The Results of Student Achievement Monitoring in Primary School in the Context of Educational Environment

Ekaterina Enchikova, Elena Kardanova

National Research University Higher School of Economics (Russian Federation)

enchicova@mail.ru, ekardanova@hse.ru

Abstract

The presentation investigates the results of Student Achievement Monitoring (SAM) in primary school within the context of educational environment. SAM was designed to characterize primary school students' subject competence (in mathematics and language) in terms of levels of mastery derived from Vygotsky's learning theory. We investigate three levels of educational environment: individual, class, and school. Since the data have a multilevel structure, we used the hierarchical regression model to investigate the interactions of variables. We pay special attention to the class and school characteristics, which may be adjusted by school or government, such as school type and learning curriculum. The teachers' practices and beliefs are other points of interest. We've found the development of students' subject competence is significantly correlated with the school type, educational program, class size, teachers' practice and experience.

Keywords: primary school, teacher, practice, subject competence, HLM

The paper investigates the development of student's subject competence in primary school[1] within the context of educational environment. The challenges of 21st century society demand the educational system to train students' skills beyond the basics. The development of student's subject competence (as opposed to simple memorization of academic material) becomes particularly important in the modern world.

Subject competence is often defined as ability to apply the subject knowledge in casual practice. Thus, knowledge is a necessary but not sufficient condition for the competence, which appears on the higher levels of the cognitive development. We operationalize the construct using the ideas of Russian psychologist Lev Vygotsky, who investigated the human mental development. This approach expanded in the works of his successors – Galperin (1998), Davydov (1972, 1996), Elkonin (1989) and Nezhnov (2007). According to Vygotsky's theory, cognitive growth can be described as a process of internalizing culturally transmitted knowledge, which involves acquisition of generalized schemas of thinking and symbolic systems (Vygotsky, 1978; 1994). Exposure to cultural models stimulates a gradual internal process of knowledge development. At the early stages of this process, individuals master specific procedures and associative links. At this level, their problem-solving very much relies on external characteristics of the problem; their ability to solve problems depends on how similar they are to the ones that had been directly taught. From this level, knowledge continues to develop to a more deep-level understanding of conceptual relations underlying learned procedures and finally, to the highest level of understanding that allows a person to

[1] Primary school in Russia corresponds to the ISCED level 1. The legal age of entry is 6-7 years; the duration of primary school in Russia is 4 years. We test children in the end of primary school, so they are 10-11 years old.

see the boundaries of the knowledge acquired and to be able to consider a multitude of possible relations within these boundaries. Three increasingly complex levels of mastery have been proposed: (1) procedural knowledge, (2) conceptual understanding, and (3) functional competence. Due to this taxonomy, subject competence refers to the second and third levels. Since the mental development goes through the acquisition of exterior experience, it is subject specific and must be examined within the subject areas.

Up to now, there have been no direct ways of assessing the different levels of subject competence within the outlined theoretical framework. The assessment instrument Student Achievement Monitoring (SAM) is the first attempt to do it. SAM was designed to characterize primary school students' knowledge (in mathematics and language) in terms of levels of mastery derived from Vygotsky's learning theory and to capture the distinction between procedural, conceptual and functional levels of knowledge (Nezhnov, P., 2011; Nezhnov, P., & Kardanova, E., 2011). To help separate participants into groups according to the level of their achievement benchmarks are used that reflect the three levels of the theoretical model.

Preliminary analysis of SAM test results revealed the majority of students achieve the second level of proficiency (conceptual understanding) by the end of primary school, but the third level (functional competence) is only starts to emerge. It corresponds to the Vygotsky's theory, which suggests that the development of the highest level of understanding of academic content continues after it was presented the recipient.

But we have found a big difference in test results and distribution of students among proficiency levels between different schools as well as between different classes of one school. So we explore the factors of the educational environment associated with the development of reasoning skills in language and mathematics. There are three levels of educational environment: (1) individual, (2) class, and (3) school. We pay special attention to the class and school characteristics, which may be adjusted by school or government, such as learning curriculum. The teachers' practices and beliefs are another point of interest, as they may have a significant impact to the students' subject competence.

Thus, the research question is – what characteristics of educational environment support the development of subject competence in primary school?

Method

Participants

The participants included 4406 fourth-grade students recruited from 192 primary schools (293 classes) in the Russian Federation. Fourth grade was chosen because it is the last year in Russian primary schools; children enter first grade at around the age of 7 years, therefore by the end of fourth grade their age range is 10-11 years. Participating schools were located in the central part of Russia where the majority of the population is ethnic Russians. Both urban and rural-area schools were included. The students represented a diverse range of socio-economic levels, matching the socio-economic structure of the whole region. The sample was closely divided by gender: 47% boys, 53% girls.

All participating schools were located in one of the central regions of Russia. This region was selected because its socio-economic characteristics (e.g., average salary, unemployment, educational level, urban-to-rural ratio) were similar to those in the entire country (Following the Results of All-Russian Population Census, 2010). For example, the distribution of the region's population by educational level (62% college and above, 30% high school, 8% below high school) was parallel to that in the country (65% college and above, 29% high school, 6% below high school). Also, the ratio of urban to rural students in the region (72% urban, 28% rural) was similar to that in the country (71% urban and 29% rural).

The regional department of education, in consultation with school principals, provided permission to conduct the study. Thus, the region's whole population of fourth grade students took part in the study. There was no selection at the school or class level. The data has multilevel structure – students are nested within the classes, schools and settlements. Table 1 shows the number of observations on each level.

Table 1 – Sample size

1	Number of students	4406
2	Number of classes	297
3	Number of schools	189
4	Number of settlements	134
5	Number of teachers who agreed to participate in the study	196 (cover 70% of students)

Instrument

Student Achievement Monitoring (SAM) toolkit (namely tests in Mathematics and Russian language) was used to assess the students' subject competence. The particular feature of SAM tests is that the test items within each content area vary systematically with respect to the depth of knowledge required. That is, each of the content areas included in the test is represented by problems tapping the three levels of subject competence. The tests include a total of 45 items each divided into 15 blocks. The majority of test items (37 out of 45, or 82%) had an open-ended format. They required either providing a brief numeric response or a simple drawing in the test booklet (for example, completing a shape pattern or placing a dot in a certain location within a figure). The remaining items (8 or 18%) had a multiple-choice format with a choice of one or more correct answers from 4-5 options. The multiple-choice items were evenly distributed across the test – they were not concentrated in any particular content area or knowledge level. All 45 items were assembled in a booklet; three items comprising each block were presented consecutively in the same order: levels 1, 2, and 3.

IRT modeling (namely Rasch dichotomous model) is used for SAM scale construction and students' estimation. There are two types of output data for SAM tests in each subject – the integral test score, which shows the general level of subject competences, and proficiency level. The benchmarks were established to help separate participants into groups according to the level of their achievement. The benchmarks reflected the three levels of the theoretical model, resulting in four distinct groups with the lowest-achieving group not having acquired Level 1 skills and the highest-achieving group having acquired

Level 3 proficiency. Specific methods of developing benchmarks are described in detail elsewhere (Kardanova & Nezhnov, 2011). Thus we have 2 indicators for each student, which can be used as depended variables.

SAM toolkit also includes a set of teachers' questionnaires which provide sociological, institutional, social and cultural information describing the context of educational environment.

Analytic model

Multilevel data structure (students are nested within classes) demands a specific method of statistical analysis. We used the hierarchical regression model (HLM) to investigate the interactions of variables. Applying HLM allows us to deal with the problem of independence of observations. Since the students are grouped in the classes and the observations are not independent, the simple linear regression is intended to underestimate the standard errors of the regression coefficients. It increases the probability of false rejection of a null hypothesis (the type 1 error hazard). We use the hierarchical regression model to identify drivers of subject competence development and possible ways of enhancing development of such skills.

Two-level hierarchical linear models (HLMs) were used to examine associations of educational environment characteristics with the subject competence development of 4406 fourth-grade students (Level 1) in 293 classes (Level 2). We don't apply three-level model, as there are 128 schools out of 189 (68%) have only one class, so the class level match the school level for those observations.

The SAM test provides us the results in Russian language and mathematics subject competence. We use the integral test score as the depended variable in the regression model. Thus we build separate models for the language and mathematics subject competence. The characteristics of educational environment come as the independent variables. There are seven categories of independent variables in this research.

1. "Gender" is the sex of the responder. It is dummy variable, 1 - is female, 0 – is male
2. "School location" is the type of settlement, where the school is located. This variable has 3 categories: 1) big city (population is more than 200 thousands), 2) town (population is less than 200 thousands ; 3) rural area . Two dummy were created, "big city" is reference category.
3. The school type "gymnasia" includes the special schools that usually have better material base, better teaching staff, deeper study of standard disciplines, and, as consequently, higher results. It is dummy variable, 1 - is gymnasia, 0 – is normal school.
4. The "class size" is the number of students in each class. In primary school it is usually around 25 students. But in the rural area the class size might be smaller, because of the low population. We identify the small classes as the classes with less than 11 students. It is dummy variable, 1 - is a small class, 0 – is normal class.
5. The "educational program" is the set of the training and monitoring tools and instructional documentation for each discipline. There are a lot of educational programs for primary school in Russia, so we selected the most popular and widespread (according to the number of students). We identify four types: 1) School of Russia, 2) School 2100, 3) System of Zankov and 4) Other school programs. So three dummy were created, "School of Russia" is the reference category.

6. The “teachers’ practices” are also correlated to the students’ results. We distinguish two pedagogical approaches – *constructivism* and *traditionalism* (Brooks & Brooks, 1993). Due to the constructivism approach, teachers intend to create a friendly atmosphere in the classroom, to involve children in the conversation, and to encourage the experimental verification of children’s hypotheses. The traditionalism teachers more emphasis on discipline in the classroom, a strong knowledge and is important to answer the question accurately as in the tutorial. We divide constructivism approach into two scales – teachers’ beliefs about the constructive way of teaching, and the implementation of it in practice. Thus there are three scales of teachers’ approaches: “constructivism beliefs”, “constructivism practice” and “traditionalism practice”. We created these scales based on the teachers’ questionnaire, using IRT approach. The scales demonstrate good psychometric quality. We use the integral scale scores as the independent variables
7. The “teachers’ experience” is the number of years, the teacher work in school.

These independent variables refer to the characteristics of the educational environment in primary school, and there are two types of them:

1. The characteristics, which can’t be adjusted by school management – school location, school type. The class size is also refers to this category, because it depends on the population of the region.
2. The characteristics, which can be adjusted by school management – the educational program, the teachers’ practice and experience.

To assess the effect of each type of environmental characteristics, there were built two regression models. The model #1 includes only the characteristics, which can’t be adjusted by school management. The model #2 includes also the adjustable characteristics.

Results

To investigate the development of subject competence in primary school, let’s start with the descriptive statistics. There are 4406 children participate in the study. The mean Russian language test score is 498.3, standard deviation is 49.9. The mean mathematics test score is 522.1, standard deviation is 48.7. The table 2 shows the distribution of sample on the proficiency levels in Russian language and mathematics in primary school. The average results in mathematics test are better than in the Russian language test. Second level is dominated in Mathematics; more than 50% of students achieved it. The Russian language test results show roughly equal amount of children on the first and second levels (about 40%).The number of children on the third level is low - 12% Russian and 18% - in mathematics. It corresponds to the theory that the functional competence only starts to emerge in the primary school.

Table 2 – the distribution of sample on the proficiency levels

Levels	**Russian language**		**Mathematics**	
	N	%	N	%
Below first level	475	10.8%	103	2.3%
1 level – procedural knowledge	1727	39.2%	1187	26.9%

2 level – conceptual understanding	1662	37.7%	2333	53%
3 level – functional competence	542	12.3%	783	17.8%

The descriptive statistics of the independent variables are shown in the table 3. We can see the mean score vary through the groups. The number of observations in each group allows us to use these variables in hierarchical regression model.

Table 3 – descriptive statistics of the independent variables

Variables		Number of students	**Russian Language** (test score) Mean	SD	**Mathematics** (test score) Mean	SD
Gender	Girls	2078	506.7	48.1	524.5	49.0
	Boys (referent)	2328	490.8	50.3	520.0	48.4
Location	Big city (referent)	1684	499.8	47.6	526.1	45.9
	Town	1482	503.6	49.3	526.5	50.6
	Rural	1240	490.0	52.7	511.6	48.6
School type	Gymnasia	647	512.7	47.3	540.9	48.0
	Normal school (referent)	3759	495.8	49.9	518.9	48.1
Class size	Big class (referent)	3952	497.8	49.7	522.6	49.1
	Small class	454	503.2	51.9	518.3	44.8
School program	School 2100	562	518.0	42.2	544.3	43.3
	School of Russia (referent)	1648	484.8	48.6	506.3	46.7
	System of Zankov	297	497.0	41.5	521.7	38.0
	Other school programs	1899	504.4	51.3	529.4	49.2

Next step we build the hierarchical regression models to investigate how the characteristics of educational environment are connected with the development of students' subject competence, which evaluated by the mean scores of SAM test. The results of the analysis are presented in table 4.

Intraclass correlation coefficients (ICC) show that approximately 50% of SAM scores dispersion is explained with the association of students in classes (2 level groups). In other words, personal and school environmental characteristics contribute equally in the development of the students' subject competence. But ICC is insufficiently decreased when we add independent variables to the models. Thus with the control of the independent variables, the differences between the classes are still significant. We can also see, that independent variables explain only 13.2% of Russian language test score dispersion, and 16.6% of mathematics score dispersion. The small percent of explained variance is expected, because there are no independent variables on personal level (level 1) in the model. The only personal characteristic is gender (girls have a significantly better results in language test).

The independent variables in model #1 explain only 5-6% of the test score dispersion. It means that the school characteristics (location, school type and the class size) don't have much effect on the development of the students' subject competence. It is important that school location doesn't connect to the test results; it means that students from different types of settlements have equal opportunities in education. However, gymnasia school type has a positive significant correlation

with the test results, but the effect size is small. And the effect of gymnasia schools decreases while controlling of educational programs and teachers' characteristics.

The question of effectiveness of education in small classes is very important. Approximately a third part of all teachers in region work in small classes, while less than 10% of students study in such classes. This research shows that education in small classes is no worse than in big ones in mathematics. And small classes have a significantly higher result in language test.

Table 4 – Hierarchical Linear Model Results

Dependent variable		**Russian Language** (test score)			**Mathematics** (test score)		
MODEL #		Null model	Model 1	Model 2	Null model	Model 1	Model 2
FIXED EFFECTS							
	CLASS MEAN (γ00)	498.7*** (2.2)	485.9*** (4.6)	456.1*** (9.7)	520.4*** (2.1)	517.4*** (4.5)	479.8*** (8.7)
Gender	Girls		13.7*** (1.2)	13.7*** (1.2)		1.6 (1.1)	1.6 (1.1)
Location (reference category – big city)	Town location		7.7 (6.1)	6.07 (5.6)		5.7 (6.1)	3.7 (5.7)
	Rural location		-3.9 (6.8)	-4.2 (6.6)		-8.9 (6.4)	-9.9 (6.2)
School type	Gymnasia		17.8** (7.2)	11.8* (6.3)		21.5*** (7.6)	15.01** (6.5)
Class size	Small class		12.7** (6)	13.8** (6)		7.1 (5.4)	7.7 (5.3)
School program (reference category – "School of Russia")	School 2100			23.1*** (6.5)			25*** (6.5)
	System of Zankov			3.1 (6.2)			6.6 (6.3)
	Other school programs			10.9** (4.7)			16.03*** (4.6)
Teacher characteristics	Constructivism teacher believes			0.88 (1.1)			1.5 (1.1)
	Constructivism teacher practice			4.3*** (1.6)			4.2 (1.6)
	Traditionalism teacher practice			-5.6* (3.1)			-2.4 (2.8)
	Teachers' work experience			0.53* (0.29)			0.69* (0.26)
RANDOM EFFECTS							
Class mean	St. deviation, u_{0j}	35.6	34.7	33.3	34.4	33.3	31.3
	Variation	1273	1209	1105	1180	1108	983
Level – 1	St. deviation , r_{ij}	35.8	35.3	35.3	33.9	33.9	33.9
	Variation	1285	1240	1240	1151	1151	1151
Percentage of variance explained	Within class		3.5	3.5		0	0
	Between classes		5	13.2		6.1	16.6
Intraclass correlation coefficient (ICC)		49.7	49.3	47.1	50.6	49.04	46.1

Note: class-level variables were grand mean centered

Standard errors in parentheses, *** p<0.01, ** p<0.05, p* <0.1

We also investigated the difference in test results between school educational programs. We found the program "School 2100" has significantly higher test results, than the other programs. Please, note that we can claim only the difference in the development of subject competence of trained students, but we can't find the causal relationship. We don't control all the environmental characteristics, and we don't know the initial level of students, so we can't argue that the high results are the consequence of the educational program. Such hypothesis need more research and might be the object of further investigations.

Investigating the teachers' characteristics we can find the constructivism teachers' beliefs are not connected to the development of students' subject competence. But teachers' practice correlate significantly with the language test results. Constructivism practice correlates positively, and traditionalism practice correlates negatively. And it is interesting, that it is not so important what teachers believe, as what they do in practice. The teachers' experience is also correlates positively with the development of subject competence.

Discussion

During the research there were designed three scales aimed to assess the teachers' beliefs and practice approach. The scales can be interpreted as traditional and constructivist beliefs and constructivist practices. These scales show good psychometric quality and can be used for further analysis.

The multilevel modeling shows, that there are approximately 50% of students variance explained on the personal level, and 50% explained on the school and class level. Thus, family and school contribute equally to the children's progress. We found, that development of students' reasoning skills is significantly correlated with the school type, educational program, class size, teachers' practice and experience.

The results interpretation is limited with the features of the data design. There is no data on the personal level, so our conclusions might claim only the connections between the examined characteristics, but not causal relationships. The coefficients of model might also change, while include the other parameters.

The model, based on the Russian data can be applied to the educational systems of other countries. There is a particular interest to confirm the discovered patterns of variables' connections on different sample in different educational systems.

This research is the first step for the international project which aims to investigate the relation of the teaching and learning approaches and development of student's reasoning skills in order to enhance teaching and learning for the benefit of education quality improvement. The project will take place in Russia, Tajikistan, Uzbekistan, Kyrgyz Republic, Armenia, Mongolia, and Moldova. It will explore the educational systems in these countries and integrate the most efficient learning approaches. The present research intends to highlight the possible opportunities of investigation in this field.

Reference

1. Brooks, J.G., Brooks, M.G. (1993). *The case for constructivist classrooms*. Alexandria, association for Supervision and Curriculum Development.
2. Galperin, P. (1998). *Psychology as an objective science*. Moscow, Russia: Institute of Applied Psychology.
3. Davydov, V. V. (1972). *Types of generalization in learning*. Moscow, Russia: Pedagogy.
4. Davydov, V. V. (1996). *The theory of developing learning*. Moscow, Russia: Intor.
5. Elkonin, D. B. (1989). *Selected psychological work*. Moscow, Russia: Pedagogy. *International Journal of Testing*, 10:4, 295 — 317.
6. Kardanova, E., Nezhnov, P. (2011): School achievements monitoring toolkit: Assessment framework. Paper presented at the *37-th Annual Conference IAEA*, Manila.
7. Nezhnov, P. (2007). Mediation and spontaneity in the cultural development model. *Moscow State University Bulletin, 14*, 133-146.
8. Nezhnov, P., & Kardanova, E. (2011). SAM Framework. Center for International Cooperation in Education Development.
9. Kardanova, E., Nezhnov, P. (2011): School achievements monitoring toolkit: Assessment framework. Paper presented at the *37-th Annual Conference IAEA*, Manila.
10. Nezhnov, P. (2011). SAM – toolkit to assess primary school students' academic achievements. CADMO. Innovations in assessment to meet changing needs. ANNO XIX, 1, pp.85-98.
11. Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes.* Cambridge, MA: Harvard University Press. (Original works published in 1930, 1933).
12. Vygotsky, L. S. (1994a). The problem of the cultural development of the child. In Van der Veer, R., & Valsiner, J. (Eds), *The Vygotsky reader*. Oxford, UK: Blackwell. (Original work published in 1929).
13. The Results of All-Russian Population Census (2010), Retrieved from URL: http://www.gks.ru/free_doc/new_site/perepis2010/croc/Documents/portret-russia.pdf.