NATURALISTA CAMPANO
ISSN: 1827-7160
Volume 28 Issue 1, 2024

Innovative Approach to Assessing University Students

Bazarbaeva Aygul Kuanishbaevna
Nukus branch of the Tashkent University of Information Technologies named after Muhammad al-Khorezmi

Abstract: This article examines innovative methods of assessment, namely multi-stage adaptive measures to improve the efficiency of student assessment procedures and obtain immediate feedback, reliable and valid assessments. The main methods that were applied in this article are mathematical models and measurements based on IRT models. Multi-stage adaptive measurements, as one of the innovative approaches that increase the effectiveness of student assessment, allow for the principle of individualization, updating in education and receiving immediate feedback to improve the learning process and educational content. Multistage adaptive measurement can be applied to blended learning, massive open online courses, and e-learning. The article may be of interest to teaching staff and experts in developing effective methods for assessing learning outcomes. In this regard, an organizational structure and block diagram of a student assessment program based on the pyramidal method have been developed.

Keywords: Adaptive measurement algorithms, assessment, multi-stage adaptive measurements, efficiency of the assessment process, software structure.

1. Introduction

Currently, a vital task of vocational education is to assess students and their learning outcomes in the context of a changing educational landscape, as well as the need to correlate vocational education programs developed on the basis of state educational standards with professional standards and labor market requirements. In an age of rapidly evolving educational technology and the accumulation of big data, it is necessary to develop a way to effectively assess students and obtain reliable results with high predictive validity.

The effectiveness of procedures for assessing learning outcomes can be achieved by choosing a method of combining quantitative and qualitative approaches in pedagogical measurements. This two-paradigm approach involves placing the results of quantitative and qualitative measurements of education on the same scale of levels. The application of this approach requires multi-stage measurements involving several stages, each of which uses measurement tools to assess knowledge, skills or competencies. Each stage of such measurements must correspond to a certain range of the level of the competence scale, the measuring instruments of which become more complex from the first to the next stage. Such multi-stage measurements are based on modern Item Response Theory (IRT)[1].

On the one hand, to obtain reliable results during the student assessment procedure, it is necessary to use a sufficient number of measuring instruments with stable parameters of their difficulty and ability to differentiate at each stage, which will lead to high accuracy of values and duration of assessment tests for each student. On the other hand, to ensure construct and predictive validity, assessment of learning outcomes within a competency-based approach requires the use of case studies, which require a significant amount of time to implement. Thus, it is necessary to look for opportunities to comply with the above conditions and move to multi-stage adaptive measurements.

Measurement is theoretically understood as the process of establishing correspondence between the characteristics being assessed and points on a scale, in which the relationship between various marks is expressed in numerical series of properties. The educational measurement process, which provides objective and comparable information, includes the measurement object (one or more hidden characteristics), measurement procedures, measuring instruments (tasks, test and scale for recording the measurement object marks), analysis and interpretation of measurement results. These components of the measurement process have their analogues in traditional educational control, but there these procedures are rather intuitive. In the case of educational measurements, each component is in the process of scientifically justifying its quality. This is especially important when it comes to final assessments, the results of which are used to make management decisions. In this case, the objects of measurement are the generated knowledge, skills and competencies, the structure and level of which are compared with the standards set out in the educational requirements as a result of their learning outcomes.

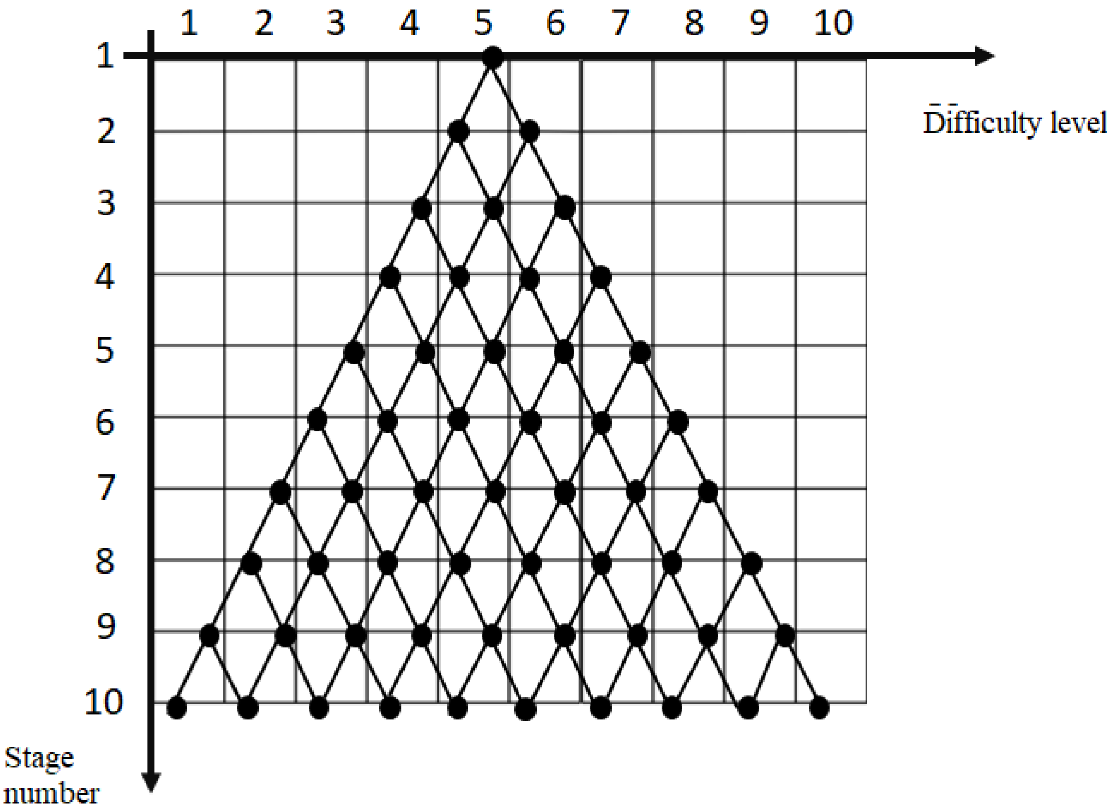


Figure 1. Job distribution for 10-step measurement.

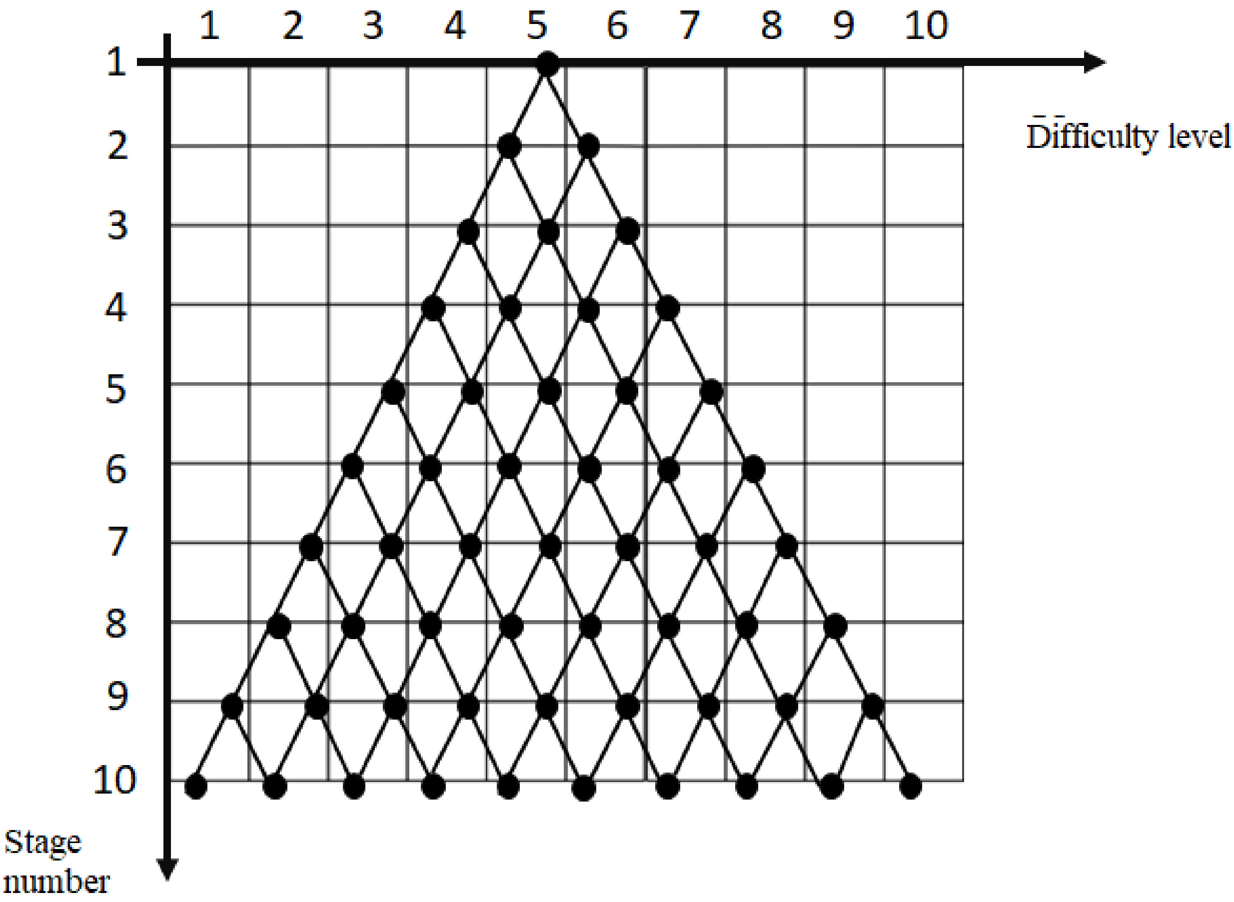
The transition to educational measurements as the most reliable and valid method of obtaining information about learning outcomes is due to the need to increase the objectivity, accuracy and efficiency of assessment processes. To obtain the most accurate and unbiased results, the final assessment procedure can lead to unnecessary costs (time, financial, human resources), so moving to multi-stage adaptive measurements seems to be the best and most effective way. This minimizes measurement error, and therefore increases its accuracy and test duration, and maximizes the reliability of marks.

The foundation of multistage adaptive measurement as an effective method for assessing learning outcomes comes from modeling approaches. The choice of strategy and development of an algorithm for providing a measuring instrument that acquires the effectiveness of educational measurements precedes the start of multi-stage adaptive measurements. Thus, adaptive measurements can be classified as two-stage and multi-stage, according to which various strategies and algorithms are developed.

Multi-stage measurement is an organization of measurement in which the subject moves along his own individual path in the process of completing sets of tasks that differ in number and complexity at each stage. The algorithm for selecting and presenting elements is based on the principle of feedback: after the subject selects

the correct answer, his next task becomes more difficult, but if the answer is incorrect, then the next element is easier than the previous one. Thus, adaptive multi-step measurement is based on context-sensitive algorithms, when the next step depends on the previous one and is carried out only after evaluating its results [2].
Multistage adaptive measurement strategies, in turn, are divided into fixed strategy and flexible strategy, depending on how the multistage adaptive measurement instruments are designed. If all students use the same set with a fixed position of measuring instruments on the difficulty axis, but each of them moves through the set depending on the results of their previous step, then the adaptive measurement strategy is deterministic-branching[3]. Difficulty measuring instruments in a set are usually spaced equidistantly from each other or choose a decreasing step to match the increasing difficulty, adjusting the pace of completion for the learner. In this article we will describe the most common strategy associated with the fixed strategy.

Pyramid method strategy. The essence of the pyramid strategy is that all students start with tasks of medium difficulty. If the student's answer is correct, then he is given a task with the next degree of difficulty. If the student answers incorrectly, he is given a less difficult subject. The procedure is repeated until the student submits the required number of answers. To implement the pyramid strategy, the number of items for each difficulty level in the test must be determined with a predetermined number of measurement stages (this coincides with the number of difficulty levels) [4].
Figure 1 shows an example of a 10-step measurement with 55 items. At the beginning, the student is given a task of medium difficulty (level 5). At the second stage, he may be given a task of the 5th or 6th level. Obviously, at each step different tasks can be given, the level of complexity of which coincides with the number of the completed step. If the test has tasks of 10 levels of difficulty, then in general each test taker is given 10 tasks out of 55 included in the test.



Picture 2. Job distribution for 9-step measurement.

It should be noted that the pyramid strategy at each difficulty level requires a certain number of tasks (Fig. 2). The largest number of tasks (equal to the number of difficulty levels) can be used at the secondary level. At the highest level, only one element is used. At adjacent levels, the number of tasks differs by 2 (with the exception of levels adjacent to the middle level). It can be seen that at the first level 2 tasks are required, at the second - 4, at the third - 6, etc. At the last level, 1 task is used, at the penultimate level - 3 tasks, etc.
You can use a general formula to determine the number of tasks at any difficulty level. Let's assume we have difficulty levels. The average number of difficulty levels is defined as the number of levels divided by 2: obviously, if X is even, then it is equal to X/2, and to get an odd X we get (X+1)/2. At a difficulty level with a number less than X/2, 2i tasks are used, where i is the number of levels. For difficulty levels with numbers greater than X/2, the number of tasks is equal to 2(X-i)+1, where i is the number of levels, but if the difficulty level is equal to X/2 then, if X\2 is odd, the number of tasks is 2 (X-i)+1, and when even it is equal to 2(X-i). In total, the test will use X(1+X)/2 levels [5].

Table 1. Number of tasks at each level of the pyramid strategy.

Difficulty level number	Total number of jobs	Questions at the corresponding number level																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
5	15	2	4	5	3	1															
6	21	2	4	6	5	3	1														
7	28	2	4	6	8	5	3	1													
8	36	2	4	6	8	7	5	3	1												
9	45	2	4	6	8	9	7	5	3	1											
10	55	2	4	6	8	10	9	7	5	3	1										
11	66	2	4	6	8	10	11	9	7	5	3	1									
12	78	2	4	6	8	10	12	11	9	7	5	3	1								
13	91	2	4	6	8	10	12	13	11	9	7	5	3	1							
14	105	2	4	6	8	10	12	14	13	11	9	7	5	3	1						
15	120	2	4	6	8	10	12	14	16	13	11	9	7	5	3	1					
16	136	2	4	6	8	10	12	14	16	15	13	11	9	7	5	3	1				
17	153	2	4	6	8	10	12	14	16	17	15	13	11	9	7	5	3	1			
18	171	2	4	6	8	10	12	14	16	18	17	15	13	11	9	7	5	3	1		
19	190	2	4	6	8	10	12	14	16	18	19	17	15	13	11	9	7	5	3	1	
20	210	2	4	6	8	10	12	14	16	18	20	19	17	15	13	11	9	7	5	3	1

Table 1 shows the number of elements at each level and in the test as a whole for various values of the measurement stage number. It is clear that the pyramid strategy can only be used if a large number of tasks of different difficulty levels are presented. However, this corresponds to a simplified understanding of multi-stage adaptive measurements.
Organizational structure of software. In the modern world, the process of learning, as well as assessing students' knowledge, plays an important role, just as it did decades ago. Many highly effective approaches to this process have been developed in the past, but it is worth noting that most of them are outdated as we currently have a huge amount of information, the processing of which costs large sums of money every year.
Multi-stage adaptive methods of knowledge assessment are designed to increase the efficiency of student assessment, allow for the principle of individualization, updating in education and receiving immediate feedback to improve the learning process.
Multi-stage adaptive measurements can be applied in blended learning, massive open online courses and online learning.

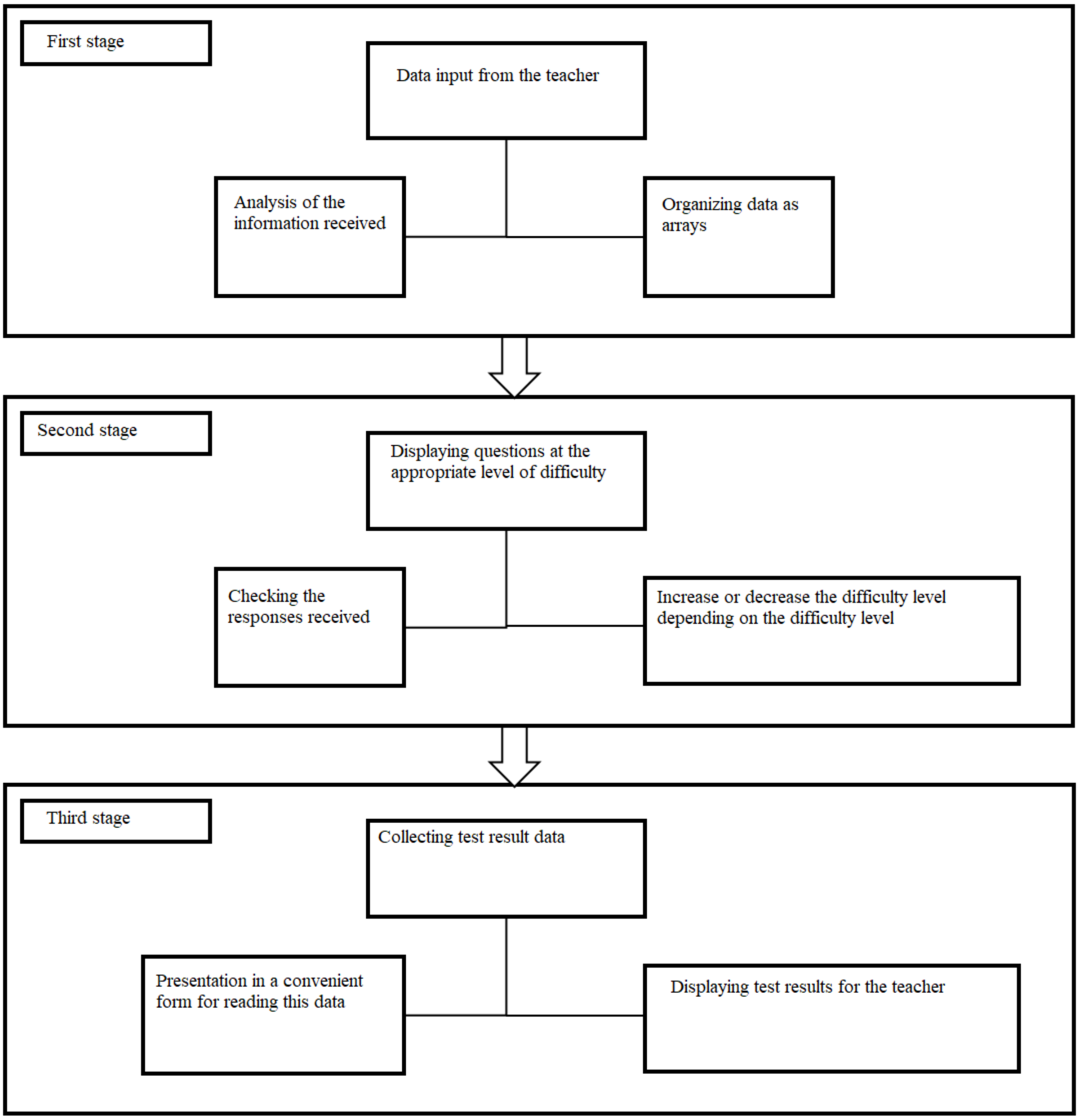


Figure 1. Structure of the student assessment program

Figure 1 shows the principle of operation of the pyramidal method; looking at this figure we can describe our method. At the first stage, the program must receive data from the teacher, such as the number of difficulty levels, questions for each difficulty level, and answers to them. At the second stage, the program sequentially asks questions to the student in accordance with the pyramidal strategy, then the algorithm checks the answers and decides to lower or increase the level of complexity of the questions asked. At the third stage, the program collects the results of the passage, structures them in a convenient form for reading and analysis, and then transfers them to the teacher.

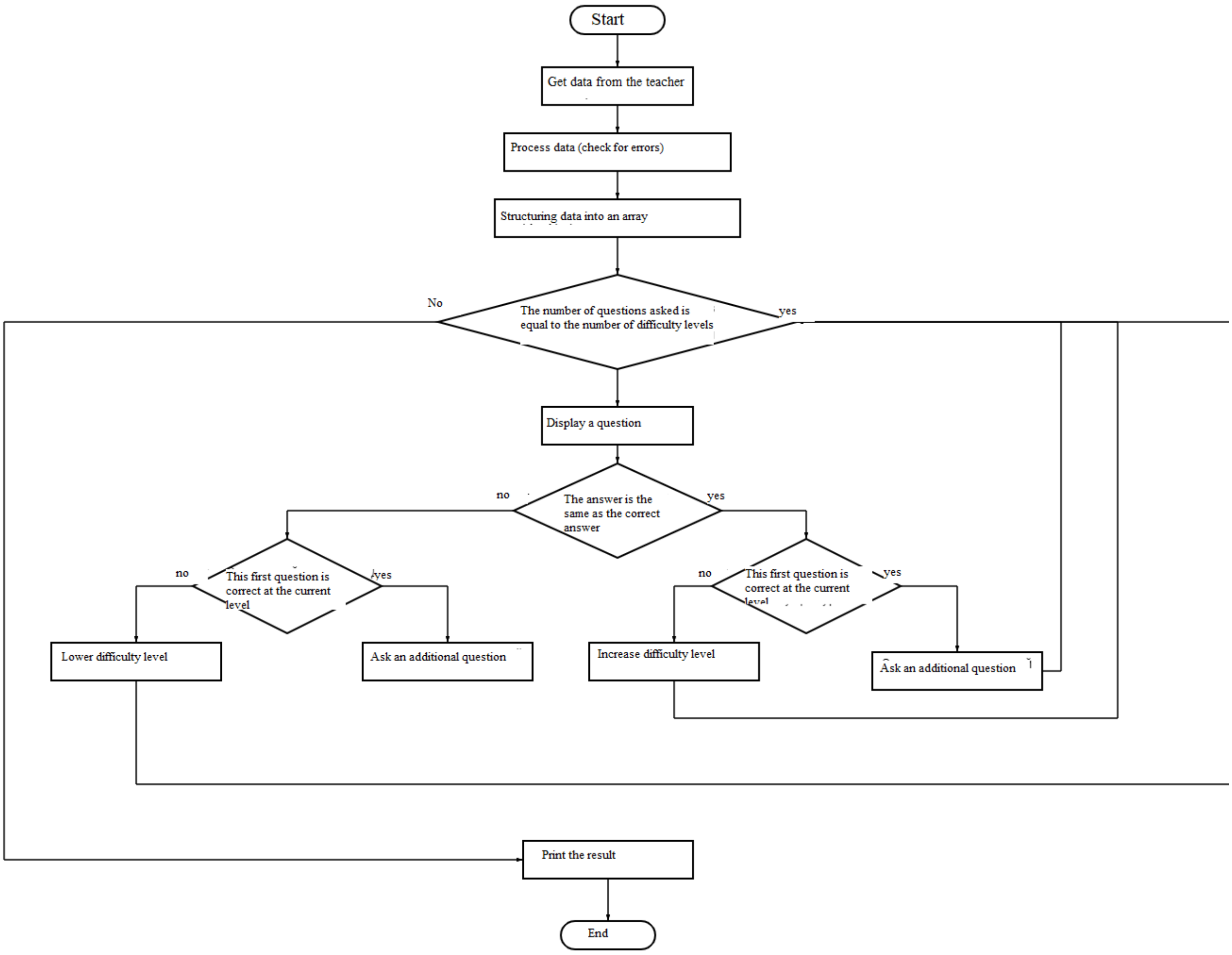


Figure 2. Flowchart of a student assessment program based on the pyramid method.

At the beginning, the program receives data from the teacher, checks the data for errors and writes it in the form of arrays. Next comes a check for the number of questions asked, as they should be equal to the number of difficulty levels, then the cycle begins its work: a question is displayed if the answer received is correct and at the same time, this is the first question that is asked at this difficulty level, then another question is displayed with given level of difficulty, and if this is the second question, then the level of difficulty increases. If the answer received was incorrect, and at the same time, this is the first question that is asked at a given difficulty level, then another question from this difficulty level is displayed, and if this is the second question, then the difficulty level is lowered. The procedure is repeated until the student submits the required number of answers, then the result is displayed.

To test the effectiveness of this method of assessing students' knowledge, an algorithm was developed, as well as a program with console output. In the first part, the teacher must enter the path to the file that contains the questions and answers, as well as the number of difficulty levels.

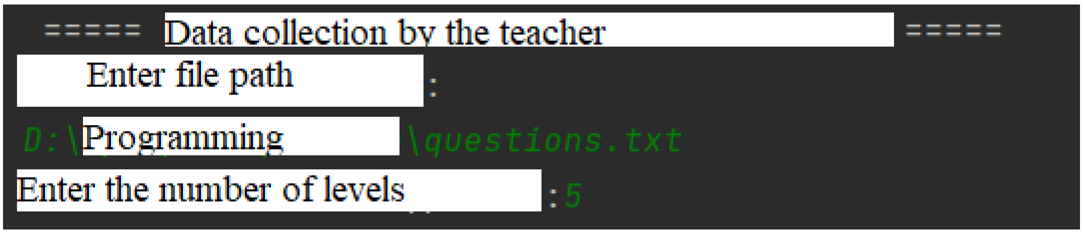
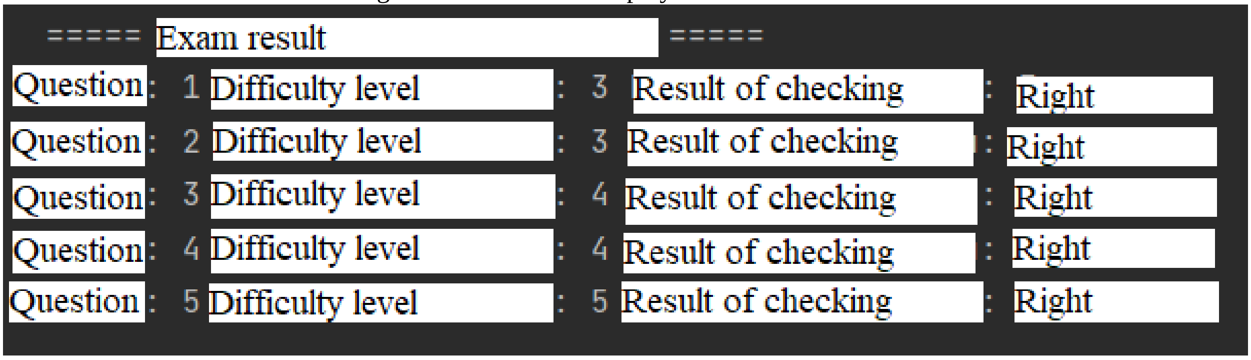


Figure 1. Example of data input from the teacher

Then, at the second stage, the algorithm processes the data and displays questions in the required sequence, focusing on the pyramidal algorithm.

It is worth noting that in this experiment we compiled 15 questions and made 5 levels of difficulty. Looking at Figure 2, we can see that the first question was asked with an average level of difficulty, so as the student answered correctly, the difficulty level of the question gradually increased with each step. At the third stage of the program, it displays the assessment result for the teacher; what data is displayed on the console can be seen in Figure 2.

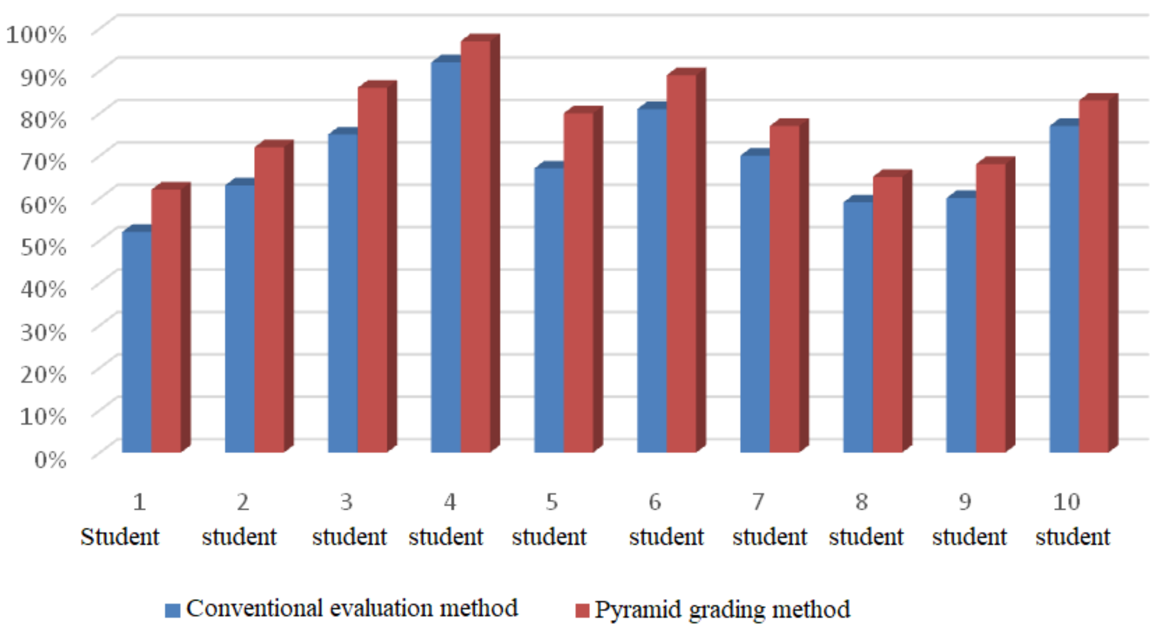
Figure 3. The result is displayed for the teacher.



As we can see, the following information is displayed in the console: what kind of question it is, the level of difficulty of this question, as well as the result of the check. If the answer was correct, then "Correct" is displayed, if the student answered incorrectly, then "Wrong" is displayed accordingly.

An experiment to test the effectiveness of the pyramidal assessment method was conducted on a group of 10 students. The subject for which the assessment will be carried out was chosen as the subject "Computer Networks"; there were a total of 5 levels of difficulty and 15 questions. We can see the result of the comparison with the conventional evaluation method and the pyramidal one in Figure 4 in the form of a diagram; then we will dwell on the advantages of the pyramidal method discovered during the experiment.

Figure 4. Comparison of percentages between conventional and pyramidal scoring methods.



Let's present in the form of a table the results of each student:

	Conventional evaluation method	Pyramid grading method
1 student	52%	62%
2 student	63%	72%
3 student	75%	86%
4 student	92%	97%
5 student	67%	80%
6 student	81%	89%
7 student	70%	77%
8 student	59%	65%
9 student	60%	68%
10 student	77%	83%

Based on this table, we can calculate the average using the usual evaluation method, it is equal to 69.6%. It is worth considering that students had to answer all 15 questions. If we calculate the average for the pyramidal assessment method, it will be equal to 77.9%, while students only had to answer 5 questions with different levels of difficulty. Now having the average indicators of the two evaluation methods, we can also determine the effectiveness of the pyramidal method relative to the usual one, it is equal to 8.3%.

2. Conclusion

In conclusion of this article, we can say that in the context of expanding the range of educational programs and developing new forms of education, the procedure for assessing students and learning outcomes should be carried out in an immediate and highly effective manner. It is argued that multi-stage adaptive measurements lead to increased efficiency of student assessment procedures through the use of IRT models (in particular, the two-parameter model), creating a situation of success for each student by selecting tasks appropriate to his level of preparation, and implementing the principle of individualization. In this article, we described the basic principles on which the program relies when assessing knowledge using the pyramid method. This method aims to reduce the time spent creating a test, as well as taking it. Creating the opportunity for each student to successfully pass knowledge assessment tests by selecting tasks that correspond to his level of preparation and implementing the principle of individualization.

3. References

1. Crocker, L. & Algina, J. (2006). Introduction to Classical and Modern Test Theory. Pacific Grove, CA: Wadsworth, 527 p.
2. Malygin, A. A. (2011). Computer adaptive testing to quality assurance of distance learning. Vestnikuniversiteta, 4, 166–169.
3. Gorbachev, V. T. (2006).Improving the efficiency of adaptive testing the quality of student learning in higher education of a humanitarian profile. PhD Thesis. Moscow: Military University of the Russian Federation Ministry of Defense, 258 p.
4. Zvonnikov, V. I. (2006). Measurement and the quality of education. Moscow: Logos. 312 p.
5. KK Seytnazarov, KI Kalimbetov. The processes of organizing teaching students' algorithms and models /ACADEMICIA: An International Multidisciplinary Research Journal 11 (2), 527-533
6. KK Seitnazarov, KI Kalimbetov. Qat'iymas bo 'lgan ma'lumotlarning ko 'pligi sharoitida otm talabalarining bilimini baholashda qarorlar qabul qilish/ Komputer texnologiyalari 1 (10), 2022.