Preliminary Data from Mathematics State Assessment for Year 2011 at for Fourth Grade in Bulgaria

Milen Zamfirov

St. Kliment Ohridski University of Sofia, Faculty of Preschool and Primary School Education, Bulgaria

milen_zamphirov@abv.bg

Abstract. This report presents results analysis from the state assessment of students' academic achievements in Mathematics at fourth grade. The assessment is realized in May 2011 with representative group of students (12 000 students). The number of participants in Mathematics state assessment from fourth class is 61 408, they are educated in 3 381 classes at 1 886 schools, in 262 municipalities in all of the 28 Bulgarian regions.

Description.

The report includes:

• Analyze the psychometric characteristics of tests in mathematics;

• General analysis of student performance;

• Analysis of differences in student performance according to their different characteristics [1].

The overall assessment in Mathematics consists of the aggregation of points – 20 choice tasks (3 alternatives).

The results of external evaluation of the 4th grade are inherently basic. The resulting differences in regions, municipalities, sex and the most frequently spoken language in the family are finding nature.

• Average score for Mathematics in the country - a criterion for ranking the regions at this stage.

• Average score for Mathematics in the region - a criterion for ranking the municipalities.

• Average score for Math group, speaking most frequently in a specific language in the family - guide to search for best practices in teaching mathematics in a multilingual environment.

The winners of best practices have the means and methods for determination of problem areas, means and methods to influence them and the means and methods to measure and stabilize the effect of its effects [2]. The achievements of the group speaking a language family to determine their own group. To enter a school or class in the group of potential winners of best practices, it must have:

1) 5 or more children in each group speaking Bulgarian, Turkish and Roma in the class.

2) Results *above average* for one of the groups should be accompanied by a *medium* and/or *above average performance* in other groups.
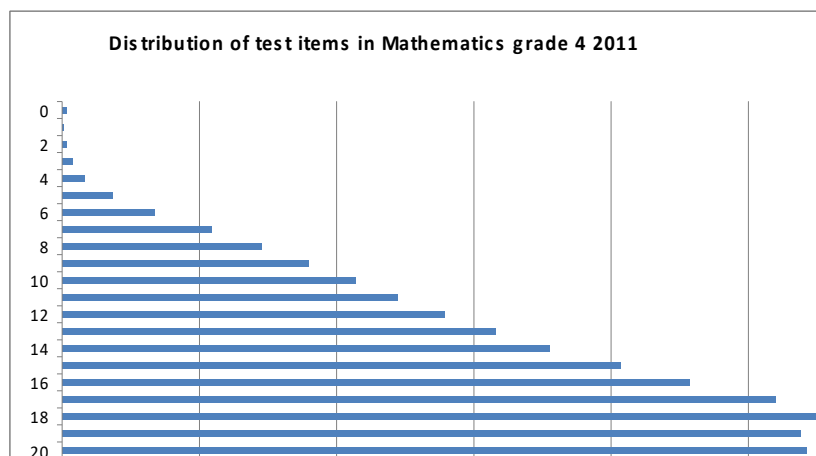
Psychometric Performance of the Test

The main parameters of psychometric tests are listed in Table 1 and Table 2. The results show very good reliability of the test for practical purposes (Alpha = 0.82). It allows you to perform comparisons of results by groups (gender, age, region, etc.). The standard error of measurement indicates how many points have to be between two test result to be statistically significantly different. In our case it is 1.68 points.

Number of people	61408	Number of questions	20
Minimum test score	0	Maximum test score	20
Average test score	14,97	Standard deviation	3,93
Average difficulty	74,85	Standard error of measurement	1,68

Alpha reliability	0,82	Standard error of Alpha		0
Reliability GLB (various questions)		0,84		

Table. 1 *Basic psychometric indicators of tests*

40 students or 0.07% have not decided any question.

The expected number of correctly solved questions using the strategy of guess is 7 (6.67) questions. Same and lower score showed 4.89% of the students (2998).

Basic psychometric performance of the test items

The psychometric data on individual questions are listed in the following table:

Question	Key	Do not reply%	Difficulty	$D=\alpha_{bis}$	α_{19}	A	Б	B
Answer 01	б	0	88	0,37	0,81	-0,24	*	-0,26
Answer 02	в	0	86	0,32	0,81	-0,17	-0,24	*
Answer 03	б	0	93	0,34	0,81	-0,20	*	-0,25
Answer 04	a	0	81	0,41	0,81	*	-0,23	-0,30
Answer 05	в	1	85	0,44	0,81	-0,30	-0,28	*
Answer 06	a	1	84	0,47	0,81	*	-0,29	-0,31
Answer 07	в	0	88	0,41	0,81	-0,24	-0,29	*
Answer 08	б	1	80	0,41	0,81	-0,23	*	-0,31
Answer 09	б	1	90	0,41	0,81	-0,28	*	-0,25
Answer 10	б	1	72	0,45	0,81	-0,23	*	-0,32
Answer 11	в	1	69	0,49	0,80	-0,29	-0,30	*
Answer 12	б	1	80	0,50	0,81	-0,32	*	-0,30
Answer 13	a	1	75	0,20	0,82	*	-0,10	-0,18
Answer 14	в	1	64	0,37	0,81	-0,20	-0,23	*
Answer 15	в	1	53	0,16	0,82	-0,09	-0,08	*
Answer 16	в	1	46	0,34	0,81	-0,23	-0,11	*
Answer 17	б	2	65	0,43	0,81	-0,25	*	-0,25
Answer 18	a	1	75	0,48	0,81	*	-0,37	-0,20
Answer 19	a	2	68	0,44	0,81	*	-0,29	-0,25
Answer 20	в	2	56	0,45	0,81	-0,29	-0,19	*

Table 2 *Basic psychometric characteristics of questions in the test*

Considering these two models we can correctly predict:

- ➢ 63% results in the group *below average*
- ➢ 36% average performance
- ➢ 84% results in the group *above average*.

Compare the regions with average results for country

Region	Number of Students	Points	Ranking

Smolyan	918	15,84	Above Average
Sofia - City	9387	15,3	
Blagoevgrad	2853	14,89	
Gabrovo	891	14,8	
Kardzhali	1395	14,66	
Yambol	1242	14,64	
Plovdiv	5635	14,62	
Pernik	948	14,58	Average
Varna	4031	14,43	
Lovech	1244	14,4	
Ruse	1752	14,33	
Kyustendil	1018	14,24	
Pazardzhik	2500	14,15	
Sliven	1967	14,05	
Shumen	1705	13,87	
Stara Zagora	3050	13,81	
Sofia – region	2114	13,81	
Silistra	1014	13,73	Below Average
Pleven	2255	13,71	
Veliko Tarnovo	1888	13,7	
Vratsa	1654	13,61	
Haskovo	2063	13,53	
Burgas	3803	13,51	
Dobrich	1682	13,36	
Vidin	801	13,31	
Targovishte	1080	13,06	
Montana	1284	12,76	
Razgrad	1163	12,63	

Table 3 *Ranking of the regions according to the average score of the country*

The results by a region compared to the national average were arranged in 3 groups: average, below average and above average. The group below average is the lowest number of points and covers 33.3% of students with lower scores. It is marked with dark gray in the table. The group above average includes students with the highest number of points. It also covers 33.3% of students (but those in the upper third of the distribution). The group is above average is marked by light gray in the table. To compare the results of a region, municipality or school with average results for the country using the following scale: up 13.73 points - performance below average for the country, 13.81 to 14.58 points – average performances; 14.62 – 20 points - performances above the national average for the country.

The scale transformation of the points in the third stage assessment for schools is given in the following table do not aim to compare the average results for the country but only with the school of its own region.

	Result to the Region		
	Below Average	**Average**	**Above Average**
Blagoevgrad	0-14	14,01 - 18	18,01 - 20
Burgas	0-13	13,01 - 18	18,01 - 20
Varna	0-15	15,01 - 18	18,01 - 20

Veliko Tarnovo	0-13	13,01 - 17	17,01 - 20
Vidin	0-14	14,01 - 17	17,01 - 20
Vratsa	0-13	13,01 - 17	17,01 - 20
Gabrovo	0-14	14,01 - 18	18,01 - 20
Dobrich	0-13	13,01 - 17	17,01 - 20
Kardzhali	0-14	14,01 - 17	17,01 - 20
Kyustendil	0-14	14,01 - 18	18,01 - 20
Lovech	0-14	14,01 - 18	18,01 - 20
Montana	0-12	12,01 - 17	17,01 - 20
Pazardzhik	0-13	13,01 - 17	17,01 - 20
Pernik	0-14	14,01 - 17	17,01 - 20
Pleven	0-13	13,01 - 17	17,01 - 20
Plovdiv	0-14	14,01 - 18	18,01 - 20
Razgrad	0-12	12,01 - 16	16,01 - 20
Ruse	0-14	14,01 - 17	17,01 - 20
Silistra	0-13	13,01 - 17	17,01 - 20
Sliven	0-13	13,01 - 17	17,01 - 20
Smolyan	0-15	15,01 - 18	18,01 - 20
Sofia – City	0-15	15,01 - 18	18,01 - 20
Sofia – region	0-13	13,01 - 17	17,01 - 20
Stara Zagora	0-13	13,01 - 17	17,01 - 20
Targovishte	0-12	12,01 - 16	16,01 - 20
Haskovo	0-13	13,01 - 17	17,01 - 20
Shumen	0-13	13,01 - 17	17,01 - 20
Yambol	0-14	14,01 - 18	18,01 - 20

Tabl. 4 *Normalizing the results separately for each region*

The lowest (tolerant) threshold for passage in the area of average achievements in the region of Montana, Razgrad, Targovishte – 12.01 points.

The highest (strict) threshold for passage in the area of average achievements in regions Varna, Smolyan, and Sofia City – 15.01 points.

The lowest (tolerant) threshold for passage in the area of above-average performance is in the region of Razgrad and Targovishte – 16.01 points.

The highest (strict) threshold for passage in the area of above – average performance in Regions Varna, Smolyan, Sofia - City, Blagoevgrad, Gabrovo, Kyustendil, Lovech, Plovdiv, Yambol, and Burgas – 18.01 points.

The influence of gender and the most widely spoken language in the family on the psychometric characteristics of test.

Effect of gender

Group	Number of persons	Min. Test Result	Average Test Result	Average Difficulty	Coefficient Alpha	Standard error of measurement
Total	61408	0	14,97	74,85	0,82	1,68
Boys	31347	0	14,8	74	0,82	1,69

| Girls | 30061 | 0 | 15,15 | 75,74 | 0,82 | 1,66 |

Tabl. 5 *Psychometric characteristics of the test group of boys and girls*

The difference in average scores between boys and girls is less than 0.5 points. It is within the standard error of measurement.

Check the influence of gender by region showed no statistically significant differences.

Influence of the most commonly spoken language in the family

Group	Number of persons	Min. Test Result	Average Test Result	Average Difficulty	Coefficient Alpha	Standard error of measurement
Total	61408	0	14,97	74,85	0,82	1,68
Bulgarian	48750	0	15,65	78,23	0,81	1,59
Turkish	7128	0	12,98	64,88	0,77	1,89
Gypsy	5320	0	11,49	57,47	0,73	1,99
Other	210	0	13,91	69,57	0,76	1,82

Tabl. 6 *Psychometric characteristics of the test according to the most commonly spoken language in the family*

The table shows that the test is not difficult for a group speaking Bulgarian - average difficulty is over 78%. For other groups the difficulty of the test is within the recommended 20% – 80%.

Conclusion

Since the percentage of correct answers can draw the following conclusions:

• The average test score is high – 14.97 points, indicating that students cope successfully with 75% of the inspected material.

• The median is 16.0 points or 50% of students have more than 18 points (90% of the inspected curriculum).

• Fashion (the ball with the highest incidence – 6662 students – 10,85%) was 20 points and coincides with the maximum possible score.

• The test is not too difficult for either group (mean number of correct answers below 20%).

• The test is not difficult for the group, speaking Bulgarian. It is mastered the average 78.23% of the inspected material.

• Group, speaking fluent Turkish wields average 64.88% of the inspected material

• The group speaks fluent Gypsy wields average 57.47% of the inspected material

• Considering the standard error of measurement, we can conclude that there is a statistically significant difference between groups speaking Bulgarian and Gypsy (4.16 points) and between the groups speaking Bulgarian and Turkish (2.67 points).

• Considering the standard error of measurement can conclude that there is a statistically significant difference.

• Among other groups there are differences within the standard error of measurement that is not statistically significant.

Guided by the heuristic rule adopted in psychology that typical of a group is what 50% of its members can do we conclude that:

•All questions are of sufficiently good discriminatory power, except question 15, which is admissible (above the recommended threshold above 0.15) discriminatory power.

•Two issues have 90 percent or more correct answers. This is № 3 (93%) and № 9 (90%).

•Average difficult are 3 questions - № 16 (46%); № 15 (53%) and № 20 (56%).

•Seven tasks were determined (№ 4, 1, 2, 10, 5, 9 and 13), which can predict the true outcome of the entire test with 60.4% accuracy. They constitute the core of the models (full and partial) that students used to cope with a test.

References

1. М. **Замфиров**, Статистически анализ на степента на усвояемост на учебен материал в културнообразователната област Природни науки и екология 5-8 клас от ученици със слухови нарушения в специализирани училища в България и Унгария, *Специална педагогика* (4) (2006).
2. М. **Замфиров**, С. **Съева**, Оценка на взаимовръзката между променливи при апробация на нови жестове за деца с увреден слух по Човекът и природата в 6 клас.