# SPECIFICATION OF THE SAM MATH TEST

## 1. OBJECTIVE

The test is intended for the assessment of subject competencies of primary school students to evaluate the level of assimilation of the *Mathematics* subject content. The test model is based on the theory of cultural development (L. S. Vygotsky, V. V. Davydov, D.B. Elkonin et al.), and assumes assessment of students' competencies at three basic levels: formal, reflective and functional.

## 2. TARGET AUDIENCE

The test is targeted for primary school graduates and can be performed by grade 4 and 5 students.

## 3. CONTENT

The test includes the main math areas from the primary school curricula.

The content was selected in accordance with Federal State Standard of Primary Education (MOED Order # 373 of October 6, 2009 *On Approving and Enactment of the Federal State Standard of Primary Education*).

Subject content of the test is divided in five areas.

**"Numbers and Calculations".** This content area is relevant to the formal aspect of the concept of natural numbers (positional representation of numbers, standard algorithms of operations with numbers, order of operations, properties of operations). It alsoincludes materials related to representation of numbers on the coordinate line. The latter is important for the understanding of real numbers and assimilation of coordinate method.

**"Measurement of Values".** This content area includes materials related to direct and indirect measurement operations, and also incorporates geometric measurements.

As to the applied aspect of this content area closely related to specific practical measurements and their representation as diagrams and charts (*data analysis*), it can rather be relevant to the *Outside World* subject.

**"Regularities".** This content area is related to construction of numerical and geometric sequences and other structured objects, and measurement of their quantitative parameters. This area is highly important for the development of mathematical thinking(first of all, algorithmic and combinatory).

**"Dependences".** This content area is related to identification and description of the mathematical structure of relations between values usually represented in test items.

**"Elements of Geometry".** This content area covers geometric materials related to identification of spatial forms and relative position of objects.

The content framework of the math test can be presented as a matrix (Table 3.1.) that includes:

■ Subject content areas (6 areas);

■ Mathematical tools (concepts, principles, formulas, algorithms, etc.) providing the orientation for mathematical operations.

***Table 1. Content of the math test***

| **Content areas** | **Orientation tools for mathematical operations** |
|---|---|
| **Numbers** | ■ *Sequence of natural numbers* |

| | |
|---|---|
| **and Calculations** | ■ *Number line*<br>■ *Positional principle*<br>■ *Properties of arithmetic operations*<br>■ *Order of operations* |
| **Measurement of Values** | ■ *Relationship between the number, value and unit*<br>■ *Whole-part relationship*<br>■ *Formula of rectangle area* |
| **Regularities** | ■ *"Induction step"*<br>■ *Recurrence (periodicity)* |
| **Dependences** | ■ *Relationship between like values (equality, inequality, multiplicity,difference, "whole-part")*<br>■ *Direct proportion between values*<br>■ *Derived values: velocity, labor productivity, etc.*<br>■ *Relationship between units* |
| **Elements of Geometry** | ■ *Form and other properties of figures (main types of geometrical figures)*<br>■ *Spatial relationship between figures*<br>■ *Symmetry* |

## 4. PRINCIPLES OF TEST DESIGN

The test was developed using two approaches: norm-referenced and criterion-referenced ones — combined in accordance with the modern test theory — Item Response Theory (hereinafter — IRT), which also allowed a criterion-referenced interpretation of the test score scale.

Norm-referenced approach enables to compare the performance of a student with that of other test takers, as well as with his/her previous performance. To this end, each test taker is assigned an integral test score obtained as a result of mathematical treatment of test results. Test scores of all test takers are shown on a common scale, irrespective of the time of testing and specific set of items done.

The second (criterion-referenced) approach allows a qualitative evaluation of the content area assimilation through indicating the type of orientation in problem solving. To this end, a graduated scale of achievements was developed based on integral scores of test takers and benchmarks classifying all test takers in groups corresponding to different qualitative levels of achievement.

## 5. TEST STRUCTURE

The test consists of three level items grouped in blocks. In total, there are 45 items and 15 blocks. The test can be viewed as a system of three subtests with each representing a set of items of the same level referring to different content areas. The overall structure of the test is shown in Figure 1.


***Figure 1. Test structure***

**6. ARRANGEMENT OF TEST ITEMS**

A block consisting of three items (levels 1, 2 and 3) corresponding to a subject content area is a structural unit of the test. Items are presented in blocks. The sequence of blocks makes no difference.

**7. RATIO BETWEEN ITEMS REFERRING TO DIFFERENT CONTENT AREAS AND LEVELS**

***Table 2. Tentative ratio between different items in the test book***

| **Content areas** | **Number of blocks** | **Number of items** |
|---|---|---|
| **Numbers and Calculations** | *4* | *12* |
| **Measurement of Values** | *5* | *15* |
| **Regularities** | *2* | *6* |
| **Dependences between values** | *2* | *6* |
| **Elements of Geometry** | *2* | *6* |

| | | |
|---|---|---|
| **TOTAL** | *15* | *45* |

**8. TYPES OF ITEMS**

The following types of items are used in the test:

■ completion items with a brief answer;

■ multiple-choice items with a choice from 4-5 offered options;

■ items requiring constructions.

The majority of items (about 80%) are the completion ones with a brief answer.

**9. NUMBER OF VERSIONS**

4 versions of the test with similar statistical parameters were developed. All versions include common items enabling to perform equating and plot the scores of all test takers on a single scale. The number of common items in different versions makes up at least 6 (at least two common blocks).

**10. SCORING OF ITEMS**

Scoring procedure uses a dichotomous approach: students get 1 point for the correct answer and 0 for incorrect (or absent) answer. Therefore, the highest raw score that the test taker can achieve for completing the test equals 45. The highest raw score that each test taker can get for each level equals 15.

**11. RECOMMENDED DURATION OF TEST TAKING**

Recommended duration of test taking is 90 minutes (two 45 minute lessons with a break). Tentative time of performing level 1 items – 1 minute, level 2 – 2 minutes, level 3 – 3 minutes. Testing can be conducted during two days: a half of the test (blocks 1-8) to be completed on the first day, and the second half (blocks 8-15) – on the second day.

**12. FORMS OF TEST TAKING**

The test can be offered in paper-based and computer-based forms. It is assumed that testing data do not depend on the form of test taking but this issue is still being studied.

**13. CONDITIONS OF TEST TAKING**

Testing can be performed by a primary school teacher (if the testing takes place in grade 4) or math teacher (is testing takes place in grade 5). The teacher in charge of testing should help the test takers to complete identification information on the front of the test booklet, explain obscure passages in the instructions, keep track of the time and maintain order in the class during the testing.

**14. PROCESSING OF TEST RESULTS**

The SAM toolkit includes an automated information system (computer module) intended to estimate test takers and automatically generate various reports (tables, charts and diagrams) both on each test taker, and different classes and schools. At his/her wish, the teacher can carry out an independent scoring (manually check the items using the keys). However, computer module would not only simplify the task but also enable to utilize the whole range of reports generated by the system to analyze the test results. In addition, the module can accumulate the data and take account of previous testing data when generating reports on subsequent testing thus allowing a comparison of results.

**15. ESTIMATION OF TEST TAKERS**

Estimation is performed using a special measuring technique based on the modern test theory IRT. The following parameters are estimated for each student:

■ Integral test score placed on the scale common for all test takers irrespective of the time of testing and specific set of items they have done. Integral test scores are presented on a 1000-point scale;

■ One of the four achievement grades:

■ **ZERO LEVEL:** even the first (formal) level has not been assimilated;

■ **FIRST LEVEL:** only the first (formal) level has been assimilated;

■ **SECOND LEVEL :** the second (reflective) level has been assimilated;

■ **THIRD LEVEL:** the third (functional) level has been assimilated.

**16. PRESENTATION OF TEST RESULTS**

For the purpose of test results presentation, the system generates two types of tables:

■ Aggregate data on schools and classes;

■ Individual data on test takers.

Automated information system allows presentation of estimation data in the form of charts and diagrams showing:

■ Distribution of students by proficiency levels;

■ Subject success of classes (profiles).

The above tables and illustrations can be supplemented with data obtained from surveys among the test takers, which enables to track the influence of various factors on the quality of students' learning achievements.