Diagnostic Testing of Learners within the Approbation of the Updated Content of Primary Education in Kazakhstan (Longitudinal Study Results)

COMPARATIVE ANALYSIS OF THE RESULTS
OF A FOUR-YEAR RESEARCH

Olga Mozhayeva (mozhayeva_o@nis.edu.kz)
Aidana Shilibekova (shilibekova_a@cpi.nis.edu.kz)
Dinara Ziyedenova (ziyedenova_d@nis.edu.kz)

DRAFT

Diagnostic Testing

Purpose of the diagnostic testing is to measure and compare the progress made by learners in the learning process.

Pilot schools (30)
Learning was carried out according to the updated content of education (SCSPE* 2015)

Control schools (16)
Learning was carried out on SCSPE* of 2012

Pilot		Control
4	**School-lyceums**	1
8	**School-gymnasiums**	4
13	**Secondary schools**	10
5	**Small-scale schools**	1
18	**Urban schools**	8
12	**Rural schools**	8

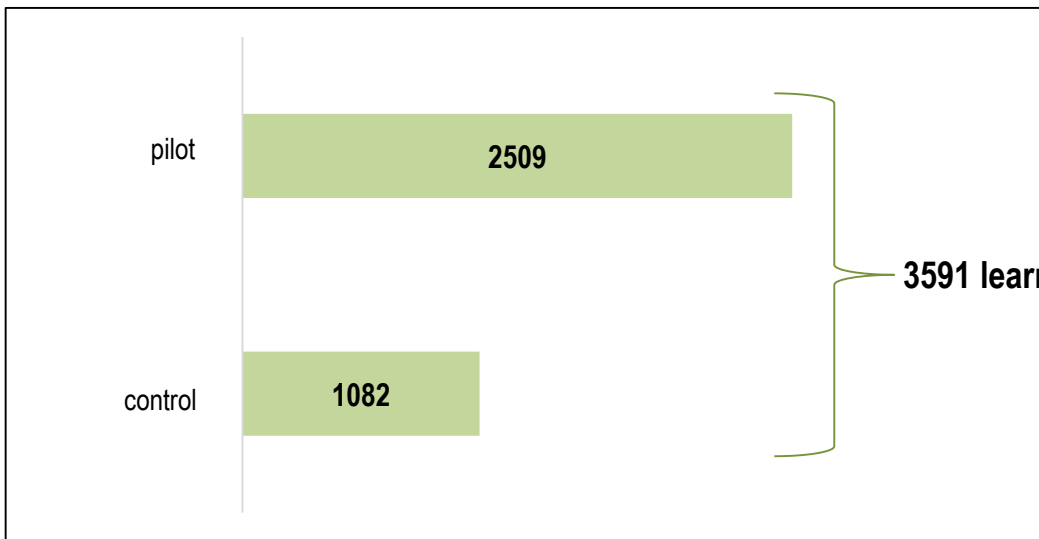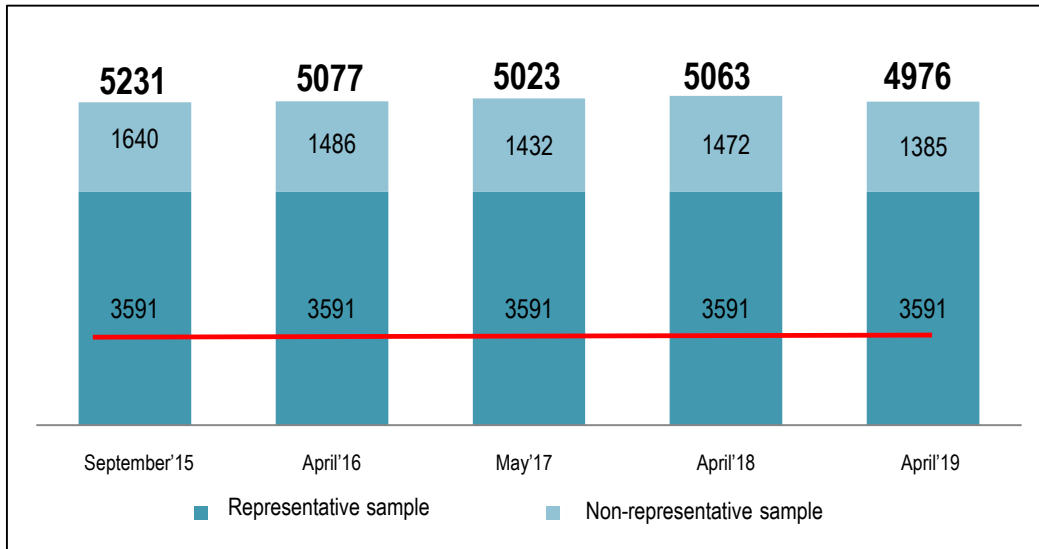*SCSPE - State Compulsory Standard of Primary Education*

STAGES AND CONTINGENT OF DIAGNOSTIC TESTING

Diagnostic testing is a longitudinal study that included 5 stages and lasted for 4 years from September 2015 to April 2019.

September 2015 → Grade 1 "at the entrance"	April 2016 → Grade 1 "at the exit"	May 2017 → Grade 2 "at the exit"	April 2018 → Grade 3 "at the exit"	April 2019 → Grade 4 "at the exit"
Pilot schools – 3420 learners	Pilot schools – 3347 learners	Pilot schools – 3390 learners	Pilot schools – 3378 learners	Pilot schools – 3301 learners
Control schools – 1811 learners	Control schools – 1730 learners	Control schools – 1633 learners	Control schools – 1685 learners	Control schools – 1675 learners
Total – 5231 learners	**Total – 5077 learners**	**Total – 5023 learners**	**Total – 5063 learners**	**Total – 4976 learners**

LEARNERS' CONTINGENT FOR COMPARATIVE ANALYSIS

5231	**5077**	**5023**	**5063**	**4976**
1640	1486	1432	1472	1385
3591	3591	3591	3591	3591
September'15	April'16	May'17	April'18	April'19

■ Representative sample ■ Non-representative sample

pilot **2509**

control **1082**

3591 learners

Comparative analysis was carried out on the basis of a sample of 3591 learners who took part in all five diagnostic tests

STRUCTURE AND CONTENT OF DIAGNOSTIC TESTING

Test sections

- Mathematics
- First language
  Kazakh / Russian / Uygur
- Natural science and World studies

Types of task used in the diagnostic testing:

- ❖ closed-ended type with the choice of one correct answer;
- ❖ closed-ended type with the choice of several correct answer;
- ❖ open-ended type that require a short answer;
- ❖ open-ended type that require an extended answer.

| Test section	Example of sub-section	Example of assessment criteria
Mathematics	Natural numbers and number 0. Fractions	Writes numbers as a bit composition
First language	Defining text types and styles	Defines text types
Natural science and World studies	Plants	Defines plants' life cycle

METHODOLOGY

❖ To compare the results, *the statistical processing of data based on the modern theory of test tasks (IRT)*, which allows to measure learners' progress level, was undertaken.

❖ In processing, *a one-parameter logistics model* (OPLM) is used to calculate the parameters of the test tasks, which allows to calculate the ability score of learners based on the discriminative properties of the test tasks and their level of complexity.

❖ *The achievements of each learner are correlated with ability scale,* which allows to calculate the level of their abilities, measure progress in learning, track the trajectory of success throughout the period of study.
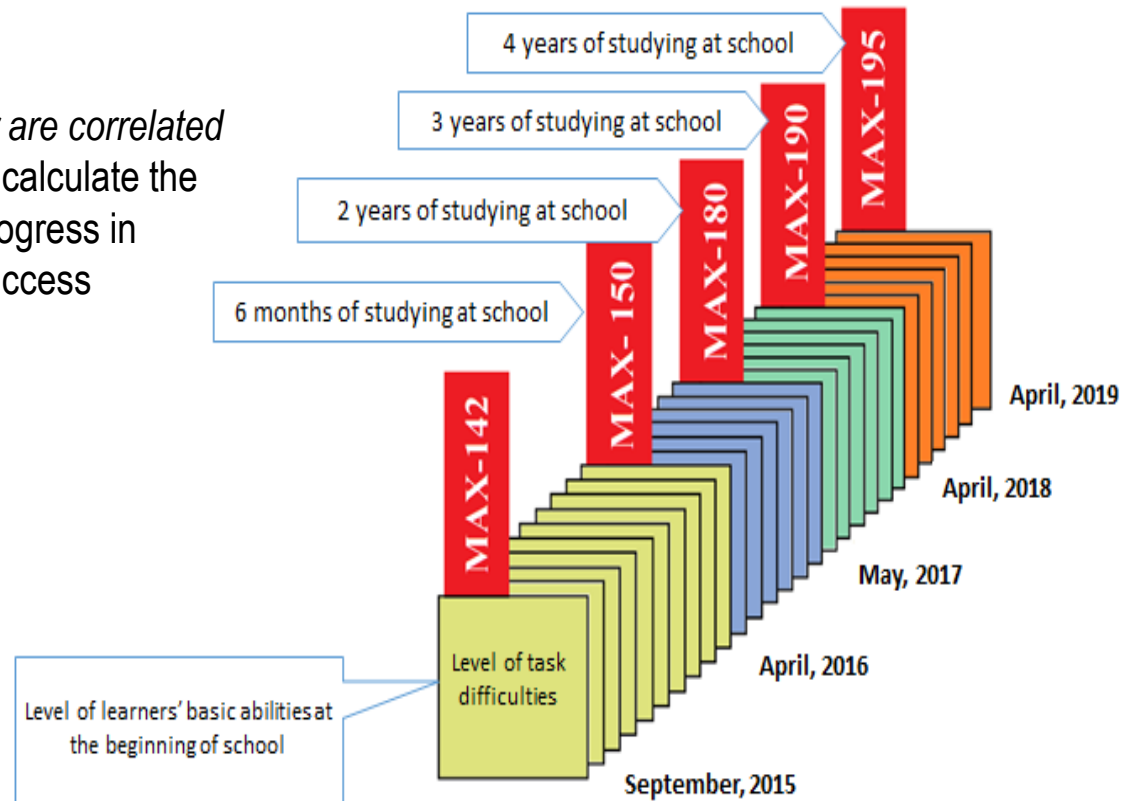
PROGRESS OF AVERAGE SCORES

A comparative data analysis of five diagnostic tests shows an increase in the average scores of learners on ability scale of pilot and control schools.
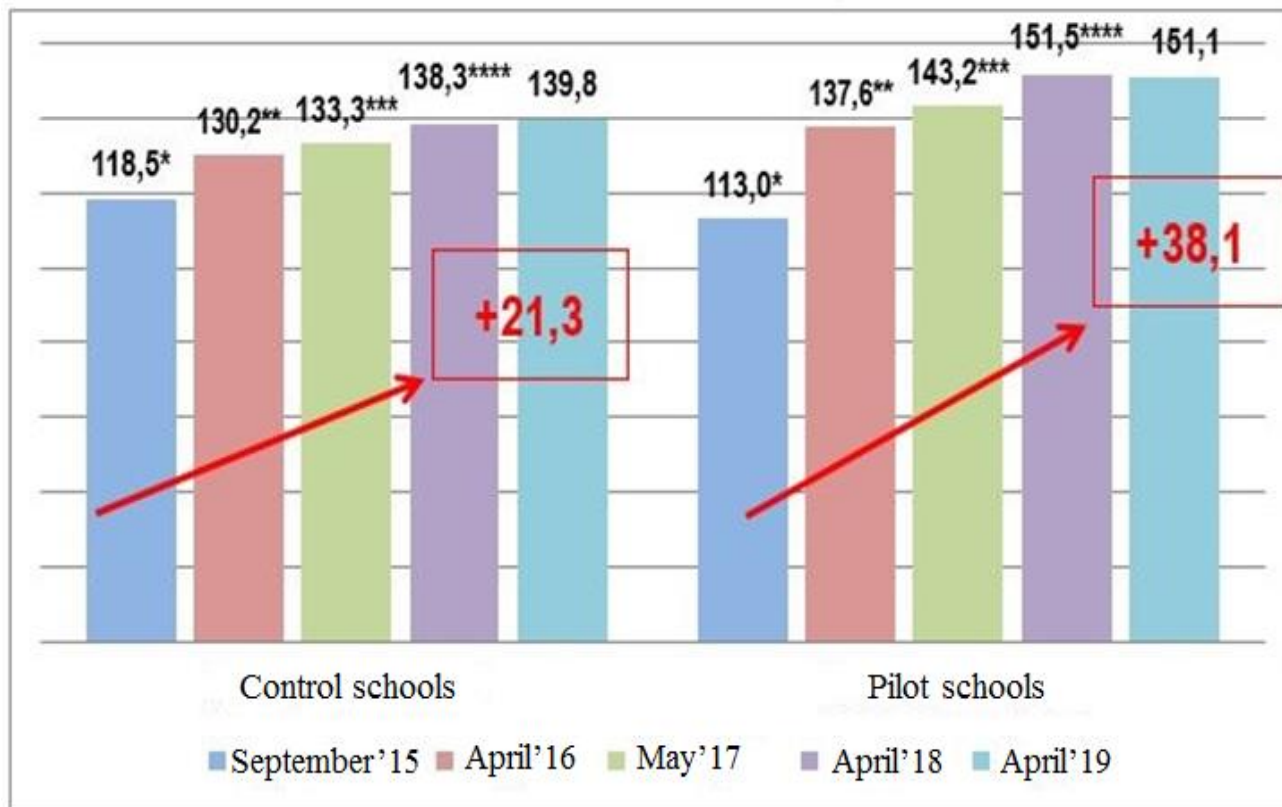
Pilot schools have more significant progress in the average score (+38.1 scores), which almost doubles the performance of control schools (+21.3 scores)

*Changes in samples lead to slight deviations in the results of previous reports.*

* Data for 2015: control schools – 118,5 scores, pilot schools – 112,6 scores
** Data for 2016: control schools – 130,4 scores, pilot schools – 137,4 scores
*** Data for 2017: control schools – 133,1 scores, pilot schools – 143,0 scores DRAFT
**** Data for 2018: control schools – 138,0 scores, pilot schools – 151,4 scores

# DIFFERENCE BETWEEN AVERAGE SCORES OF CONTROL AND PILOT SCHOOLS
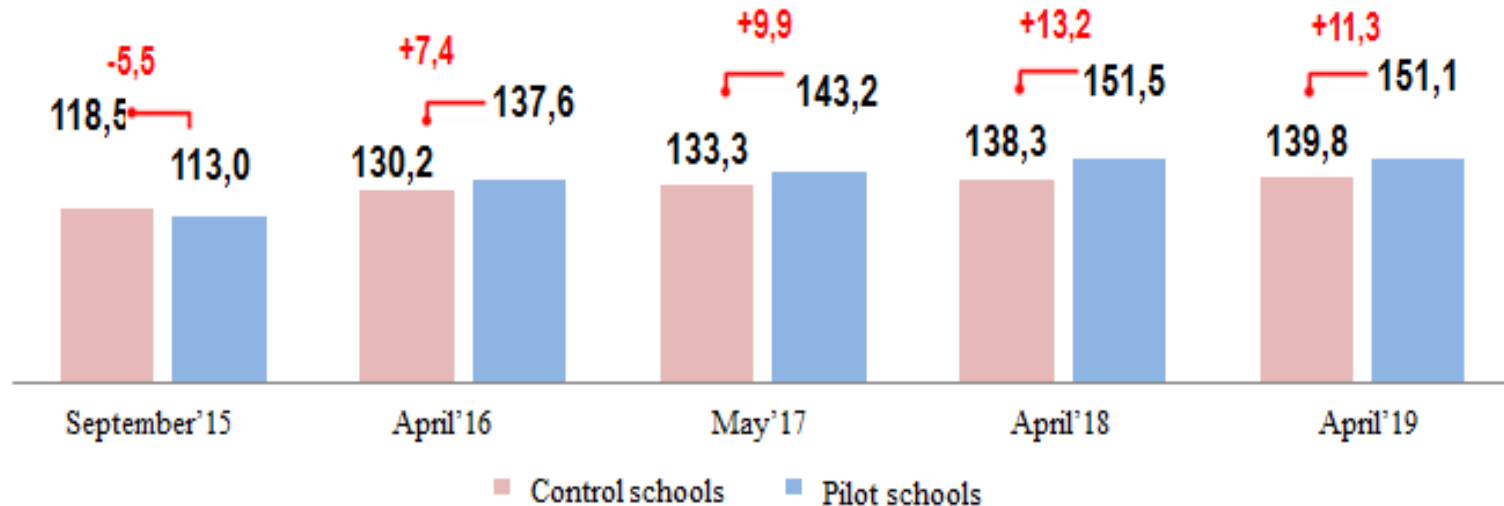
❖ Performance of pilot schools "at the entrance" test were lower than control schools by 5.5 scores.

❖ From the results of subsequent tests, there is an increase in the difference of average scores in favor of pilot schools (+7.4; +9.9; +13.2, +11.3).

In general, the analysis of the results showed a comparative advantage in all tracked indicators and increase of pilot schools' average scores on the ability scale.

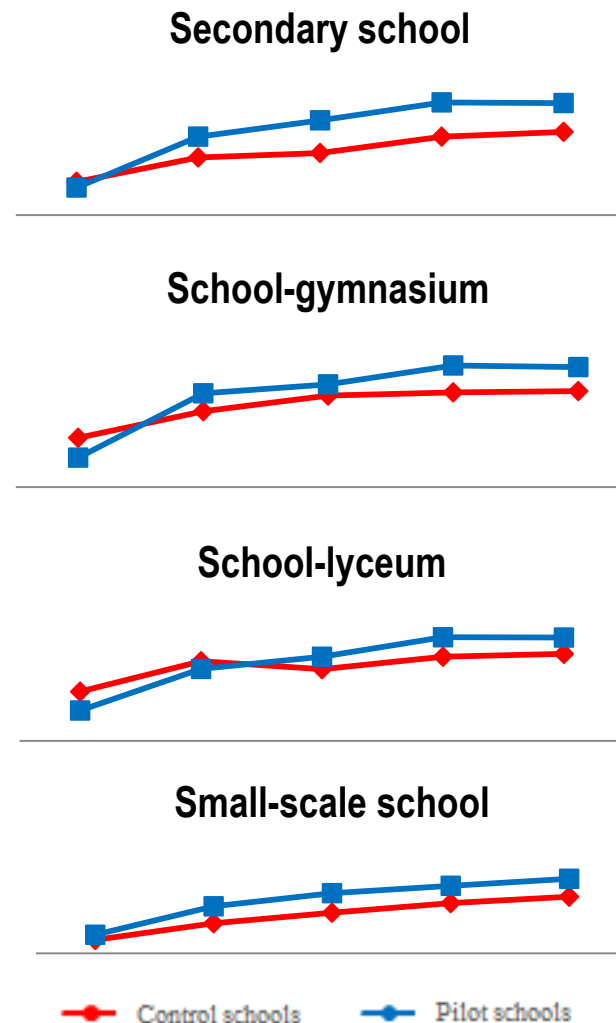# PROGRESS OF AVERAGE SCORES BY LANGUAGE OF INSTRUCTION

	September' 15	April' 16	May' 17	April' 18	April' 19	Difference September' 15 / April' 19
Kazakh						
control	118,6	127,7	132,7	136,9	140,0	+21,3
pilot	113,2	138,5	143,1	151,5	151,6	+38,4
Russian						
control	119,1	133,2	137,3	141,4	140,7	+21,6
pilot	112,3	136,7	143,5	151,4	150,3	+38,0
Uygur						
control	115,5	130,2	120,8	133,2	135,4	+19,9
pilot	125,8	131,6	138,1	156,5	160,3	+34,6

Kazakh

Russian

Uygur

Control schools ——— Pilot schools
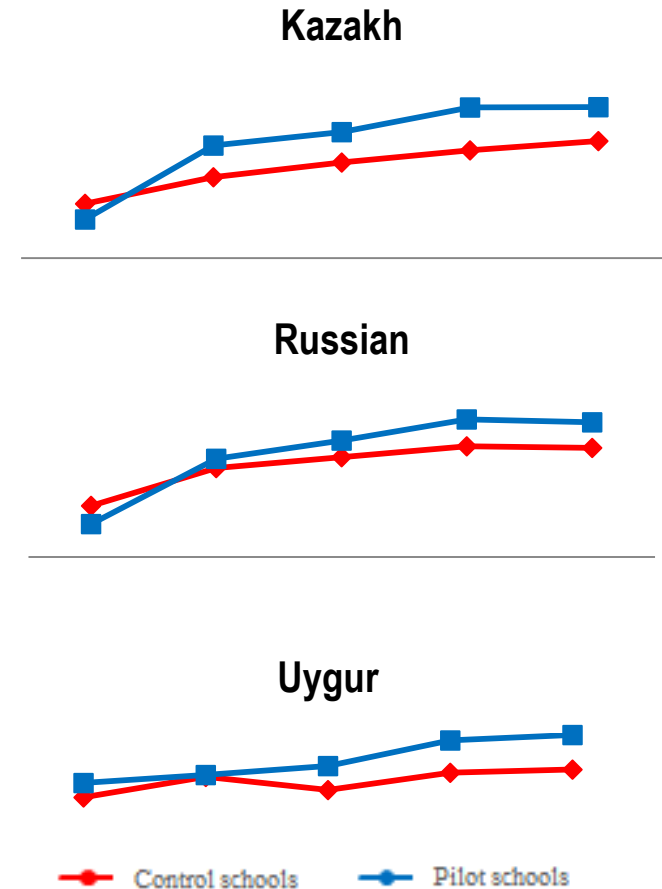
- ❖ The range of differences in progress indicators of pilot schools varies from 5.1 to 13.5 scores (within 8.4 scores), while in control schools it varies from 5.3 to 16.5 scores (within 11.2 scores).
- ❖ There is a less marked difference in the pilot sample.

PROGRESS OF AVERAGE SCORES BY SCHOOL TYPES

	September' 15	April' 16	May' 17	April' 18	April' 19	Difference September' 15 / April' 19
Secondary school						
control	115,4	126,6	128,7	136,2	138,4	+23,0
pilot	112,7	136,3	143,8	152,1	151,8	+39,0
School-gymnasium						
control	120,9	132,2	138,8	140,2	140,7	+19,8
pilot	112,5	139,8	143,5	151,5	150,9	+38,4
School-lyceum						
control	124,3	139,1	135,2	141,4	142,8	+18,5
pilot	115,0	135,4	141,4	151,0	150,8	+35,8
Small-scale school						
control	108,5	119,1	125,9	131,9	136,1	+27,6
pilot	111,7	129,9	138,2	143,0	147,4	+35,7

Secondary school

School-gymnasium

School-lyceum

Small-scale school

Control schools   Pilot schools

❖ The range of differences in progress indicators of pilot schools varies from 3.3 to 9.9 scores (within 6.6 scores), while in control schools it varies from 6.7 to 20.0 scores (within 13.3 scores).

❖ This indicates a decrease in segregation in the pilot sample between lyceums, gymnasiums, secondary schools and small-scale schools.

PROGRESS OF AVERAGE SCORES BY LOCATION TYPES

	September ' 15	April' 16	May' 17	April' 18	April' 19	Difference September' 15 / April' 19
Urban						
control	119,7	133,3	133,9	138,4	140,2	20,5
pilot	113,8	139,0	143,7	152,2	151,1	37,3
Rural						
control	115,3	122,2	131,5	138,0	138,7	23,4
pilot	110,2	132,8	141,6	149,1	151,3	41,1

Urban

Rural

Control schools Pilot schools
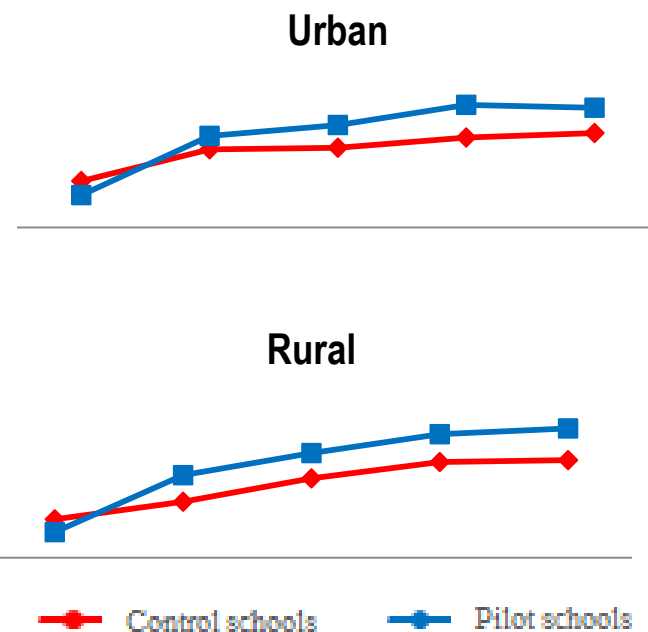
- ❖ The range of differences in the progress indicators of pilot schools varies from 0.2 to 6.2 scores (within 6.0 scores), while in control schools - from 0.4 to 11.1 scores (within 10.7 scores).

- ❖ In the pilot sample, indicators of rural and urban schools are more close.

AVERAGE SCORE BY REGIONS

CONTROL SCHOOLS

Regions	September' 15	April' 16	May' 17	April' 18	April' 19
Akmola	108,5	126,5	124,5	136,8	140,3
Aktobe	122,8	142,3	146,3	148,1	140,5
Almaty	115,5	130,2	120,8	133,2	135,4
Atyrau	125,2	135,3	135,0	140,7	143,3
East Kazakhstan	115,1	127,6	133,2	137,2	141,5
Zhambyl	113,9	125,5	129,5	135,9	139,7
West Kazakhstan	107,8	122,6	127,8	134,2	138,5
Karagandy	105,2	120,3	129,9	137,3	133,4
Kostanay*					
Kyzylorda	130,2	126,1	140,0	145,7	143,2
Mangystau	118,6	120,2	138,4	143,4	143,1
Pavlodar	113,7	115,6	124,6	132,2	130,2
North Kazakhstan	108,5	119,1	125,9	131,9	136,1
Turkestan	120,7	128,6	133,4	126,8	139,2
Almaty city	120,1	131,0	134,3	144,5	141,3
Nur-Sultan city	124,3	139,1	135,2	141,4	142,8

ПИЛОТНЫЕ ШКОЛЫ

Regions	September' 15	April' 16	May' 17	April' 18	April' 19
Akmola	116,1	134,4	140,3	147,8	148,6
Aktobe	116,1	137,5	142,7	149,4	151,4
Almaty	112,4	132,6	140,0	148,2	151,0
Atyrau	116,0	136,2	144,1	154,6	150,6
East Kazakhstan	112,3	130,0	139,6	151,6	149,5
Zhambyl	110,7	137,3	140,2	151,1	154,3
West Kazakhstan	115,7	140,1	144,0	151,7	146,5
Karagandy	115,5	132,4	137,5	152,9	148,4
Kostanay*	109,7	138,3	144,4	152,1	151,2
Kyzylorda	111,4	141,3	148,2	154,8	152,0
Mangystau	110,4	128,2	139,6	148,5	149,9
Pavlodar	111,8	128,8	135,0	149,5	156,1
North Kazakhstan	108,5	125,0	133,4	134,8	141,2
Turkestan	107,6	138,2	144,9	143,9	148,6
Almaty city	116,1	146,9	147,8	154,6	150,8
Nur-Sultan city	121,9	149,4	145,6	168,3	157,5

* T*here are no learners in the control school of Kostanay region who participated in all five diagnostic tests*.

MINIMUM AND MAXIMUM SCORES

Control schools

Pilot schools

- Minimum score
- Maximum score
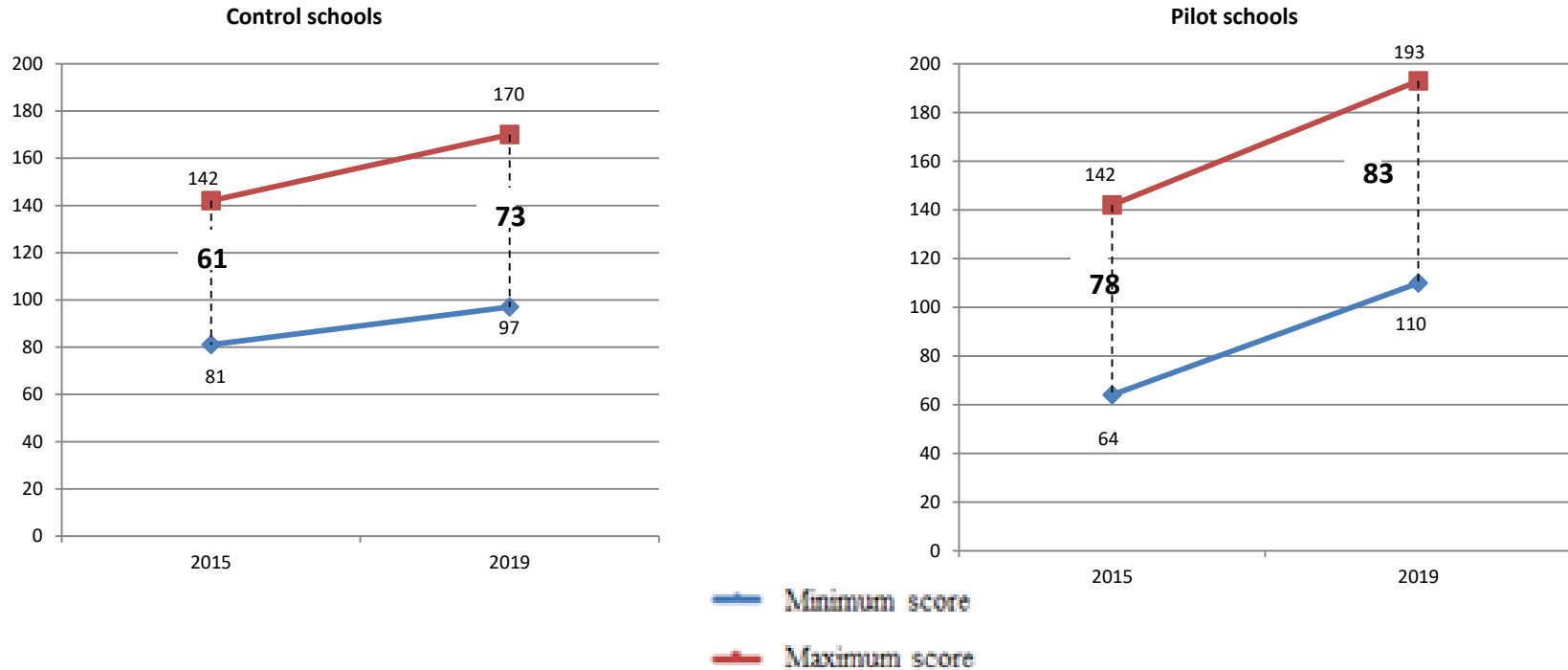
❖ The difference between high and low-performing learners in pilot schools for 4 years increased from 78 scores to 83 scores (+5 scores), in control schools it increased from 61 scores to 73 scores (+12 scores).

❖ In pilot schools minimal score increased from 64 (September 2015) to 110 (April 2019), i.e. the increase was 46 scores. In control schools this indicator increased from 81 scores (September 2015) to 97 scores (April 2019), i.e. the increase was 16 scores. Accordingly, the difference between low-performing learners in pilot and control samples is 13 scores. This may indicate effective support for low-performing children as part of the updated content of education.

CONCLUSIONS

The analysis demonstrates a comparative advantage in all monitored indicators and an increase in average scores on the ability scale in pilot schools

❖ The average progress in pilot schools is almost twice as high as in the control schools

❖ Learners' results demonstrate a more substantial average score progress in pilot schools than in control ones by languages of instruction

❖ There is a decrease in the difference of average scores in pilot schools between different types of schools (small-scale, secondary, gymnasiums, lyceums, rural and urban)

❖ In terms of diagnostic testing sections in pilot schools, a more approximate distribution of results is observed

❖ The results of completed tasks for the examples presented demonstrate a stable advantage of pilot schools

❖ In pilot schools, there is a narrowing gap between the minimum and maximum scores

USE OF DIAGNOSTIC TESTING RESULTS

Based on the results of each stage of diagnostic testing, to update and increase the practical significance of the results of diagnostic testing, an analysis of the results of learners was conducted.

Analysis of the results allowed:

❖ to confirm the effectiveness of the developed curricula, teaching materials and criteria-based assessment system in practice

❖ to analyse the learning results of the updated program

❖ to provide methodological support to teachers in the real educational process of 30 pilot schools

❖ to develop recommendations for improving the structure and content of subject programs and course plan of updated content of education, teaching materials, criteria-based assessment system, the content of teacher training courses

❖ to determine possible difficulties were predicted and prospects for the development of educational policy