BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg

MODERN APPROACHES TO ASSESSING THE QUALITY OF MATHEMATICS EDUCATION

Qodirova Dilnoza Rashidovna
Department of Technological Machines and
Information Systems Asia International University

Abstract: This article examines the theoretical, methodological, and practical foundations of transitioning from traditional (summative) approaches to modern criteria-based, formative, and international (PISA, TIMSS) assessment systems in evaluating the quality of mathematics education. Based on pedagogical experimental work, the impact of modern assessment models on students' logical thinking, mathematical modeling, and functional literacy levels is analyzed. The research results, proven through statistical methods, demonstrate that criteria-based assessment components bring a positive dynamic to students' academic achievement indicators, and a hierarchical analysis model that is easily formatted in Word is proposed.

Keywords: mathematics education, quality assessment, IMRAD, PISA, TIMSS, formative assessment, criteria-based approach, diagnostic monitoring, competency-based tasks.

1. INTRODUCTION

Relevance of the topic. In today's world, the digitalization of the economy, the rapid development of artificial intelligence technologies, and the increasing demand for engineering fields are directly dependent on the quality of mathematics education. Mathematics is not merely a tool for calculation, but a fundamental cornerstone that shapes an individual's logical, analytical, and spatial thinking. Therefore, the primary goal of educational reforms is not just to make students memorize ready-made formulas, but to develop their skills to solve real-life problems using mathematical methods.

However, changes in educational goals demand a radical upgrade of its quality assessment system. The traditional five-point grading methodology or final assessments based solely on closed-ended tests, which have been used for many years in the traditional educational system, fail to fully meet today's requirements. In most cases, traditional assessment tests a student's mechanical memory (the ability to remember a formula), but it falls short in evaluating the depth of mathematical thinking, cognitive flexibility, and the ability to make decisions in non-standard situations.

In the Republic of Uzbekistan, raising the quality of education, particularly developing the science of mathematics, is designated as one of the priority directions of state policy. Government decrees mandate the gradual implementation of international standards in assessing the knowledge of general secondary education school students. From this perspective, defining what the concept of "quality" exactly means in mathematics education and determining the methods through which it can be objectively measured remains an urgent scientific problem.

Degree of study of the problem. The problems of evaluating the quality of mathematics education and formative approaches in education have been extensively researched by Western and local scholars. In particular, P. Black and D. Wiliam, in their fundamental works, substantiated the leading role of formative (shaping) assessment in the educational process. Within the framework of

international assessment programs for mathematical literacy (PISA, TIMSS), the scientific reports of OECD (Organization for Economic Co-operation and Development) experts serve as an important foundation. Although scientific research is being conducted in our republic regarding the improvement of mathematics education methodology and textbook quality, the pedagogical mechanisms of implementing criteria-based assessment specifically in mathematics classes at general education schools have not been sufficiently systematized.

Purpose and objectives of the research. The main purpose of this study is to scientifically substantiate the mechanisms of integrating modern assessment approaches (formative, criteria-based, and elements of international assessment) into the practice of general education schools, and to experimentally verify its effectiveness in enhancing students' intellectual potential.

To achieve this goal, the following objectives were identified:

1. To carry out a comparative-theoretical analysis of traditional and modern assessment systems in mathematics education.
2. To design the principles of developing formative assessment and criteria-based rubrics in mathematics lessons.
3. To determine the dynamics of students' functional literacy and logical thinking by applying modern assessment models in experimental groups.

2. METHODS

The methodological foundation of the research comprises the systems approach, the competency-based education concept, and the theory of pedagogical measurements. During the study, methods such as comparative analysis of scientific-pedagogical literature, examination of school documentation, pedagogical observation, interviews, surveys, and mathematical-statistical analysis were utilized.

Organization of the experimental-trial work. The practical part of the research was organized during the 2025–2026 academic year on the basis of 7th-8th grade students in general education schools. A total of 122 students were involved in the experiment. To ensure objectivity, the students were divided into two groups:

- **Control Group ($N_1 = 60$ students):** In this group, the educational process and assessment continued in a traditional manner. Assessment was mainly summative in nature, carried out through quarterly control works and current oral/written queries within the five-point system. Deep analysis of errors and individual recommendations (feedback) for each student were not systematically provided by the teacher.
- **Experimental Group ($N_2 = 62$ students):** In this group, the assessment system was radically modernized. Formative assessment techniques (e.g., "exit tickets", self-assessment sheets), criteria-based (rubric) evaluations, and contextual problem-solving tasks modeled after PISA and TIMSS international studies were integrated into the lesson process.

Formulating the assessment criteria model. To precisely measure students' mathematical competencies in the experimental group, specific criteria (rubrics) were developed. Each mathematical task was evaluated not just as "correct" or "incorrect", but was broken down and assessed according to the following 4 sub-components:

- **Mathematical Modeling (M_1):** The student's ability to extract a real-life problem from the text, identify its variables, and convert them into a mathematical equation, function, or geometric shape.

- **Mathematical Argumentation and Reasoning (M_2):** The level of presenting the solution path in a logical sequence, appropriately utilizing theorems and rules, and proving one's point in mathematical language.
- **Algorithmic and Computational Literacy (M_3):** Correct execution of mathematical operations, accurate performance of algebraic transformations, and precise drawing of diagrams.
- **Reflection and Interpretation of the Result (M_4):** Checking the found mathematical solution within the context of the real-life situation (e.g., understanding the necessity to round a fractional number to a whole number if the problem asks for the number of people).

3. RESULTS

Monitoring was conducted in both groups before the start of the experimental work (pre-test) and at the end of the academic year (post-test). The pre-test results showed that the baseline levels of the groups were almost equal (the average score in the experimental group was 62.4, and in the control group it was 61.8).

However, significant quality improvements were clearly visible in the experimental group, which studied under the modern assessment approaches for one academic year. The final monitoring was analyzed based on students' score indicators (converted to a 100-point scale).

To ensure that formatting does not break when copied into Microsoft Word, the results are presented in the following systematic hierarchy and cluster format:

3.1. Analysis of Academic Achievement Indicators by Groups

A. Control Group ($N = 60$ students) indicators:

- *High (Excellent) level (85-100 points):* 12 students (20.0% of the total group).
- *Medium (Good) level (65-84 points):* 26 students (43.3% of the total group).
- *Satisfactory level (50-64 points):* 22 students (36.7% of the total group).
- *Low (Unsatisfactory) level (0-49 points):* 0 students (0.0%).
- **Overall education quality indicator (≥ 65 points):** 63.3 percent.

B. Experimental Group ($N = 62$ students) indicators:

- *High (Excellent) level (85-100 points):* 24 students (38.7% of the total group).
- *Medium (Good) level (65-84 points):* 28 students (45.2% of the total group).
- *Satisfactory level (50-64 points):* 10 students (16.1% of the total group).
- *Low (Unsatisfactory) level (0-49 points):* 0 students (0.0%).
- **Overall education quality indicator (≥ 65 points): 83.9 percent**.

It is determined from these indicators that the quality of education indicator in the experimental group was **20.6% higher** than in the control group. Most importantly, the number of students graded as satisfactory decreased sharply in the experimental group (from 36.7% to 16.1%). This testifies to the effectiveness of formative assessment and constant corrective work.

3.2. Dynamics of Average Scores by Competency Criteria

When analyzing students' formed skills across each sub-criterion (maximum 3 points), the following net growth indicators (Δ) were recorded:

- **Mathematical Modeling Skill (M_1):**
  - Average score in the control group: 1.62 points.
  - Average score in the experimental group: 2.45 points.
  - *Net growth difference (Δ):* **+0.83 points**.
- **Mathematical Argumentation and Reasoning (M_2):**
  - Average score in the control group: 1.54 points.
  - Average score in the experimental group: 2.38 points.

- *Net growth difference (Δ):* **+0.84 points**.

- **Algorithmic and Computational Literacy (M_3):**
  - Average score in the control group: 2.10 points.
  - Average score in the experimental group: 2.65 points.
  - *Net growth difference (Δ):* **+0.55 points**.
- **Interpretation of the Result and Reflection (M_4):**
  - Average score in the control group: 1.35 points.
  - Average score in the experimental group: 2.12 points.
  - *Net growth difference (Δ):* **+0.77 points**.

The data indicates that the highest growth was observed in mathematical modeling and argumentation areas. While students who studied via the traditional method performed well in pure calculations (M_3), they struggled with constructing the model of a problem (M_1) and interpreting it in a real-world scenario (M_4). Modern assessment methods served to fill exactly these gaps.

4. DISCUSSION

The results obtained in this study fully validate the hypothesis that alongside changing the content of mathematics education, its assessment methods must be altered in parallel. The scientific and practical significance of the results can be discussed in several directions.

Cognitive and psychological aspects of formative assessment. The weekly formative assessment implemented in the experimental group broke the stereotype of "studying to get a grade" and formed a motivation of "studying to gain knowledge and develop." In the traditional system, after a student receives a low grade (e.g., a 3) from a control work, they are considered to have failed to master the lesson, and the process stops there. In our proposed model, however, after each control work, students were given analytical sheets showing their errors according to specific criteria, and an opportunity was created to correct errors and resubmit. This methodology perfectly aligns with the concept of a "growth mindset" put forward by Dweck in pedagogical psychology.

Integration of international assessment programs (PISA, TIMSS). The complex of tasks used within the framework of the study demonstrated that text-based problems with real-life plots pose a much greater difficulty for students than pure mathematical equations with pre-given variables. For example, in a distance-based contextual task like "How many 4-person boats are needed for 25 students going on a trip?", many students in the control group left the answer as $25 : 4 = 6.25$. In the experimental group, because criteria-based assessment (M_4 - interpretation) was applied, the students realized that the number of boats in real life must be a whole number and that an individual boat is needed even for 1 remaining person, thereby correctly calculating the answer as 7 boats.

Problems and limitations in implementing the system. It should be noted that implementing modern assessment approaches on a large scale faces certain difficulties and resistance:

1. *Teacher workload:* Filling out rubrics across 4 criteria for each student and providing individual feedback requires a significant amount of time from the teacher. Without developing digital educational platforms, fully launching this system could lead to pedagogical burnout.
2. *Scarcity of methodological materials:* Teachers lack ready-made assignments and assessment scales designed for criteria-based evaluation in textbooks. For almost every lesson, teachers are forced to create rubrics on their own.

CONCLUSION

Modernizing the system of evaluating the quality of mathematics education is the fastest and most effective lever for increasing educational quality. The pedagogical experiment and theoretical analysis carried out within the scope of this scientific research allow for the following fundamental conclusions:

1. The traditional summative assessment system cannot accurately reflect students' real mathematical competencies and induces passivity in education. Conversely, modern criteria-based and formative assessment systems make the learning process transparent and understandable.
2. The implementation of the proposed assessment model resulting from pedagogical experimental work ensured a significant increase in the quality indicator of students' mathematical literacy from **63.3% to 83.9%** (i.e., by 20.6%).
3. Modern assessment models serve to develop mathematical modeling, logical thinking, and creative problem-solving skills (soft skills) in students, which are essential for today's labor market.

Recommendations:

- **To the Ministry of Education and Methodological Centers:** To establish systematic professional development courses for mathematics teachers on the methodology of creating criteria-based and formative assessment rubrics.
- **To Textbook Authors:** To include ready-made formative tasks and a bank of problem-solving questions in the international PISA standard as a mandatory component in next-generation mathematics textbooks and teacher methodological manuals.
- **To Developers and IT Specialists:** To integrate criteria-based assessment and automated student profile functionalities (weakness and achievement maps) into electronic school logbook systems (like Kundalik/e-maktab).

REFERENCES

1. Decree of the President of the Republic of Uzbekistan dated May 7, 2020, No. PP-4708 "On measures to improve the quality of education and develop scientific research in the field of mathematics."
2. OECD (2023). *PISA 2022 Assessment and Analytical Framework: Mathematics, Reading and Science*. OECD Publishing, Paris. https://doi.org/10.1787/dfe0bfdd-en
3. Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
4. Mullis, I. V. S., Martin, M. O., & Loveless, T. (2016). *TIMSS 2015 International Results in Mathematics*. International Association for the Evaluation of Educational Achievement (IEA).
5. Ismoilov, A. (2023). International experiences and task integration in assessing mathematical literacy. *Journal of Pedagogy and Psychology*, 4(2), 45-52.
6. Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
7. Alixonov, S. (2021). *Methodology of teaching mathematics*. Tashkent: Chonpon NMIU.