Institute of Vocational Education of NAES of Ukraine

ADAPTIVE TESTING IN THE CONTEXT OF USING ELECTRONIC LEARNING TOOLS: ESSENCE, DEVELOPMENT AND ASSESSMENT

Oleksandr Radkevych

Doctor of Pedagogical Sciences, Senior Researcher, Chief Research Associate of the Department of Monitoring and Evaluation of the Quality of Secondary Education, Institute of Pedagogy of the National Academy of Pedagogical Sciences of Ukraine https://orcid.org/0000-0002-2648-5726,
e-mail: mr.radkevych@gmail.com

Abstract

The relevance of the article is determined by the appropriateness of using adaptive testing in education based on electronic learning tools, in particular Google Forms for accurate measurement of knowledge and skills of learners.

Aim: to illuminate the general concept of adaptive testing of learners in the context of using electronic learning tools.

Methods: literature analysis was aimed at a detailed study of scientific works of foreign and domestic researchers, articles, books and other sources of information related to the object of research - to clarify the current state of the problem, identify unresolved issues and determine directions for further research; case study - for analyzing a specific case, or a series of cases in the context of the study; formation of conclusions.

Results: the significance of adaptive testing in the educational process, which adapts to the needs of each learner, is highlighted. Its advantages and disadvantages are disclosed. Varieties of adaptive testing are identified: linear, computer-based, and combined. The importance of adaptive testing using artificial intelligence is highlighted. The requirements for the preparation of adaptive tests are considered, in particular the importance of evaluation criteria and complexity parameters. The importance of feedback from learners and the need to review tests to maintain their relevance and validity are emphasized. The general safety rules when working with Google Forms and the importance of automatic grading of learner responses are characterized. The process of analyzing learner responses and displaying test results is defined. The possibilities of integrating Google Forms with educational platforms are considered. The advantages and limitations of using Google Forms for adaptive testing in general secondary education institutions are emphasized.

Conclusions: It is determined that adaptive testing is an important tool for deep analysis and creation of accurate complexity parameters of test questions. The need to implement feedback from learners for continuous improvement of the testing process is emphasized. The importance of regular review and updating of tests to ensure their relevance is noted. Attention is drawn to taking into account the differentiation of questions by level of complexity, relevance to the learning context, and general educational goals.

Introduction. Adaptive testing is becoming increasingly relevant in education due to its flexibility and ability to cater to the individual needs of learners (Wainer, 2000). Its core concept involves real-time test modifications based on each learner's responses. This approach has many advantages, including the ability to more accurately assess learners' knowledge, develop effective learning programs, and enhance their interest and motivation. However, teachers face certain challenges when implementing adaptive testing: *firstly*, determining the appropriate difficulty level for each test is a challenge for educators, especially with large groups of learners with varying levels of knowledge and skills. It is necessary to develop systems that can automatically adapt to changes in learners' knowledge levels during the testing process (Van der Linden, 2016); *secondly*, there are technical difficulties associated with developing adaptive tests. This requires a significant number of resources, including a large number of test questions, a system for processing learners' responses, and technology to ensure a proper testing process (Rasch, 1960; Magis & Barrada, 2017); *thirdly*, there's a need to create a vast amount of test questions. Teachers have to develop many questions of different difficulty levels, which is a labor-intensive process (Rasch, 1960); *fourthly*, it is necessary to regularly update the question bank. Questions must reflect the current content of study programs, which implies their constant updating, which can also be quite time-consuming. *fifthly*, there is a need for careful data analysis. The use of adaptive testing generates a large amount of data that requires analysis for the correct interpretation of results (Embretson, 1996). *sixthly*, there is the danger of overvaluing adaptive testing over other forms of learning. There is a risk that a teacher may focus too much on testing, ignoring other important aspects of the learning process, such as the development of critical thinking and creativity skills.

All this presupposes significant practical efforts, and scientific research in this field will contribute to the improvement of education quality and the expansion of individualized learning opportunities. Moreover, adaptive testing is aimed at improving the efficiency and validity of assessing educational outcomes and learners' skills, which helps better understand their needs and capabilities.

Sources. The issue of adaptive testing in education holds a central place in scientific discussions in recent years. This topic is the subject of numerous studies aimed at exploring the effectiveness of adaptive testing, as well as identifying potential problems and challenges that may arise during its implementation.

It should be noted that there is a large number of scientific works dedicated to the creation of methodological foundations for adaptive testing. They cover the development of algorithms for question selection, models for response evaluation, and systems for data processing. However, there are still aspects that require further research. Scientists continue to look for ways to the most accurate assessment, and also develop methods that could help withstand the possibilities of underestimating or overestimating the levels of knowledge of learners during adaptive testing (Embretson & Reise, 2000).

Another area of research is identifying ways to effectively implement adaptive testing in general secondary education practice. This involves the issues of developing technical support, training teachers, updating the question database, as well as interpreting and using the obtained results. Despite the significant amount of research conducted, some aspects of this problem still require further study (Wainer, 2000). One such study is the work of F. Jensen and T. Nielsen (2003), which focuses on studying the impact of adaptive testing on learners' motivation. The authors hypothesize that adaptive tests, which better correspond to the level of learners' knowledge, can increase their motivation for learning. Important in this context is the Handbook of Item Response Theory Modeling (Reise & Revicki, 2004), dedicated to exploring ways of using adaptive testing for evaluating learners' abilities in the context of critical thinking. The authors have developed a model of adaptive testing that can assess critical thinking at different levels of complexity. The work of B. Cheng and colleagues (2006) reveals the peculiarities of using adaptive testing in the context of distance learning. The authors explore how adaptive testing can assist in evaluating the learning process in an online environment.

Another important study is the scientific inquiries of M. Buehler and other researchers (2007), who revealed the possibility of using adaptive testing for evaluating learners' cognitive abilities. This has opened new perspectives for the use of adaptive testing from the standpoint of learning psychology. In particular, the results of positive psychology research (Furlong et al., 2009) are valuable, which has enabled determining the influence of adaptive testing on inequality in learning. It also helped understand how adaptive testing can influence levels of opportunity in learning. A. McCormick and colleagues (2015) determined the impact of adaptive testing on learners' learning. They investigate how adaptive testing affects learners' motivation, encompassing their attitudes to learning and participation in the learning process. S. Stone and B. Zhang (2017) focus on

developing a new approach to adaptive testing, which they call "feedback-based adaptive testing". This approach involves the use of feedback from learners for tuning the test in real-time.

S. Choi and N. Chung (2020) investigated the use of adaptive testing in the context of online learning. This work highlights how adaptive testing can support distance learning, as well as uncovers various challenges that learners may face when using this method in an online environment.

Aim: of this article is to highlight the general concept of adaptive testing for learners in the context of using electronic learning tools.

Methods: The literature analysis was aimed at a detailed study of scientific works by foreign and domestic researchers, articles, books, and other sources of information related to the object of research. The purpose was to clarify the current state of the problem, identify unresolved issues, and determine the directions for further research. Case studies were used to analyze a specific case, or a series of cases, in the context of the research; and to form conclusions.

Results and Discussion. Adaptive testing belongs to the computer-oriented method of evaluating educational outcomes, which adapts to the abilities of each specific learner during the testing process. This technology creates personalized test versions based on a learner's previous responses, allowing for more precise evaluation of knowledge and skills than traditional methods. Adaptive testing has many advantages. For example, it allows learners to feel challenged but not overwhelmed, as the difficulty level of each subsequent question corresponds to the learner's current level of knowledge. This can make the evaluation process less stressful and increase motivation for learning.

There are different types of adaptive testing. For instance, linear adaptive testing is a sequence of questions that change depending on the responses to previous questions. Usually, the testing starts with a question of average difficulty level, and then, based on the learner's answers, the difficulty level of the following questions is adjusted (Baraud et al., 2003). It is called "linear" because the questions are selected immediately before the testing and are presented in a certain order. The advantages of linear adaptive testing include: effective measurement of a learner's level of knowledge, as the questions adapt to the learner's existing knowledge; the tests are often shorter in duration but maintain a high level of accuracy. The disadvantages of linear adaptive testing include: the need for a large database of questions to ensure the validity of the test (for 10 test questions, the teacher needs to develop 30 questions, 10

questions for each level of difficulty); learners do not have the opportunity to return to previous questions.

Computer adaptive testing is a complex system of question selection based on previous responses from learners. That is, the computer, based on internal work algorithms, determines which question should be provided to the learner next (Lyashenko et al., 2017; Lyashenko, 2019). In other words, the test adapts to the learning outcomes of the learner, increasing the level of difficulty of questions for correct answers and reducing it for incorrect ones. The advantages of this type of testing include: the ability to more quickly and accurately assess learner outcomes, as the test focuses on questions that correspond to their level of knowledge, they may feel less stress and more motivation to continue the test; questions are usually selected in a way that most effectively assesses the abilities of a learner, reducing the time needed to complete the test. The disadvantages of such testing are: the involvement of significant efforts and resources, as it is necessary to create a large number of questions for different levels of difficulty and develop algorithms that will determine which questions to give based on the responses of the learner; for testing, access to computers or other digital devices, as well as a reliable internet connection, is necessary; learners cannot return to previous questions, which may limit the possibility of feedback.

Another type of adaptive testing is combined, which combines elements of linear and computer adaptive testing. Its essence lies in the block distribution of questions. At each stage of testing, usually, a small block of questions is presented and based on their answers, the system determines which block of questions should be given next. This approach provides more flexibility in question selection, reducing the computational power requirements compared to computer adaptive testing. It can also provide a more stable and accurate assessment of learners' knowledge and skills, especially in short tests. However, the test may be less adaptive, as each block of questions is chosen based on the answers to the previous block, not to each individual question. The advantages of combined testing include: greater flexibility in question selection, as "blocks" of questions can be developed in advance considering different strategies; better control of the content of the test, as blocks of questions can be balanced in content, complexity, and other characteristics; greater stability and accuracy of assessment, as it uses more information to select the next stage of questions. The disadvantages of this type of testing are: it requires complex planning and design compared to traditional adaptive tests; it may

cover a larger number of questions to create different blocks; it's not flexible enough to adapt to learner responses in real-time, as happens in traditional adaptive tests.

The modern approach to assessing learners' knowledge and skills, which uses machine learning algorithms and other technologies for automating the process of test creation and adapting them to the needs of each learner, is based on adaptive testing with the use of artificial intelligence. Artificial intelligence allows for automatic creation of new questions based on responses to previous ones without the involvement of a teacher, which ensures significant individualization of testing (Somenko et al., 2023). Its advantages include: creating a test that corresponds to the specific level of knowledge and skills of a learner, and also making the testing process more fair and accurate; AI systems can automatically generate and adapt questions, reducing the need for manual work and providing quick and effective testing; artificial intelligence can provide large-scale data processing to detect patterns and trends aimed to assist in identifying problem areas of knowledge or skills of learners. The disadvantages of this type of testing include: the results will be as high-quality as the data is well prepared and processed; artificial intelligence can make mistakes, which are difficult to detect and correct, especially if its algorithms are a "black box"; the implementation of AI systems requires significant resources, which may be unavailable for educational institutions with a limited budget; the use of artificial intelligence in educational contexts can raise ethical issues, such as information privacy, fairness, and transparency. All these types of adaptive testing have their own characteristics, advantages, and disadvantages, but what is common for them is that they allow the testing process to adapt to the individual needs of learners. We will substantiate the requirements for the preparation of adaptive tests. First of all, the test should consist of a large number of questions of different difficulty levels, covering the entire spectrum of skills or knowledge being assessed. A large number of questions is necessary so that the system has a sufficient number of options for choosing the next question, depending on the learner's answers. In addition, each question must be clearly defined by the level of difficulty (Radkevich, 2023). This requires preliminary testing of questions on a sample of learners to determine their level of difficulty, as well as to ensure that they are valid and reliable indicators of the knowledge they are intended to measure. An important aspect of preparing tests for adaptive testing is also the development of clear evaluation criteria. Since adaptive testing can lead to great

variability in the tests that different learners take, it is important to have clear criteria so that learner responses can be correctly and consistently evaluated. Finally, creating an adaptive test involves engaging experts from different fields, encompassing psychologists and experts in data processing and machine learning. They work together to not only accurately measure knowledge and skills but also to be fair and accessible to all learners. In this context, the process of developing a list of questions for adaptive testing becomes relevant. An important aspect is understanding that questions in adaptive testing serve not only to measure learners' knowledge and skills but also to dynamically adjust the testing process. Creating a list of questions begins with defining key learning goals and competencies that need to be assessed (Schäfer et al., 2023). Each question or task must be directly related to one or several of these goals. When the goals are defined, test developers create a wide range of questions that reflect different levels of difficulty – from simple to complex. Special attention should be paid to creating questions that measure deep understanding and critical thinking, not just memorization of facts. After the questions are created, they undergo a pre-testing process, during which the level of their quality and difficulty is evaluated. This process may include collecting data on learner responses, statistical analysis of these data, as well as feedback from teachers and other experts. Based on the data, each question is assigned a level of difficulty. The last step is to include the questions in the "question bank" for adaptive testing, where they can be used to assess learner learning outcomes. Let's emphasize the importance of defining the complexity parameters of the tests, which is a critical component of adaptive testing, as it allows the system to dynamically adjust the test for each individual based on responses to previous questions. The complexity parameters contain several key elements related to the questions: what level of knowledge or skills is needed to answer them? For example, a question that involves simple memorization of facts may be defined as less complex than a question that involves critical thinking or solving complex problems. However, complexity parameters are not determined only by the questions themselves. They also have to be related to real test performance data. This means that it is necessary to collect and analyze data on how learners respond to questions in real testing conditions. This data may include such information: the percentage of learners who answer the question correctly; the average time to answer the question; feedback from learners and teachers about the perceived difficulty of the question. Using this data to define complexity

parameters helps ensure that they reflect the real complexity of questions for learners, not just theoretical expectations.

Learners feedback is also an important element of the adaptive testing process, as it helps refine the testing process and content while ensuring the best understanding and engagement of learners. The scientific approach to evaluating learners feedback is based on several key principles: first, it is important to develop clear, consistent criteria for evaluating feedback. These criteria may be based on parameters such as the completeness of the feedback, its relevance to the test content, constructiveness, and objectivity; second, it is important to consider the context in which feedback is given. For example, learners who took the test at a high level may have a different perspective on the test compared to those who took it at a low level. Contextual factors, such as a learner's age, his previous testing experience, and general attitude to learning, can also affect feedback. Analysis of feedback from learners should include quantitative and qualitative analysis. Quantitative analysis may include assessing general trends, such as overall satisfaction with testing or common problems that arise. Qualitative analysis, on the other hand, involves a deeper dive into individual feedback to search for more detailed information about learner experience. An important aspect of evaluation is the development of an action plan for improving the testing process based on the information obtained (Noroozi et al., 2023). This can include revising test questions, changing instructions or the test format, as well as changing the way the test is administered.

To ensure the relevance and validity of tests, it is necessary to review and update the test. This process involves a number of stages that are based on careful analysis and a systemic approach. The first step is to develop a clear understanding of the goals of adaptive testing. Do the goals of the test really match the needs of the learners and the requirements of the curriculum? If not, then such test goals need to be reviewed or modified. The next step is to analyze test performance. This can include analyzing feedback from learners, analyzing data on learner performance, and also analyzing data on test usage. Data on learner performance can indicate questions or areas of the test that need attention and review.

Particular attention should be paid to differentiation by the level of complexity of questions, their relevance to the educational context, and their compliance with general education goals. It is also important to consider changes in the curriculum, new teaching methods, or learning technologies that could affect the relevance or validity of the test, which in turn is aimed at revising the test, its structure, format. The final step in the process of reviewing and updating an adaptive test is planning and implementing changes, which includes making changes to the assessment system or testing methodology.

In this context, the Google Forms electronic system, which is a free tool for creating tests provided as part of the Google Workspace package, is of great importance. This resource provides a wide range of possibilities for adaptive testing in general secondary education institutions, enabling easy creation and administration of tests, as well as processing their results (Miller et al., 2009). Google Forms is easy to use, thanks to its intuitive interface, teachers can easily create tests, and learners can take them. Google Forms has built-in analytical tools that allow you to quickly analyze test results, including automatic checking of answers and scoring. Google Forms also has built-in capabilities for adaptive testing. For example, you can use response-based transition logic, which allows you to automatically direct learners to questions depending on their previous answers (Chrysafiadi et al., 2018). In addition, Google Forms ensures security and privacy. All data is stored on reliable Google servers, and access to it is granted only to those given permission (Sclater et al., 2016). Also, to avoid fraud during testing, limits can be set on the number of attempts and test duration. Let's characterize the approaches to creating an adaptive test in Google Forms. To do this, several sections with questions need to be prepared. Each section should include questions that correspond to a certain level of difficulty. You can configure transitions between sections based on learner responses. For example, if a learner correctly answered a question in one section, you can redirect him to the next section with more difficult questions. If the answer is incorrect, the learner can be redirected to a section with less difficult questions or provided with additional study materials (Almond et al., 2015). Google Forms also allows for automatic evaluation of learner responses and provision of feedback. This can be useful for learners as they can immediately see their results. It also relieves the teacher of the need to manually grade tests. Using Google Forms for adaptive testing can make the assessment process more personalized and effective as it allows for better consideration of individual needs and knowledge levels of each learner. To this end, the system offers users a simple and intuitive interface that includes a main workspace and a toolbar containing relevant options for creating new tests and editing existing ones (Creswell, 2009). After registering an account on Google network and going to Google Forms, the teacher is presented with a home screen page where

all tests created by him will be displayed. To create a new test, simply press the "Create Form" button, located at the bottom right of the page (fig. 1). This allows opening a new screen with an empty form and a toolbar for setting its parameters.
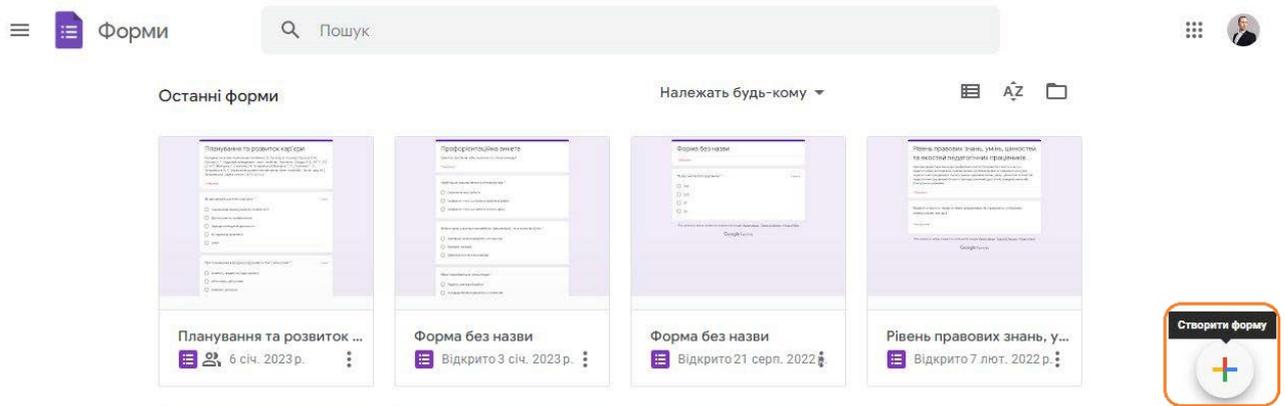

Fig. 1 Creating a new test

Initially, an empty form appears with one "Multiple Choice" type question. After that, the text of the question and potential answers are entered. Using the drop-down question type menu, the teacher can select various types of questions that cover: short answers, paragraphs, multiple-choice, checkboxes, drop-down, linear scale, multiple-choice grid, checkbox grid, date, time, and file uploads. Each type of answer has its own features. For example, the "Multiple Choice" answer type allows learners to select one answer from the provided options, while the "Checkboxes" type allows to select one or several answers. The "Scale" type is used for evaluating the answer to a certain question on a given scale (Medina-Díaz & Verdejo-Carrión, 2020).
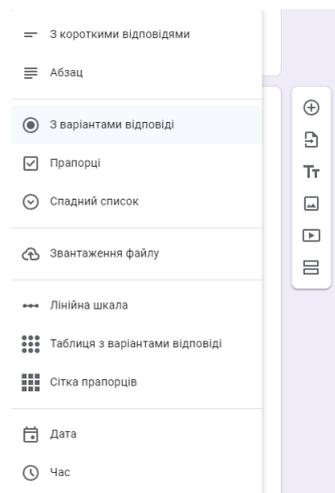

Fig. 2 Choosing question type in Google Form

To the right of the form, there is a toolbar (fig. 2), which allows you to add new questions, descriptions, create new sections, and implement various types of answers. Moreover, there are tools for incorporating media (images, video), as well as adding sections, which structure questions by difficulty levels (Iftakhar, 2016). Furthermore, you can adjust the form's parameters, including its color scheme, background image, font, and add a logo. Given this, the interface is easy to navigate, regardless of the teacher's technical skills, allowing a wide range of tasks to be performed from simple question creation to detailed adjustment of the form's visual layout.

To create an adaptive testing system in educational institutions, it is necessary to consider different levels of question difficulty. Note that they are all built on Bloom's taxonomy, which is designed to systematize educational objectives (Bloom, 1956). Hence, the three primary levels are easy, medium, and hard.

Easy level: this level usually aims to check the basic knowledge and understanding of the specific educational material being tested. Questions at this level come down to reproduction of facts,

ideas, or simple processes. They can include questions that require defining terms, enumerating facts, or explaining basic concepts. For example, in history, a question might be: "List the causes of World War I". This question checks if learners know the basic facts and can reproduce learned information.

Medium level: questions at this level require learners to have a deep understanding of the learning material and the ability to analyze information. They can include questions that require comparison and contrast of ideas, explaining cause-and-effect relationships, or data analysis. For example, in math, a question might be: "Explain why the result of adding two negative numbers is a negative number". Such a question checks if learners understand the concepts behind the procedures and if they can analyze mathematical relationships.

High level: questions at this level require learners to demonstrate a high level of critical thinking, covering the ability to evaluate, synthesize, or present information in new contexts. For example, in literature, a question might be: "How does the theme of loneliness manifest in Shakespeare's works, and how does it affect the characters?". Such a question expects from learners not only an understanding of the theme and characters but also analysis of how the theme affects the characters and the plot of the work, as well as the ability to draw conclusions based on this analysis. It's important to note that levels of difficulty aren't necessarily tied to the complexity of the question itself. They more so reflect the depth of thinking that the question anticipates, not its relative complexity.

Note that Google Forms provides the ability to adjust different types of answers, covering single choice, multiple choice, and text answer. Each of these types of answers is used to collect different types of data and can be useful in different scenarios (Proskura & Lytvynova, 2020). Single choice, also known as multiple choice, allows learners to select one answer from a given list of options. When creating a single-choice question, you need to enter the text of the question and then indicate the answer options. You can also add as many answer options as you need by clicking on "Add option".

Multiple choice, which in Google Forms is called "Checkboxes", allows learners to select several answers from a given list. Creating a multiple-choice question is similar to creating a single-choice question, but learners can choose several answer options instead of one.

Text response is another type of answer in Google Forms, which allows learners to enter open text as an answer. There are two types of text answers: short answers and paragraphs. Short answers are intended for short, specific responses, such as a name or date. Paragraphs can be used for longer responses, for example, for answers to open-ended questions. These types of answers in Google Forms provide great flexibility during test creation for surveys, allowing learners to choose the most suitable type of answer for each question. The above allows teachers to adjust different types of responses, covering single, multiple choice, and text answer, making this tool flexible and effective for test creation, especially in the context of adaptive testing.

The issue of choosing the complexity level of the first question is pertinent in this context. Like most scholars, we adhere to the stance of starting adaptive testing from an average difficulty level, which is supported by several scientific reasons: firstly, effectiveness: beginning with an average difficulty level ensures rapid identification of the approximate knowledge level. If the learner answers the question of average difficulty correctly, the system may proceed to more complex questions. Conversely, if the answer is incorrect, the system moves to less complex questions. This helps optimize the testing process, quickly determine the learner's knowledge level, and provide them with relevant testing; secondly, motivation: starting the test from an average difficulty level serves as motivation for learners. If the test begins with overly simple or overly complex questions, it may disconcert or disappoint too much; thirdly, time saving: the system can quickly narrow down the range of possible knowledge levels, reducing the time needed to determine this level; fourthly, accuracy: adaptive tests that start at an average difficulty level often prove to be more accurate in assessing learners' knowledge level compared to tests that start at extreme difficulty levels. Given this, we recommend beginning testing from the average difficulty level by selecting the first question of average difficulty after the introductory word (fig. 3).

Fig. 3 Transition to an average difficulty level at the start of the test

During adaptive testing, it is necessary to follow a logic that includes automatically directing learners to certain questions or sections of questions depending on their previous answers. This function is a crucial tool, especially when it comes to adaptive testing, as it allows the adaptation of the questionnaire to the needs of each learner based on his previous performance. The process of creating logic based on responses involves: firstly, creating a new form with different sections. Each section can include one or more questions. Note the possibility of considering each section as a separate "path" in the survey; secondly, adjusting each question so that, depending on the learner's answer, the corresponding "path" is chosen. To do this, the teacher needs to select the question they want to use as a transition, click the three dots and select the "Go to section based on answer" option (fig.4). It's also important to choose "Shuffle options" to prevent cheating during testing.
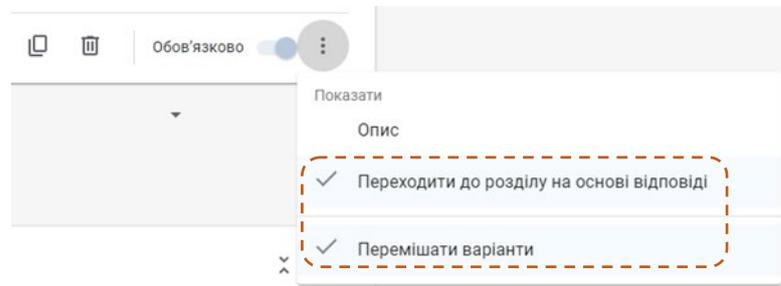

Fig. 4 Transition to the section based on the answer

After that, for each answer option, you need to select the section to which the transition will be made (Zhuang, 2022). This will allow setting transitions between questions: from easy to difficult (fig. 5). It is recommended to mark each section with a number with the question number to simplify the transition to the necessary section and question.
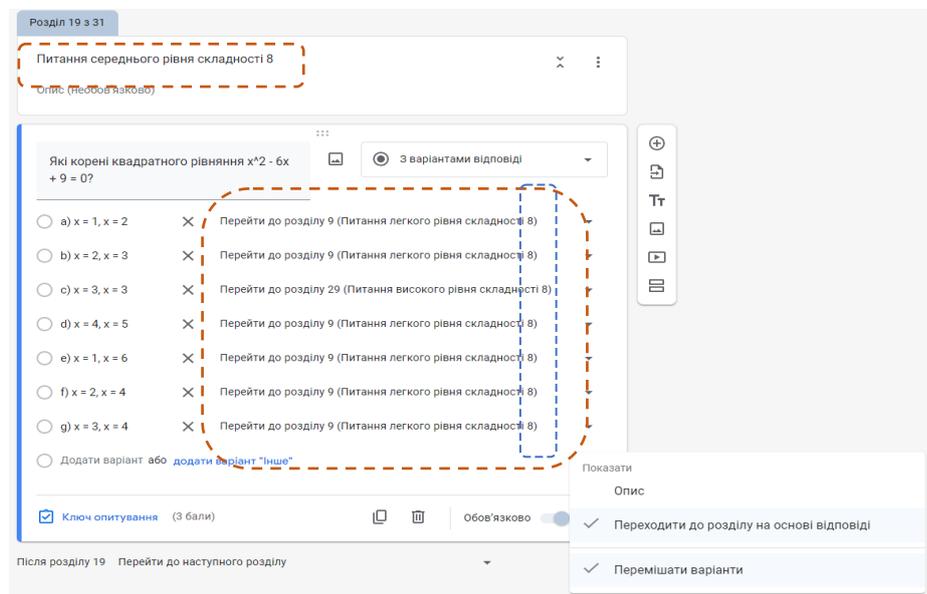
65

Fig. 5 Transition between questions by difficulty levels

Note that Google Forms automatically saves all changes made during the creation or editing of a form, so you don't need to worry about saving your work (Chen & Bryer, 2012). After all, all forms that are created are stored in Google Drive, where they can easily be found and edited. Regarding the publication of the form, Google Forms offers several options. The first way is to create a link to the form, which can be sent by email or through any messenger system. To create a link, you need to click on the "Send" button in the upper right corner of the form editing screen, and then copy the link that appears in the pop-up window.

The second way to publish the form is to embed the form on a website (fig. 6). This can be effective if you have your own website or blog and need to make the form accessible to authorized learners. For this, you need to click on the "Send" button and select the "< >" option, which will provide a code for embedding. Before this, you need to choose the width and height of the form (Tess, 2013).
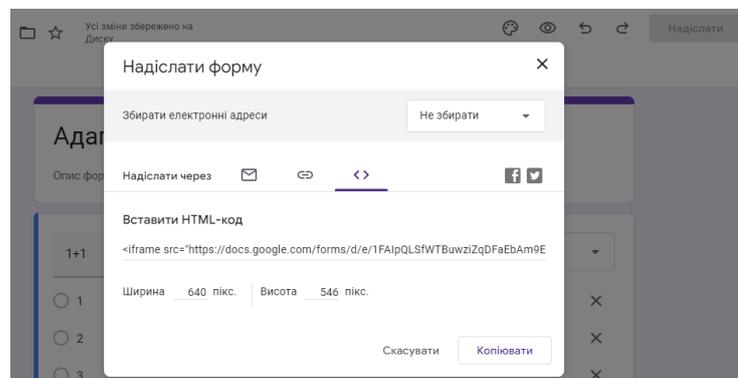

Fig. 6 Distribution of the form by embedding the code on the website

It is also important to consider the parameters of access to the form. The teacher can allow access to the form for all learners or restrict access only for learners of a specific class. It is also necessary to check whether learners fill out the form multiple times, or only once.

During adaptive testing using Google Forms, it is necessary to follow safety rules to protect personal data. Let's highlight a few key aspects of safety to consider: firstly, the data that is entered into Google Forms is stored on Google's servers. This means that caution should be exercised when dealing with confidential information and data (Leenes & Lucivero, 2014); secondly, an important aspect of security is the management of access to the form. Google Forms allows you to control who can view and fill out the form. The teacher can restrict access to the form only for learners from his class or make the form accessible to everyone. It is important to choose the right level of access, depending on who should have access to the form (Blythe & Camp, 2012); thirdly, it is necessary to set a reliable

66

password for your Google account, which is used to access Google Forms. This is the first level of data protection. In addition, it is recommended to enable two-factor authentication for an additional level of protection (Hadnagy, 2010); fourthly, it is important to regularly check for updates in Google Forms. Let us emphasize that Google Forms allows you to automatically assess learners' responses, which can enhance teachers' productivity. To set up automatic grading, you first need to enable the "Enable/disable grading" option in the form settings (Fig. 7).
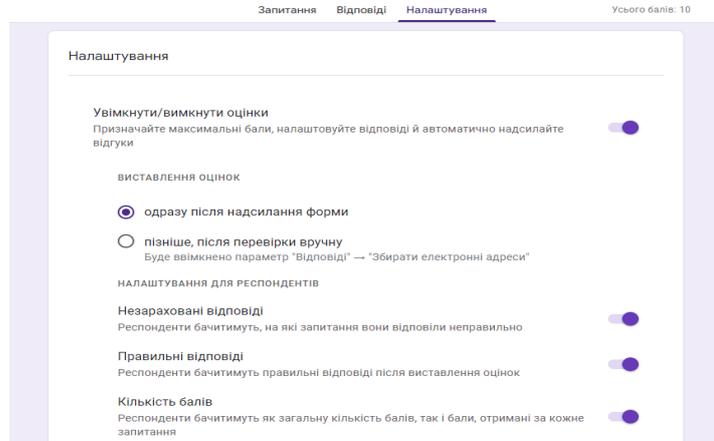

Fig. 7 Automatic Scoring

Afterwards, you can set the correct answers and scores for each question. This is done by selecting a question and clicking the "Survey Key" button. You can choose the correct answer and set the number of points that will be assigned for it (fig. 8). Note that according to the complexity, the number of points for the correct answer should also change. In this context, it is crucial to have a clear distribution of points for each question. Thus, if evaluation is carried out on a 100-point scale, they need to be divided into 30 +1 questions, into three levels of complexity: easy, medium, and high. Note that the first question should be of formative complexity and preferably should test the logic of the learner, considering this there will be not 10 questions, but 11.

For other questions, you can assign 2 points for easy, 3 points for medium difficulty, and 5 points for high difficulty questions. As a result, this will give 100 points. Note that this is just one of the possible variants of points distribution. Other answers will be considered incorrect and will decrease the level of complexity of the questions, and hence the number of points. In addition, you can also enter feedback for correct or incorrect answers.
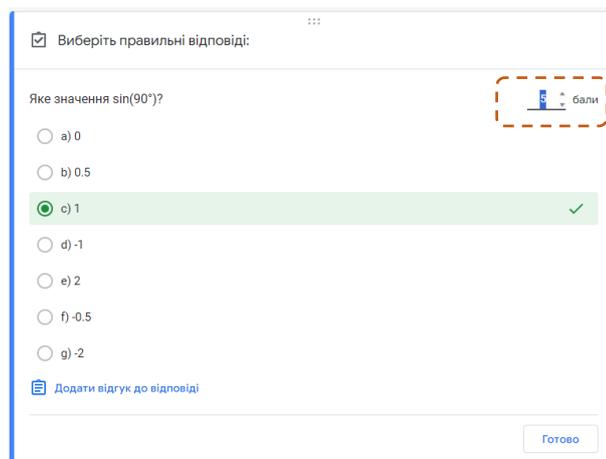

Fig. 8 Choosing the correct answer and assigning points for it

It is important to note that for the proper functioning of adaptive testing in the Google Forms system, a separate section (Fig. 9) should be created for each question. It should contain only one question. After answering the first question, the system will guide the learner along the route pre-prepared by the teacher. If the answer is correct, they will be directed to the first question of medium complexity, or to the second question of easy complexity.
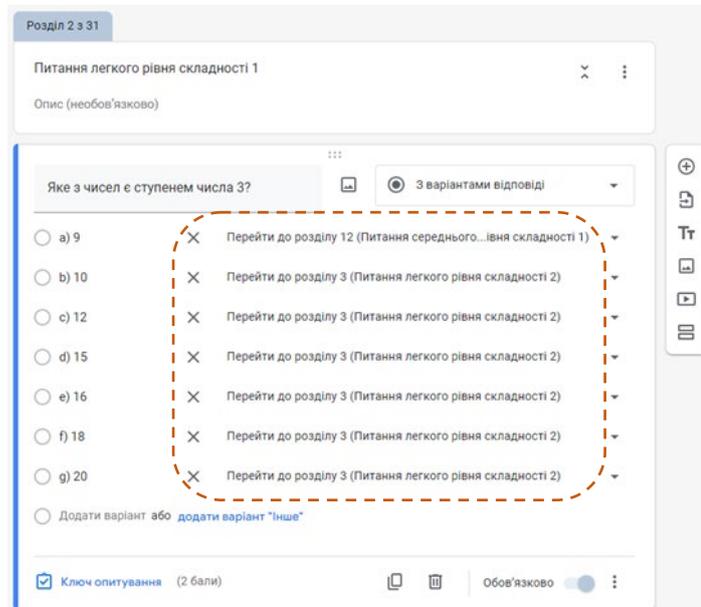
Fig. 9 Creating sections for each question and setting up transitions

If this is the last question of each level of complexity, it is necessary to choose "Submit form" in the answer transitions (Fig.10). This will successfully complete the test.
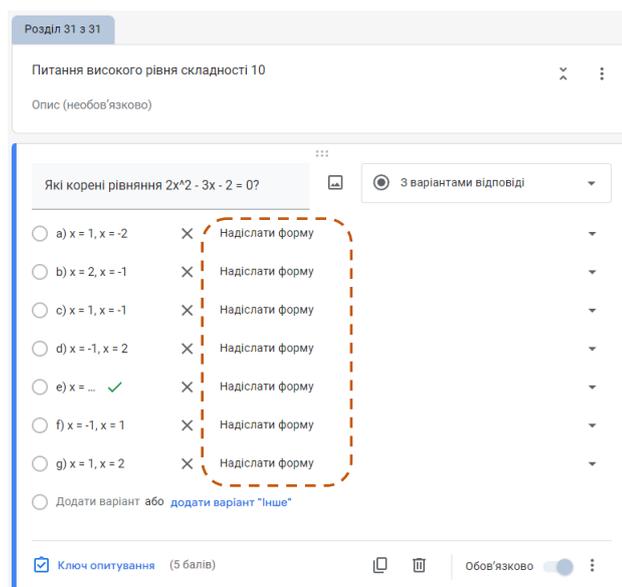

Fig. 10 Submitting the form

Google Forms provides convenient tools for analyzing learner responses and displaying test results. After completing the test, you can go to the "Responses" tab, which displays general answer statistics as well as detailed information about each learner's answers. You can view the number of answers to each question, as well as in percentages. This helps identify questions that were most difficult for learners, or to determine topics that may require additional explanation or repetition (Walvoord, 2010). Moreover, there is the possibility of viewing the answers of each learner separately to assess their level of understanding of the material and identify areas in which they may need additional help (Angelo & Cross, 2012). Not only does this enable teachers to quickly and easily assess test results, but it also helps identify gaps in learners' knowledge. In the context of supporting remote/distance learning in general secondary education institutions, Google Forms (Hodges et al., 2020) allows teachers to send forms online via email, link or embed them on websites. Learners can respond to them at their own pace, and teachers can receive feedback in real time (Moore, 2013). Google Forms also helps track learner progress. Teachers can create tests with automatic scoring and quickly send results to learners. Using Google Forms can enhance learning by providing interactivity and quick feedback.

Conclusions. Adaptive testing is an important and effective tool in the educational process, helping to better determine the level of knowledge and skills of learners. Its main principle is the dynamic selection of questions based on previous learner responses. The adaptive testing process covers many aspects, including developing and evaluating questions, determining complexity parameters, collecting and analyzing learner feedback, and regularly reviewing and updating tests. While there are types of adaptive testing - linear, computer-based, combined, and adaptive testing using artificial intelligence - they all have their own advantages and disadvantages. Their main common task is to provide a more accurate and individual assessment of the knowledge and skills of learners. In the educational process, adaptive testing is a critically important tool that involves conducting deep analysis and creating accurate test complexity parameters. These parameters should be based on actual test performance data, not just on theoretical requirements for questions. Integrating feedback from learners, the testing process can be continuously improved to meet the current needs of learners and the curriculum. This includes reviewing and updating tests to ensure they remain relevant and valid. In this process, it is important to consider the differentiation by the level of complexity of tasks, the relevance of the educational context, and compliance with the general goals of education. Therefore, adaptive testing is a complex but extremely important means of improving the quality of education, including general secondary. An important part of the adaptive testing process is the development of quality tests that require a large number of questions of varying complexity and their detailed preparation. In the context of adaptive testing, a number of key advantages of Google Forms have been identified, which is a flexible and effective tool for creating a variety of tests with questions of varying complexity, according to Bloom's taxonomy. This tool provides a large selection of answer types and customization options, covering single choice, multiple choice, and text answer. Of particular value is the logic of transition based on answers, which allows to automate the processes of directing learners to different sections of the test depending on their previous answers. This feature is extremely useful for implementing adaptive testing. Automatic grading of learner responses in Google Forms reduces the load on teachers, simplifying the grading process. In addition, using Google Forms to analyze learner responses allows teachers to identify gaps in learner knowledge and adapt the learning process to learner needs. In the context of remote/distance learning, Google Forms is a valuable tool that allows you to create and send forms online, providing teachers with the ability to track learner progress in real-time. However, when using Google Forms, it is important to consider security issues, including personal data protection, access management to the form, setting a reliable password, and regular checking for updates. Given this, the system can be very useful in the context of organizing adaptive testing in educational institutions, if used properly and adhere to recommendations for data protection. Considering the continuous development of technology, in the future, one can expect improvements in the adaptive testing process, in particular using artificial intelligence and big data analytics.

As a practical expression of our scientific research, we have developed a Comprehensive Adaptive Test for 9th grade in subjects: algebra, geometry, trigonometry, statistics, and probability using Google Form. You can familiarize yourself with it and take it at the link:
https://forms.gle/TpKQBojbFHZnjPT58

List of references

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian networks in educational assessment*. Springer.

Angelo, T. A., & Cross, K. P. (2012). Classroom assessment techniques. *Jossey Bass Wiley*.

Baraud, Y., Huet, S., & Laurent, B. (2003). *Adaptive tests of linear hypotheses by model selection*. The Annals of Statistics, 31(1), 225-251.

Bloom, B. S. (1956). Taxonomy of Educational Objectives. *Handbook I: Cognitive Domain Text*.

Blythe, J., & Camp, L. J. (2012, May). Implementing mental models. *In 2012 IEEE symposium on Security and privacy workshops* (pp. 86-90). IEEE.

Buehler, M., Griffin, P., & Ross, M. (2007). A *Self-Adaptive Algorithm for Context-Aware Ubiquitous Learning*. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Eds.), Educational Data Mining.

Chen, B., & Bryer, T. (2012). *Investigating instructional strategies for using social media in formal and informal learning*. International Review of Research in Open and Distributed Learning, 13(1), 87-104.

Cheng, B., Wang, M., Mørch, A. I., Chen, N. S., Kinshuk, & Spector, J. M. (2014). *Research on E-Learning in the Workplace 2000–2012: A Bibliometric Analysis of the Literature*. Educational Research Review, 11, 56–72.

Choi, S. W., & Chung, H. (2020). Online learning environments in higher education: Connectivism vs. dissociation. *Education and Information Technologies*, 25(6), 5991–6010.

Chrysafiadi, K., Troussas, C., & Virvou, M. (2018, October). A framework for creating automated online adaptive tests using multiple-criteria decision analysis. *In 2018 IEEE international conference on systems, man, and cybernetics (SMC)* (pp. 226-231). IEEE.

Creswell, J. W. (2009). *Research designs: Qualitative, quantitative, and mixed methods approaches*. Callifornia: Sage.

Embretson, S. E. (1996). The new rules of measurement. *Psychological assessment*, 8(4), 341.

Furlong, M. J., Gilman, R., & Huebner, E. S. (Eds.). (2009). *Handbook of positive psychology in schools*.

Hadnagy, C. (2010). Social engineering: The art of human hacking. *John Wiley & Son*s.   Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). *The difference between emergency remote teaching and online learning*.

Iftakhar, S. (2016). Google classroom: what works and how. *Journal of Education and Social Sciences*, 3(1), 12-18.

Jensen, F. V., & Nielsen, T. D. (2007). Bayesian networks and decision graphs (Vol. 2). *New York: Springer*.

Leenes, R., & Lucivero, F. (2014). Laws on robots, laws by robots, laws in robots: Regulating robot behaviour by design. *Law, Innovation and Technology*, 6(2), 193-220.

Magis, D., Yan, D., von Davier, A. A., Magis, D., Yan, D., & von Davier, A. A. (2017). An overview of computerized adaptive testing. *Computerized Adaptive and Multistage Testing with R: Using Packages catR and mstR*, 35-51.

McCormick, A. C., Kinzie, J., & Gonyea, R. M. (2013). Student engagement: Bridging research and practice to improve the quality of undergraduate education. *In Higher Education: Handbook of Theory and Research*: Volume 28 (pp. 47-92). Dordrecht: Springer Netherlands.

Medina-Díaz, M. D. R., & Verdejo-Carrión, A. L. (2020). Validity and reliability in student learning evaluation throughout active methodologies. *ALTERIDAD. Revista de Educación*, 15(2), 270-284.

Miller, M. D., Linn, R. L., & Gronlund, N. E., (2009). Measuring and Asssessment in Teaching (10th ed.) *Merrill Prentice Hall*, Upper Saddle River. NJ

Moore, M. G. (Ed.). (2013). Handbook of distance education. *Routledge*.

Noroozi, O., Banihashem, S. K., Biemans, H. J., Smits, M., Vervoort, M. T., & Verbaan, C. L. (2023). Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. *Education and Information Technologies*, 1-28.

Proskura, S. L., & Lytvynova, S. H. (2020). *The approaches to Web-based education of computer science bachelors in higher education institutions*.

Rasch, G. (1960). Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.

Reise, S. P., & Revicki, D. A. (Eds.). (2014). Handbook of item response theory modeling: Applications to typical performance assessment. *Routledge*.

Schäfer, M., Nadi, S., Eghbali, A., & Tip, F. (2023). Adaptive test generation using a large language model. arXiv preprint arXiv:2302.06527.

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education. London: Jisc. Accessed February, 8(2017), 176.

Stone, C. A., & Zhang, B. (2003). Assessing goodness of fit of item response theory models: A comparison of traditional and alternative procedures. *Journal of Educational Measurement*, 40(4), 331–352.

Tess, P. A. (2013). The role of social media in higher education classes (real and virtual) – A literature review. *Computers in human behavior*, 29(5), A60-A68.

Van der Linden, W. J. (2016). Unidimensional logistic response models. In Handbook of item response theory. *Chapman and Hall/CRC*.

Walvoord, B. E. (2010). Assessment clear and simple: A practical guide for institutions, departments, and general education. *John Wiley & Sons*.

Zhuang, Y., Liu, Q., Huang, Z., Li, Z., Shen, S., & Ma, H. (2022, June). Fully Adaptive Framework: Neural Computerized Adaptive Testing for Online Education. *In Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 36, No. 4, pp. 4734-4742).

Ляшенко, О. І. (2019). Оцінювання навчальних досягнень учнів за допомогою адаптивного тестування. *Освіта для миру= Edukacja dla pokoju*, 1, 178-189.

Ляшенко, О. І., Жук, Ю. О., Ващенко, Л. С., Гривко, А. В., & Науменко, С. О. (2017). Тестові технології оцінювання компетентностей учнів. https://undip.org.ua/wp-content/uploads/2021/08/test_2017.pdf

Радкевич, О. (2023). Еволюція електронних засобів внутрішнього контролю та оцінювання якості освіти: від початку до сучасності. *Rozwój nowoczesnej edukacji i nauki–stan, poblemy, perspektywy*, 115-127.

Соменко, Д. В., Трифонова, О. М., & Садовий, М. І. (2023). Штучний інтелект та нейромережі в освітньому процесі: переваги та недоліки. актуальні проблеми та перспективи технологічної і професійної освіти, 78.

Translated & Transliterated.

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). Bayesian networks in educational assessment. Springer. [in English].

Angelo, T. A., & Cross, K. P. (2012). Classroom assessment techniques. Jossey Bass Wiley. [in English].

Baraud, Y., Huet, S., & Laurent, B. (2003). Adaptive tests of linear hypotheses by model selection. The Annals of Statistics, 31(1), 225-251. [in English].

Bloom, B. S. (1956). Taxonomy of Educational Objectives. Handbook I: Cognitive Domain Text. [in English].

Blythe, J., & Camp, L. J. (2012, May). Implementing mental models. In 2012 IEEE symposium on Security and privacy workshops (pp. 86-90). IEEE. [in English].

Buehler, M., Griffin, P., & Ross, M. (2007). A Self-Adaptive Algorithm for Context-Aware Ubiquitous Learning. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero, & J. Stamper (Eds.), Educational Data Mining. [in English].

Chen, B., & Bryer, T. (2012). Investigating instructional strategies for using social media in formal and informal learning. International Review of Research in Open and Distributed Learning, 13(1), 87-104. [in English].

Cheng, B., Wang, M., Mørch, A. I., Chen, N. S., Kinshuk, & Spector, J. M. (2014). Research on E-Learning in the Workplace 2000–2012: A Bibliometric Analysis of the Literature. Educational Research Review, 11, 56–72. [in English].

Choi, S. W., & Chung, H. (2020). Online learning environments in higher education: Connectivism vs. dissociation. Education and Information Technologies, 25(6), 5991–6010. [in English].

Chrysafiadi, K., Troussas, C., & Virvou, M. (2018, October). A framework for creating automated online adaptive tests using multiple-criteria decision analysis. In 2018 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 226-231). IEEE. [in English].

Creswell, J. W. (2009). Research designs: Qualitative, quantitative, and mixed methods approaches. Callifornia: Sage. [in English].

Embretson, S. E. (1996). The new rules of measurement. Psychological assessment, 8(4), 341. [in English].

Furlong, M. J., Gilman, R., & Huebner, E. S. (Eds.). (2009). Handbook of positive psychology in schools. [in English].

Hadnagy, C. (2010). Social engineering: The art of human hacking. John Wiley & Sons. [in English].

Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). The difference between emergency remote teaching and online learning. [in English].

Iftakhar, S. (2016). Google classroom: what works and how. Journal of Education and Social Sciences, 3(1), 12-18. [in English].

Jensen, F. V., & Nielsen, T. D. (2007). Bayesian networks and decision graphs (Vol. 2). New York: Springer. [in English].

Leenes, R., & Lucivero, F. (2014). Laws on robots, laws by robots, laws in robots: Regulating robot behaviour by design. Law, Innovation and Technology, 6(2), 193-220. [in English].

Magis, D., Yan, D., von Davier, A. A., Magis, D., Yan, D., & von Davier, A. A. (2017). An overview of computerized adaptive testing. Computerized Adaptive and Multistage Testing with R: Using Packages catR and mstR, 35-51. [in English].

McCormick, A. C., Kinzie, J., & Gonyea, R. M. (2013). Student engagement: Bridging research and practice to improve the quality of undergraduate education. In Higher Education: Handbook of Theory and Research: Volume 28 (pp. 47-92). Dordrecht: Springer Netherlands. [in English].

Medina-Díaz, M. D. R., & Verdejo-Carrión, A. L. (2020). Validity and reliability in student learning evaluation throughout active methodologies. ALTERIDAD. Revista de Educación, 15(2), 270-284. [in English].

Miller, M. D., Linn, R. L., & Gronlund, N. E., (2009). Measuring and Asssessment in Teaching (10th ed.) Merrill Prentice Hall, Upper Saddle River. NJ [in English].

Moore, M. G. (Ed.). (2013). Handbook of distance education. Routledge. [in English].

Noroozi, O., Banihashem, S. K., Biemans, H. J., Smits, M., Vervoort, M. T., & Verbaan, C. L. (2023). Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. Education and Information Technologies, 1-28. [in English].

Proskura, S. L., & Lytvynova, S. H. (2020). The approaches to Web-based education of computer science bachelors in higher education institutions. [in English].

Rasch, G. (1960). Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. [in English].

Reise, S. P., & Revicki, D. A. (Eds.). (2014). Handbook of item response theory modeling: Applications to typical performance assessment. Routledge. [in English].

Schäfer, M., Nadi, S., Eghbali, A., & Tip, F. (2023). Adaptive test generation using a large language model. arXiv preprint arXiv:2302.06527. [in English].

Sclater, N., Peasgood, A., & Mullan, J. (2016). Learning analytics in higher education. London: Jisc. Accessed February, 8(2017), 176. [in English].

Stone, C. A., & Zhang, B. (2003). Assessing goodness of fit of item response theory models: A comparison of traditional and alternative procedures. Journal of Educational Measurement, 40(4), 331–352. [in English].

Tess, P. A. (2013). The role of social media in higher education classes (real and virtual)–A literature review. Computers in human behavior, 29(5), A60-A68. [in English].

Van der Linden, W. J. (2016). Unidimensional logistic response models. In Handbook of item response theory (pp. 41-58). Chapman and Hall/CRC. [in English].

Walvoord, B. E. (2010). Assessment clear and simple: A practical guide for institutions, departments, and general education. John Wiley & Sons. [in English].

Zhuang, Y., Liu, Q., Huang, Z., Li, Z., Shen, S., & Ma, H. (2022, June). Fully Adaptive Framework: Neural Computerized Adaptive Testing for Online Education. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 4, pp. 4734-4742). [in English].

Lyashenko, O.I. (2019). Otsinyuvannia navchalnykh dosiahnen uchniv za dopomohoiu adaptivnoho testuvannia [Assessment of students' achievements using adaptive testing]. Osvita dlia myru=Edukacja dla pokoju, 1, 178-189, [in Ukrainian].

Lyashenko, O. I., Zhuk, Yu. O., Vashchenko, L. S., Hryvko, A. V., & Naumenko, S. O. (2017). Testovi tekhnolohii otsiniuvannia kompetentnostei uchniv [Test technologies for assessing students' competencies]. https://undip.org.ua/wp-content/uploads/2021/08/test_2017.pdf [in Ukrainian].

Radkevych, O. (2023). Evolutsiia elektronnykh zasobiv vnutrishnoho kontrolu ta otsiniuvannia yakosti osvity: vid pochatku do suchasnosti [Evolution of electronic tools for internal control and assessment of education quality: from the beginning to the present]. Rozwój nowoczesnej edukacji i nauki–stan, poblemy, perspektywy, 115-127, [in Ukrainian].

Somenko, D.V., Tryfonova, O.M., & Sadovyi, M.I. (2023). Shtuchnyi intelekt ta neiro merezhi v osvitnomu protsesi: perevahy ta nedoliky [Artificial intelligence and neural networks in the educational process: advantages and disadvantages]. Aktualni problemy ta perspektyvy tekhnolohichnoi i profesiinoi osvity [Actual problems and prospects of technological and vocational education], 78, [in Ukrainian].

АДАПТИВНЕ ТЕСТУВАННЯ В КОНТЕКСТІ ВИКОРИСТАННЯ ЕЛЕКТРОННИХ ЗАСОБІВ НАВЧАННЯ: СУТЬ, РОЗРОБЛЕННЯ ТА ОЦІНЮВАННЯ

Олександр Радкевич

доктор педагогічних наук, старший дослідник, головний науковий співробітник відділу моніторингу та оцінювання якості загальної середньої освіти, Інститут педагогіки НАПН України, https://orcid.org/0000-0002-2648-5726 e-mail: mr.radkevych@gmail.com

Реферат

Актуальність статті зумовлена доцільністю застосування адаптивного тестування в освіті на основі електронних засобів навчання, зокрема Google Forms для точного вимірювання знань і вмінь здобувачів освіти.

Мета: полягає у висвітленні загальної концепцію адаптивного тестування здобувачів освіти в контексті використання електронних засобів навчання.

Методи: аналіз літератури спрямовувався на детальне вивчення наукових робіт зарубіжних та вітчизняних дослідників, статей, книг та інших джерел інформації, які стосуються об'єкта дослідження – для з'ясування існуючого стану проблеми, виявлення невирішених питань та визначення напрямів подальших досліджень; кейс-стаді (вивчення випадків) – для аналізу конкретного випадку, або ряду випадків у контексті дослідження; формування висновків.

Результати: висвітлено значення адаптивного тестування в освітньому процесі, що підлаштовується під потреби кожного здобувача освіти; розкрито його переваги та недоліки; з'ясовано різновиди адаптивного тестування (лінійне, комп'ютерне та комбіноване); висвітлено значимість адаптивного тестування з використанням штучного інтелекту; розглянуто вимоги до підготовки адаптивних тестів, зокрема важливість критеріїв оцінювання і параметрів складності; наголошено на значущості зворотного зв'язку від учнів і необхідності перегляду тестів для підтримання їх релевантності та валідності; схарактеризовано загальні правила безпеки під час роботи з Google Forms та важливість автоматичного оцінювання відповідей учнів; визначено процес аналізу відповідей учнів та відображення результатів тестування; розглянуто можливості інтеграції Google Forms з освітніми платформами; підкреслено переваги та обмеження використання Google Forms для адаптивного тестування в закладах загальної середньої освіти.

Висновки: з'ясовано, що адаптивне тестування є важливим інструментом для глибокого аналізу і створення точних параметрів складності тестових питань; підкреслено необхідність впровадження зворотного зв'язку від учнів для постійного вдосконалення процесу тестування; показано важливість регулярного перегляду та актуалізації тестів для забезпечення їх релевантності; акцентовано увагу на врахуванні диференціації питань за рівнем складності, релевантності навчального контексту та загальних освітніх цілей

Ключові слова: *адаптивне тестування; Google Forms; зворотний зв'язок; критерії оцінювання; автоматичне оцінювання; загальна середня освіта.*