evaluating reading skills through technology were the main theme of the research – "Issues in Computer-Adaptive Testing of Reading Proficiency" by Micheline Chalhoub-Deville (4, 12). The initial section of this research article includes two chapters on the L2 reading construct. In the first chapter – "If reading is reader-based, can there be a computer-adaptive test of reading?" he gives the opinion on this issue that given our current state of knowledge regarding L2 reading, it is doubtful that test items can tap the individual components of reading or the interactive nature of the reading process. In the second chapter, "Developments in reading research and their implications for computer-adaptive reading assessment" William Grabe offers a comprehensive review on this reading research and synthesizes recent findings in the field (5, 22). He draws a line under the lack of connection between reading research and assessment and, like Bernhardt, questions the validity of current reading assessment methods. Grabe concerns that the prevalence of traditional reading comprehension tests is due in part to the success, according to psychometric criteria, of traditional approaches to testing due to the fact that traditional tests are easy to administer, to score, and to scale, and they are economical. This chapter suggests that reading assessment practices should move beyond their limitations by means of a serious research agenda.

Howard Wainer's book – "Computerized Adaptive Testing: A Primer" concludes by turning a critical eye on computerized adaptive testing, asking whether any real merits actually outnumber the problems (like cost, security risks, bias, etc.) that have been observed in operational test use (6, 25). It is vital for the sphere of language assessment to do the same. A diversity of CATs, including placement exams for EFL programs, large-scale academic language ability testing, and L2 proficiency assessment involving receptive as well as productive skills have been begun to be operated by language testers.

Yet, there are some drawbacks of such assessment as those mentioned in the conclusion of this book, and a number of further questions remain to be answered about the function of CATs in language assessment, including

• To what extent are language learning/ability constructs, most of which are multidimensional by nature, compliant to adaptive models that assume the testing of one-dimensional phenomena?

- How can computer- based testing (or CBTs) reply to the increasing requisition for performance assessments in language education and professional contexts (e.g., Norris (7, 100)).
- Are computerized renditions of currently subsisting tests (e.g., "New" TOEFL) doing a better more efficient job of addressing the inferences we need to be making, and if not, is the effort/expense guaranteed?

However, this book doesn't directly address such issues, *Computerized Adaptive Testing: A Primer* should nevertheless be claimed reading for anyone to attempt to answer these and related questions. It has congregated the essential

theoretical, statistical, and practical basics that cannot be provided in any other single resource, and it will enable language testers to assess the extent to what kind of CATs are applicable for the kinds of consequences and intentions we call for to address in language testing.

 Conclusion

The literature review provides data only some features of technology-adapted testing. The sphere of t assessment is wide open that ripe for further investigation. With the enhancement of Web 2.0 and 3.0, huge amount of opportunities for novelties, development and change are waiting to implement. Majority of the very technologies that play a significant role in language teaching are also put to work in language assessing, but they can take on new meaning. Grasping their meaning requires knowledge of basic concepts in language testing, which to date have not become common currency among language teachers and other professionals in applied linguistics. In these developments teachers' role, as well as students', and learning itself are continuously changing to suggest existing possibilities for foster elaboration, as it was mentioned in the records of Kristina Mullamaa (8, 112).

Professional development and teacher education are key areas of interest for those who want to develop best pedagogical practices on how to integrate useful assessments in a computer-assisted language learning (CALL) environment.
As Brown predicted in his review of technology and language assessment in the first issue of LLT, the technology-related issue of language assessment will continue to increase in their detail and complexity, and the need will continue to grow for professionals capable of negotiating the considerations that come into play in the design of computer-assisted language tests (9, 10).

In searching answers for the questions related to technology-based assessment, its characteristics, professional development, tools and the impact on students it

seems this type of assessment requires more than just new technology, but also new attitudes and skillsets.

REFERENCES

1. Chapelle, C. A., & Voss, E. 20 years of technology and language assessment in Language Learning & Technology, Language Learning & Technology. June 2016, Volume 20, No 2, 116–128 pp.

2. http://llt.msu.edu/issues/june2016/chapellevoss.pdf

3. Taher Bahrani. Technology as an Assessment Tool in Language Learning. International Journal of English Linguistics Vol. 1, No. 2; September 2011

4. www.ccsenet.org/ijel

5. Garcia Laborda, J. Introducing standardized EFL/ESL exams. Language learning&Technology. Vol 11 No 2 p3-9 Jun 2007

6. https://www.researchgate.net/publication/234747595_Introducing_standardized_EFLESL_exams

7. Micheline Chalhoub-Deville. Issues in computer-adaptive testing of reading proficiency. Cambridge University Press 1999, 257p

8. www.cambridge.org/97805221653800

9. William Grabe. Reading in second language-moving from theory to practice. Cambridge University Press, 2008. – 485 https://doi.org/10.1017/CBO9781139150484

10. Howard Wainer, Computerized adaptive testing: A Primer. Language Learning & Technology May 2001, Vol. 5, No. 2 23-27 pp

11. http://llt.msu.edu/vol5num2/review2/default.html

12. John M. Norris, Concerns with computerized adaptive oral proficiency assessment. Language Learning & Technology May 2001, Vol. 5, No. 2 99-105 pp   http://llt.msu.edu/vol5num2/norris/default.html

13. Mullamaa Kristina ICT in language learning. International education studies vol 3 num1 february 2010 112p

14. www.ccsenet.org.ies

15. Douglas Brown Language assessment principles and classroom practice. Pearson Education. 324p

16. https://doi.org/10.1177%2F0265532207086784