



NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

*Brun Irina, Ivanova Alina, Kardanova Elena,
Orel Ekaterina*

NONCOGNITIVE DEVELOPMENT OF FIRST GRADERS AND THEIR COGNITIVE PERFORMANCE

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: PSYCHOLOGY
WP BRP 57/PSY/2016

This Working Paper is an output of a research project implemented within NRU HSE's Annual Thematic Plan for Basic and Applied Research. Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE

*Brun Irina¹, Ivanova Alina², Kardanova Elena³,
Orel Ekaterina⁴*

NONCOGNITIVE DEVELOPMENT OF FIRST GRADERS AND THEIR COGNITIVE PERFORMANCE*

It is well known that performance, both cognitive and noncognitive, in primary school is very important for children's future outcomes. In this study we attempt to classify patterns of cognitive, social, emotional and personal development based on data from first-graders beginning their schooling. We use complex iPIPS (international Performance Indicators in Primary Schools) data – a large-scale assessment of first-year pupils, which includes math and reading tasks along with noncognitive assessment – gathered from 1202 children from the Republic of Tatarstan. We describe 5 clusters of first-year pupils and give background information about family and preschool experience which may influence performance in each cluster.

Key words: iPIPS; noncognitive development; primary school; first-graders; large-scale assessment, cognitive development

JEL Classification: Z.

¹ Center of Education Quality Monitoring, Institute of Education, National Research University Higher School of Economics, junior research fellow, E-mail: ibrun@hse.ru

² Center of Education Quality Monitoring, Institute of Education, National Research University Higher School of Economics, research assistant, E-mail: aeivanova@hse.ru

³ Center of Education Quality Monitoring, Institute of Education, National Research University Higher School of Economics, head of Center of Education Quality Monitoring, PhD, E-mail: ekardanova@hse.ru

⁴ Center of Education Quality Monitoring, Institute of Education, National Research University Higher School of Economics, research fellow, E-mail: eorel@hse.ru

* The study was implemented in the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) in 2015. Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.

1. Introduction

Interest in PSED – personal, social and emotional development (e.g., team-work, communication, perseverance, optimism, emotional control, gratitude, social intelligence, curiosity, etc.) has a long history. Researchers in psychology and education, school administrators, teachers, parents, and even children themselves have long been aware that education involves interactions among people. This awareness has prompted scientific research devoted to personal, social and emotional development and the relation between such development and cognitive development and job performance. Modern education is based not only on knowledge and skills transfer, but also on fostering communication skills, teamwork, self-confidence, cultural awareness and so on. Children who have good inter-personal skills and can self-regulate their behavior and emotions are more likely to associate well with their peers and their teachers, and to reap the benefits of their education (Merrell, Bailey, 2008).

In what spheres does PSED matter? One example is academic performance. School systems around the world focus on both knowledge acquisition and developing social-emotional skills in youth. In an impressive meta-analysis that summarized the results of 213 school-based social-emotional learning programs conducted with pupils (N= 270,034) from kindergarten to high school, Durlak and colleagues (2011) reported effects of social-emotional learning ranging from 0.22 (conduct problems) to 0.57 (social-emotional skills). They also found that social-emotional skills training had direct effects on academic performance. Social, emotional and personal skills are hence both a means and an end in formal education programs. These results were confirmed in many other studies focused on different levels of education (at the primary level (Miles, Stipek, 2006, Elias, Haynes, 2008), secondary level (Parker et al., 2004, Zins et al., 2004), and higher education level (Poropat, 2009)).

Research conducted by Hindman, Skibbe, Miller, & Zimmerman (2010) focused on relations between child and family characteristics, classroom factors, and literacy and mathematics skills during preschool and the first grade. They showed that a child's gender and age together account for approximately 16% of the variance in literacy skills, and preschool experience and social skills explained 21% of the variance. Language, vocabulary knowledge and social skills together account for 28% of variance in math skills. They also showed that math progress during the first year is less rapid compared to preschool years. The only important predictor for math growth is vocabulary skills. One of this study's most interesting findings is that no family and classroom factors predicted significantly math or literacy scores and growth (Hindman, Skibbe, Miller, Zimmerman, 2010). Another study looked at gender differences in self-regulatory skills and academic achievement among kindergarten children. They found no

significant differences in academic achievement between boys and girls, but measures of self-regulation predicted significant variance in math and phonological awareness gains during the preschool year (Matthews, Cameron Ponitz, Morrison, 2009). Other researchers found that classroom behavior of first-graders predicts their academic performance 4 years later (Alexander, Entwisle, Dauber, 1993). Hamre and Pianta (2001) stated that a conflict between a child and a teacher and the level of a child's dependency in kindergarten have long-term consequences, and result in poor academic outcomes (especially for boys) in the eighth grade. The results from this study also emphasize the role of teacher-child relationships in children's behaviour problems and educational attainment in primary and middle school (Hamre & Pianta, 2001). These results are consistent with previous research (Birch & Ladd, 1997; Pianta & Nimetz, 1991). Eleven years later Denham, et al. (2012) showed that negative/aggressive social-emotional behaviour during kindergarten (3 and 4-year olds) is related to early school success and kindergarten academic indicators.

Merrell and Bailey used a PSED questionnaire (personal social and emotional development) and PIPS results to investigate how strong the relationship between noncognitive and cognitive development of four-to-seven-year-olds living in England is. According to their research, the association between cognitive and noncognitive development weakens with age. At the age of 4, the correlation is 0.5 for Reading and 0.49 for Math; at the age of 7, the correlation coefficients are 0.37 and 0.41 for Reading and Math, respectively. Among 11 PSED areas, Communication has the highest association with educational achievement (0.5 Reading, 0.47 Math), and the lowest association was found between Comfort and cognitive scales (0.25 Reading, 0.23 Math). As for predicted progress in Reading and Math, PSED at the start of school (at the age of four) is a significant predictor for Reading and Math scores at the age of seven (controlled for scores at the beginning of school).

All these findings confirm that different aspects of children's early social and emotional development play a significant role in later educational attainment throughout primary and middle school.

Another area where social and emotional skills are of great importance is job performance. In a meta-analysis of personality measures and job performance, Hurtz and Donovan (2000) reported estimated mean operational (true) validity coefficients (corrected for unreliability in the criterion measure and range restriction on the personality measures) for self-reported personality (according to the Five-factor model) of 0.22 (Conscientiousness), 0.14 (Emotional Stability), 0.10 (Agreeableness), 0.09 (Extroversion), and 0.05 (Openness). However, in an analysis aligning "Big Five" dimensions to specific job-performance aspects, Hogan and Holland (2003) reported true estimated validities of 0.43 (Emotional Stability), 0.35 (Extroversion-Ambition),

0.34 (Agreeableness), 0.36 (Conscientiousness) and 0.34 (Openness to Experience). These findings unequivocally demonstrate that the underlying broad dimensions of personality description are related to various performance indicators valued by employers. So early personality development is highly important for future employment outcomes.

Social and emotional development is also an important predictor of future health. In a recent review, Friedman and Kern (2014) distinguished six core health outcomes covered in public health policy research: physical health, subjective well-being, social competence, productivity, cognitive function, and longevity. At present, it is evident that social and emotional skills are associated with a wide range of health outcomes, such as smoking (Munafò, Zetterler, and Clark, 2007), obesity (Sutin, Ferrucci, Zonderman, and Terracciano, 2011), alcohol craving and consumption (Papachristou et al., 2013; Stautz and Cooper, 2013), resilience to Alzheimer's disease (Terracciano et al., 2013) and health status in cardiovascular populations (Versteeg, Spek, Pedersen, and Denollet, 2012).

As mentioned above, previous research has shown that social and emotional development influences various outcomes in different spheres in life. Thus, it is important to estimate levels of social and emotional development at the beginning of schooling and establish groups of students with similar patterns of social and emotional development. Establishing groups, we believe, will support teachers in developing strategies to help children better adapt to the school environment, correct problematic behaviour, etc.

This paper has been prepared in the framework of iPIPS (international Performance Indicators in Primary Schools)¹, which gives an opportunity to assess both cognitive (basic knowledge in math and reading) and noncognitive (social and emotional) development of children when they start schooling.

The purpose of this study is to define and describe groups of first graders based on a combination of cognitive and noncognitive development. To describe these groups we examined the demographic characteristics of students' families and children's preschool experience. The main research question of our study is whether or not we can define meaningful patterns of cognitive, personal, social and emotional development in the early years of schooling

¹<http://www.ipips.org/>

2. Method

2.1 Instrument

Throughout this study, we have employed iPIPS (international Performance Indicators in Primary Schools) as an instrument to assess children's cognitive and noncognitive development.

The iPIPS is based on the Performance Indicators in Primary Schools monitoring system (PIPS), which was developed nearly 20 years ago in the UK (Tymms 1999a) with the intention of providing baseline data for primary schools and producing objective, nationally-based, value-added scores at a later stage in children's education. The assessment was designed to be carried out twice a year. The instrument provides high quality information for teachers about what their pupils know and can do at the beginning of their formal education in order to help the school improve provision of services. The follow-up assessment at the end of the first year of schooling provides reliable evaluation of children's educational progress.

The use of the PIPS baseline and follow-up assessments has been expanding, and the assessments have been translated into and adapted into different languages, including German, Chinese and now Russian. The assessment has proved to be reliable and valid in a number of countries. More recently, the international Performance Indicators in Primary Schools (iPIPS) project has been established to help policymakers evaluate education policy and practices in their countries. Since its development, the instrument has been used to assess the cognitive and noncognitive development of more than three million children across several countries. Full documentation outlining the iPIPS assessment is available on the project's website (www.ipips.org).

In 2013-2014, iPIPS was translated and adapted for use in Russia (Hawker, Kardanova, 2014; Ivanova, Nisskaya, 2015).

In Russia, as well as in the UK, the assessment procedure is conducted via computer. The content of the cognitive assessment is very comprehensive but no individual child sees all of it, because it is presented in a series of sequences with stopping rules. This adaptive nature means that a very wide range of children can be assessed in an efficient, reliable and enjoyable way.

The assessment scheme of Russian iPIPS is the same as in original version and consists of two main parts – cognitive and noncognitive assessments. The first part is designed to evaluate those aspects of juvenile cognitive development which predict later success or difficulties in education (Tymms, 1999a and 1999b; Tymms, Brien, Merrell et al, 2003; Tymms, Merrell, Henderson et al, 2012). The noncognitive section of the instrument assesses personal, social, and

emotional development, and is based on teachers' observations of their pupils' behavior in the school setting.

The cognitive section is made up from the following sub-units:

- Handwriting, in which the child is asked to write his/her own name
- Vocabulary, in which the child is asked to identify objects within pictures
- Ideas about reading, in which concepts about text structure are assessed
- Phonological awareness: rhymes and repeats
- Letter identification
- Word recognition
- Reading as decoding
- Reading comprehension
- Counting
- Addition and subtraction problems presented with pictures
- Digit identification
- Mathematics problems including sums with symbols.

The noncognitive section is comprised of the PSED survey, which uses the teacher's knowledge of each child as gained through general day-to-day interactions and observations. The assessment involves determining the place of each child on each of eleven items of the PSED questionnaire including: Comfort, Independence, Concentration (teacher-directed activities), Concentration (self-directed activities), Actions, Relationships with Peers, Relationships with Adults, Rules, Cultural Awareness, Communication.

These eleven items are arranged into three sections. The section "social development" consists of five items, "personal development" consists of four items, and two items are dedicated to measure "adjustment to the school environment." Each child is assessed on each of the eleven items using a five-point scale. For each point on the scale there is a detailed description of typical behavior for the teacher to rely on. The teacher then decides which descriptor provides the closest match to the behaviour of the child in question. Items in the assessment are summarized in the table 1.

Table 1. iPIPS PSED item description

Item	Description
------	-------------

Adjusting to the school environment	
Comfort	This item seeks to measure the extent to which a child is comfortable with his/her separation from his/her main carer, ability to cope with transitions between locations and activities, and generally how settled he/she is during the day.
Independence	Dependency on adults or other children for guidance and support is measured, as well as the extent to which a child needs help with dressing and going to the bathroom.
Personal development	
Confidence	A child's willingness to talk and ability to join group activities is measured.
Concentration (teacher-directed activities)	This item addresses concentration on tasks issuing from the teacher. Is the child able to maintain concentration and remain undisturbed in the face of competing activities?
Concentration (self-directed activities)	Similar to the item above, but focuses on activities chosen by the child rather than determined by the teacher.
Actions	A child's impulsivity is measured with this item. Does he/she act without consideration for him/herself or others? Does he/she demonstrate appropriate behavior and interact well with others?
Social development	
Relationship to peers	This item measures the child's ability to communicate, make friends and take notice of the feelings of others.
Relationship to adults	This item measures the child's ability to approach and communicate with adults and to interact appropriately and confidently.
Rules	This item measures the extent to which the child obeys rules or instructions and restrains from distracting peers.
Cultural awareness	This item measures the extent to which the child understands that others may have a different way of life and that this should be respected.
Communication	Is the child able to communicate fluently and coherently, listen to the views of others, respond appropriately and take turns in conversation?

Each of these items can be seen as measuring an element of a child's personal, social or emotional skills which is likely to affect his or her educational experience and relationships with

their peers, teachers and others. It is likely that a classroom teacher is aware of the lack of these skills in students without the need for assessment, but by measuring these skills and then repeating the assessment later in time, teachers can judge whether appropriate noncognitive progress is made. If the progress is less than expected, it could indicate a deficit which is likely to affect children in their future years of schooling.

2.2 Participants

The target population for this study is children enrolled in the 1st grade of school on the 1st of September 2014. The Russian sample consists of 1202 children recruited from 29 schools (20 regular schools, 9 grammar schools) in the Tatarstan Republic. This region was selected for analysis because in important aspects it represents the Russian Federation as a whole (although it can not be considered as a representative sample). Tatarstan's socio-economic characteristics (e.g., average salary, unemployment rate, educational level, urban-to-rural ratio) were similar to the national average (based on 2012 census, REF). It should be noted that sample used in analyses was smaller than initial sample because of the response rate for PSED and parents' questionnaires. There were 409 boys (82 aged up to 7 years old, 215 from 7 to 7.5, 112 older than 7.5) and 447 girls (109 aged up to 7 years old, 214 from 7 to 7.5 and 124 older than 7.5). Gender data were collected via parents' questionnaire, the response rate for parents is 71.2%.

2.3 Analysis

We used IRT modelling for item analysis, our dimensionality and reliability study, scale construction for cognitive and noncognitive data, as well as for student estimation for all scales. In particular, the one-parameter dichotomous Rasch model (Wright B.D.& Stone M.N., 1979) was used for cognitive data to transform students' raw scores in mathematics and reading into measures on an objective interval scale. The Rating Scale Model (RSA), which is an extension of the Rasch dichotomous model for polytomous items, was used for noncognitive data (Wright B.D.& Masters G.N., 1979). The Rasch models were chosen for both psychometrical and practical reasons. Firstly, Rasch models have optimal metric properties, and secondly, from a practical point of view, they are useful for empirically determining the quality of test items and their response categories, constructing scales and carrying out different problems of testing (Bond & Fox, 2001). Winsteps software (Linacre J. M., 2011) was used for parameters estimation and data analysis under the Rasch models.

To define different groups of children based on their cognitive and noncognitive development, we used k-means cluster analysis. Prior to this, a clustering correlation matrix for

four clustering parameters was computed (the Pearson correlation coefficient, r , was used). The goal of cluster analysis is to discover groups in the data based on the patterns of the responses. In keeping with the purpose of the study, cluster analysis was used to identify groups of first-graders with similar levels of cognitive and noncognitive development. Partitional method was used because the sample size is too big for hierarchical clustering. The k-means method is widely known and is considered one of the simplest and most popular algorithms for pattern recognition (Jain, 2010). It has a number of disadvantages, however. Unlike hierarchical cluster analysis, the k-means method requires a number of clusters to be identified beforehand. There are several approaches for determining the number of clusters to be used, including the meaningfulness of established groups (Jain, 2010). Our hypothesis was that our data contains four or five clusters: one of the clusters was expected to show high results on all four scales, another to show low results on all four scales. The remaining two or three clusters were expected to display different tendencies depending on which domain (cognitive or noncognitive) is more developed. We conducted cluster analysis with 3, 4 and 5 clusters, and experts in pedagogy and psychology determined which cluster solution is more meaningful.

To describe the social and demographical information of these groups we used association coefficients: chi-square and Cramer's V statistics. SPSS package was used to assess the extent of association.

3. Results

The Results section will be organized as following. Firstly, we will briefly present the results of psychometric analysis via iPIS instruments to ensure that both cognitive scales and the noncognitive PSED questionnaire provide reliable measures of children's skills in mathematics and reading, as well as a reliable assessment of children's social and emotional development. Secondly, we use cognitive and noncognitive measures to conduct cluster analysis aimed at establishing groups of children with different patterns of cognitive and noncognitive development.

3.1. Psychometric analysis of cognitive scales

Because of space limitations, only the results of the math scale analysis are presented here. There were 46 items in this test which included items such as recognition of numbers, ideas about mathematics, and use of arithmetical operations. Two items were removed from the scale because of their poor fit to the model. Both items were extremely easy for the Russian children and didn't input much to the estimation of their math abilities. The remaining items (44 in total)

formed a good scale which was unidimensional, with all items fitting the model and displaying good psychometric characteristics.

Figure 1 presents the Rasch variable map, which shows the relative distribution of items and test participants in a common metric. Specifically, the variable map depicts the joint distribution of items operationally defining the mathematics variable and the locations of children, based on their total correct scores, along this variable. The left column is the “logit” unit of measurement scale. On the map, students are represented on the left side and the items are on the right. More difficult items and higher-performing students are located in the upper part of the map (positive logits), while easier items and lower-performing students are placed in the lower part of the map (negative logits).

The distribution of students is wide and displays good differentiation between higher and lower scoring students. The distribution of item locations is also satisfactory because the span includes very easy items appropriate for less-able students, and very difficult items appropriate for advanced students. Furthermore, the progression of items from easier to more difficult represents a smooth, uniform, and progressive continuum of increasing difficulty. The student sample is well-located relative to the mathematics variables, which means that the test was well-targeted for the sample.

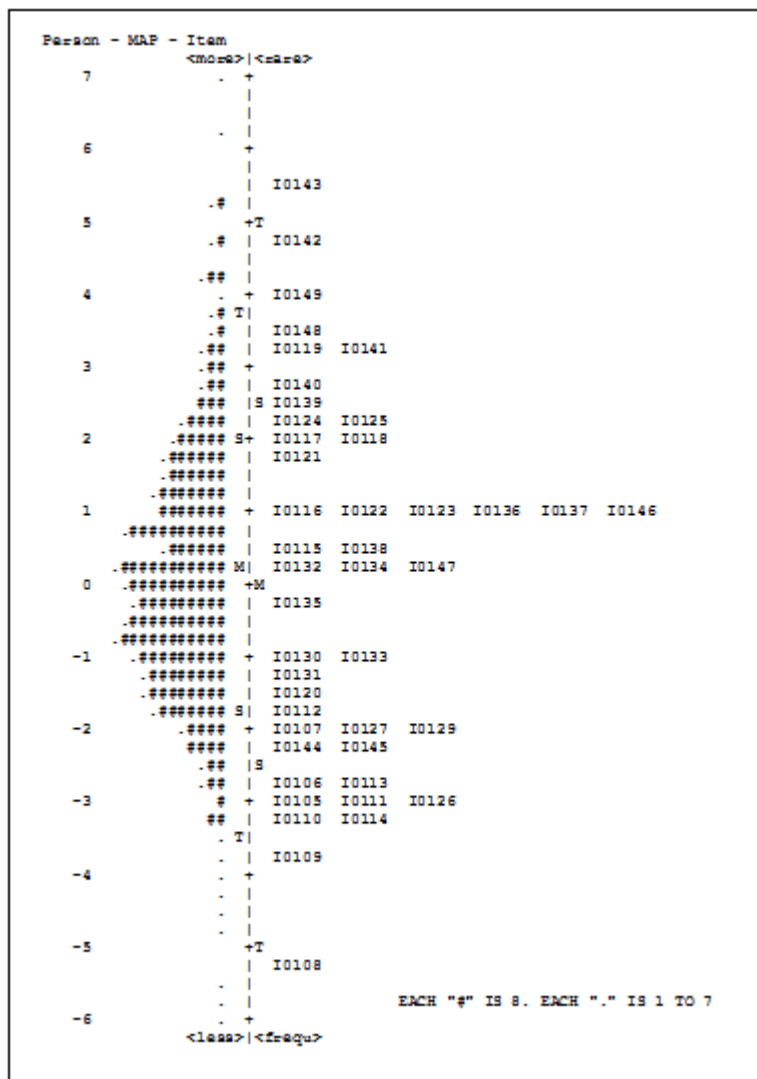


Figure 1. The iPIPS math baseline assessment variable map

Our analysis produced a “person reliability” of 0.90, meaning that the proportion of observed student variance considered true was 90%. (This index is close in value and interpretation to classical reliability alpha, which is 0.92 for the math baseline scale).

Thus, our analysis found that iPIPS math test items, and the test overall, display satisfactory psychometric qualities. This test can be used to measure students’ math ability at the beginning of first grade.

The results for the reading scale are substantively similar (classical reliability alpha is 0.95 for the reading scale).

Estimation of students’ measures was conducted separately for each scale, resulting in two different estimates for each individual: a measure of math ability and a measure of reading

ability. Table 2 presents descriptive statistics of students' measures (in logits). For each subscale, the mean of item difficulties is set to be equal to 0 to fix the scale origin.

Table 2. Measures of math and reading abilities

	Mean	Standard Deviation	Min	Max
Math	0.00	1.75	-5.62	5.99
Reading	0.09	1.61	-6.05	5.37

3.2. Psychometric analysis of PSED questionnaire

In the first stage of our analysis we decided to examine the empirical structure of the PSED questionnaire. The original three-dimensional structure (Table 1) is based on the statutory framework for early years foundation stage provided by the UK Department of Education (Statutory Framework, 2014), which does not provide any empirical evidence for these dimensions. Russian educational standards for preschool education also refers to social and personal development, but does not correspond to British standards. Because of the lack of the empirical evidence for the PSED structure, we ignored its original three-dimensional structure and conducted a dimensionality study of the PSED questionnaire.

We examined the dimensionality of PSED scale by conducting a principal component analysis (PCA) on the standardized residuals, which represents the difference between the observed responses and the responses expected under the model (Ludlow, 1985; Smith, 2002). Theoretically, if the assumption is withheld, then correlations between item-level residuals should be close to zero. If there is no second dimension remaining in the residual variation, then the PCA should generate eigenvalues that are all close to one and the percentage of variance across the components should be uniform. The analysis of the eigenvalues of the PSED scale residual correlation matrix for 11 components indicated that there is one component with an eigenvalue of 2.7, whereas the eigenvalues for the other components ranged from 1.5 to 0.47 (with the exception of the eigenvalue for the last component, which was 0.015). In addition, the percentage of variance for the first component was 24%, whereas the variance for the other components was roughly evenly split across the components. This provided us with a basis for treating the PSED scale as two-dimensional, consisting of two subscales. Table 3 shows the standardized residual loadings for all items.

Table 3. Standardized residuals loadings for the PSED scale

Subscale	Item	loading
Self-confidence	Confidence	0.66
	Comfort	0.52
	Independence	0.40
	Relationship to peers	0.39
	Relationship to adults	0.26
	Communication	0.23
Classroom behaviour	Concentration 1: Teacher- directed activities	-0.65
	Rules	-0.61
	Concentration 2: Self-directed activities	-0.61
	Actions	-0.61
	Cultural awareness	-0.07

“Classroom behaviour” refers to the skills which help children maintain concentration and the ability to follow school rules and timetable, as well as to broader cultural awareness (i.e., the understanding that others may have different ways of life and that this should be respected). The “self-confidence” factor describes a child’s independence, social skills and ability to maintain relationships with peers and adults within the school and broader community.

Further analysis of the Self-Confidence and Classroom Behaviour subscales was conducted separately for each scale.

Self-Confidence subscale analysis: This scale consists of 6 items (Table 2). First, we conducted a dimensionality study for the subscale using the same approach as for the total PSED scale, namely PCA of the standardized residuals. The subscale is unidimensional; therefore, we conducted a series of analyses aimed at investigating the psychometric quality of the items and of the subscale as a whole. The results indicate that all items in our subscale fit the model and have display satisfactory psychometric quality. The reliability coefficient alpha is 0.80 for the subscale.

Lastly, we investigated the quality of response categories. In the RSM context, a set of criteria was posited to verify where the rating scale functioned usefully and diagnose where it malfunctioned (Linacre J.M., 2002). Figure 2 shows the category probability curves for the scale,

with technical details omitted. We see that the thresholds are ordered and that each response category is the most probable category for some part of the respondents. This finding implies that as an individual moves up the ability continuum, each category, in turn, becomes the most probable response.

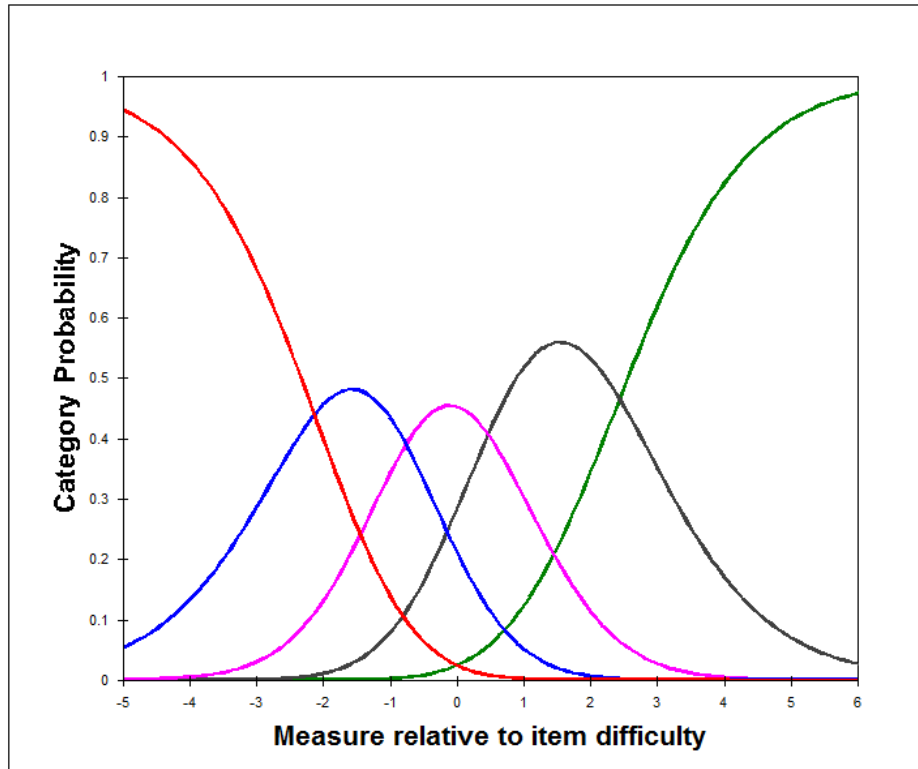


Figure 2. Category probability curves for the rating scale (Self-confidence subscale)

To sum up, the evidence suggests effective functioning of the categories in the Self-Confidence subscale.

Classroom Behaviour subscale analysis. The Classroom Behaviour subscale consists of five items with negative loadings (Table 2). All steps of this subscale analysis were similar to those in the analysis of the Self-Confidence subscale. The subscale is unidimensional, all items fit the model, and all categories function effectively. The reliability coefficient alpha is 0.86 for the subscale.

To sum up, the IRT analysis of the PSED scale demonstrates that the scale consists of two subscales: Self-Confidence and Classroom Behaviour. Both subscales exhibit good psychometric properties and can be used for estimation of students.

The estimation of students' measures was conducted separately for each subscale, resulting in two different estimates for each individual: a measure of Self-Confidence ability, and a measure of Classroom Behavior. Table 4 presents descriptive statistics of students' measures (in logits). (For each subscale, the mean of item difficulties is set equal to 0 to fix the scale origin).

Table 4. Measures of Self-confidence ability and Classroom behavior

	Mean	Standard Deviation	Min	Max
Self-confidence	1.26	1.31	-3.96	4.19
Classroom behaviour	0.67	1.93	-4.60	4.88

3.3. Clusters of cognitive and noncognitive development

Table 5 describes the correlations among all scales and sub-scales that we used in cluster analysis. The parameters used to establish groups of students appeared to be correlated linearly (sig. 0.05). The linear association between cognitive scales is 0.6, and the association between noncognitive scales is 0.57. Correlation between cognitive and noncognitive parameters is not as strong (0.24-0.29), but is still statistically and substantively significant.

Table 5. Correlation of parameters used for cluster analysis

	Mathematics	Reading	Self-confidence	Behaviour
Mathematics	1	0.6	0.25	0.24
Reading	0.6	1	0.29	0.25
Self-confidence	0.25	0.29	1	0.57
Classroom Behaviour	0.24	0.25	0.57	1

To conduct cluster analysis, all scores were converted from logit scales to z-scores. Five groups of first graders with different patterns of cognitive and noncognitive development were established (using cluster analysis). Final cluster centers were achieved at the 8th step.

Table 6. Final Cluster Centers (z-scores)

	Cluster				
	1	2	3	4	5
Mathematics	0.11	0.24	-0.65	-1.09	1.55
Reading	0.28	0.39	-0.74	-1.26	1.25
Self-confidence	1.39	-0.38	-0.10	-1.05	0.36
Classroom Behaviour	1.13	-0.45	0.15	-1.11	0.52
Number of students (% of cases)	214 (18%)	403 (33%)	270 (22%)	164 (14%)	151 (13%)

Cluster 1 - Sociable

Students in this cluster have high or medium-high levels of noncognitive skills and, on the contrary, medium-to-low levels of reading and math abilities. They can be described as self-confident children, who feel quite well in school environment, do not experience difficulties in maintaining contacts with peers and adults, and demonstrate a medium level of school performance. Sometimes they fail in complex tasks which require concentration (because behavioral measures are lower than their self-confidence), both self- and teacher-directed, especially when they are tired or don't feel well.

When working with these children, teachers can employ team-based tasks, using these students' communication strengths to develop their behavioural and cognitive abilities.

Cluster 2 – Disorganised

Students in this cluster display medium levels of cognitive development and poor results on noncognitive scales. This means that they perform moderately well in class, but still do not adjust to the school environment, do not make friends, and do not feel comfortable at school.

Kids from this cluster have trouble following school-established rules, and are not adjusted to the school environment. Communication skills and self-confidence are also low. These students have problems interacting with other children in class and even with their teacher.

When working with pupils from this cluster, the teacher should pay more attention to communication and self-regulation skills. They should receive small tasks (large ones can scare them) so that they do not lose interest in a given subject. Teachers should pay attention to these students' motivation levels: negative feedback can make them lose interest, deepening their problems.

Cluster 3 – Active

Students in this cluster demonstrate medium levels of noncognitive skills and medium-to-low levels of cognitive abilities. These students are quite sociable, but not properly adjusted to the school environment. When working with these students, teachers should pay attention to their math and reading results, and rely on these children's ability to follow rules and respect discipline. Teachers can help these students by fostering their communications skills.

Obviously, these children feel quite comfortable in an environment which they are accustomed to, but lose self-control when facing new people or circumstances. To assist with their noncognitive development, teachers should foster their communication skills and

independence. These children need special attention and support in situations which are beyond their daily life. They may experience difficulties in team work and need assistance in solving group tasks.

Cluster 4 - Latent

Students in this cluster display low levels in all the characteristics assessed. Teachers should give these students as much attention as possible. Trusted channels of communication can help teachers establish relationships with children from this group and help them in their studies.

Children from this group may require individual work with a teacher. Teachers should give these children positive feedback on every-day tasks that they perform correctly, even the simplest ones. It is desirable for the pupils to be involved in classroom activities (for example, helping the teacher from time to time in handing out printed materials). The teacher should be very careful when assigning these students to teams, and give them tasks suitable for their abilities.

Cluster 5 - Stellar

Students from this cluster display the highest levels of cognitive development among all the groups. Noncognitive measures are also relatively high, but lower than in cluster 1. When working with this group, attention to both noncognitive factors should be given equally.

Students from this group are quite independent. They may tend to lead groups of children, friends, or classmates, but it could be difficult for them to coordinate their positions and plans with others and make compromises, which may become grounds for conflict. Adults should instruct these students on how to work with peers, first in pairs and then in small groups. Assignments should be difficult, and include a need to find agreement or distribute tasks among group members. It is important that children have an opportunity to change roles: everyone should try the role of leader, performer, consultant, and person receiving help from others. In the first grade, pupils from this group should join different workgroups to train their communicational skills with different people.

4. Family background and preschool experience

In our study we also aim at describing defined groups using background information including family social and economic status, children's preschool experience, and so on. In this section we will try establish what contextual variables can be considered important predictors of students' cognitive and noncognitive development. This contextual information was selected

according to chi-square and Cramer's V (see table 7) and forms three domains – demographical, economical and pre-school experience.

Table 7 presents a description of the chosen contextual variables, with the percent of cases in brackets in each cluster.

Table 7. Background description for clusters

Background information (case), Cramer's V		Cluster (% of cases)				
		1	2	3	4	5
Socio-demographical	Gender (boys), 0.28	27.4	57.0	38.1	67.7	60.0
	Number of books at home (more than 100), 0.16	24.1	31.4	16.8	18.2	43.0
	Mother's education (higher), 0.15	58.4	62.8	50.0	32.8	68.7
Economical	Income (more than 20 000 RUR per month), 0.11	83.8	85.3	77.2	69.5	81.1
Pre-school experience	Nursery in the pre-school year (yes), 0.11	91.3	87.4	86.6	79.3	92.0
	Special training before school (yes), 0.17	24.1	31.4	16.8	18.2	43.0

The background of children from the fourth cluster differs from that of children from other groups. On average, they have fewer books at home, most of them did not receive any special training before school, their mothers in general did not receive a higher education, and so on. In cluster five, this picture is largely reversed. This group features the highest percentage of books at home, the greatest proportion of mothers with higher educations, and most children have received special training and attended kindergarten. So we can assume that the roots of the inter-cluster differences are in the children's background and pre-school experience.

Discussion

In our study we classified students from Tatarstan into five groups which differ from each other in terms of both cognitive and noncognitive development. These groups also differ by their socio-demographic characteristics and pre-school experience. In this paper we also proved that noncognitive development, as measured by the iPIPS PSED questionnaire, is a two-dimensional construct. It consists of 2 subscales: communication and classroom behavior. Although this

finding was not our primary focus, it provides us with a better understanding of the instrument and allows us to reduce the number of variables in the analysis.

Our results can serve as the basis for recommendations to teachers on how to deal with children from each group. Because children from the 5 groups in our study display different characteristics, different pedagogical strategies are required to help different children take full advantage of their first year of school.

Though we still need further research on children with the lowest (cluster 4) and the highest (cluster 5) results, we can state that each group requires an individualized approach to ensure children experience all the potential benefits of education.

References

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1993). First-grade classroom behaviour: Its short- and long-term consequences for school performance. *Child Development*, 64, 801–814

Birch, S. H., & Ladd, G. W. (1997). The teacher–child relationship and children’s early school adjustment. *Journal of School Psychology*, 35, 61–79.

Bond Tr.G., Fox C.M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. – New Jersey.: Lawrence Erlbaum Associates, Inc., Publishers, 2001, 255 p.

Denham, S. A., Bassett, H. H., Thayer, S. K., Mincic, M. S., Sirotkin, Y. S., & Zinsser, K.(2012). Observing preschoolers’ social-emotional behavior: Structure, foundations, and prediction of early school success. *Journal of Genetic Psychology*, 173,246–278.

Durlak, J. A. et al. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405-432.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children’s school outcomes through eighth grade. *Child Development*, 72,625–638.

Hindman, A. H., Skibbe, L. E., Miller, A., & Zimmerman, M. (2010). Ecological contexts and early learning: Contributions of child, family, and classroom factors during head start, to

literacy and mathematics growth through first grade. *Early Childhood Research Quarterly*, 25, 235–250.

Hogan, J., and Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88(1), 100-112.

Hawker D., Kardanova E. (2014). Base-line assessment of pupils on the entrance to the primary school and the measurement of their progress during the first year: iPIPS international study. / In: Trends in the development of education: what are the effective school and nursery? Moscow: “Delo”, pp. 311-320 (in Russian).

Hurtz, G. M., and Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology*, 85(6), 869-879.

Ivanova A. E., Nisskaya A. K. (2015). Starting diagnostics on entrance to the primary school and progress evaluation during the first year. *School technologies*, 2, p. 161 – 168.

Jain A. K. Data clustering: 50 years beyond K-means //Pattern recognition letters. – 2010. – T. 31. – №. 8. – C. 651-666.

Linacre J. M. (2011). A User's Guide to WINSTEPS. Program Manual 3.71.0. <http://www.winsteps.com/a/winsteps.pdf>

Linacre J.M. (2002). Optimizing Rating Scale Category Effectiveness. *Journal of Applied Measurement*, Vol.3, № 1, 85-106

Ludlow, L.H. (1985). A strategy for the graphical representation of Rasch model residuals. *Educational and Psychological Measurement*, 45(4), 851-859.

Matthews, J. S., Cameron Ponitz, C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, 101, 689–704.

Merrell, C., Bailey K. Predicting Achievement in the Early Years: How Influential is Personal, Social and Emotional Development? URL: <http://www.oerj.org/View?action=viewPDF&paper=55>

Munafò, M. R., Zettler, J. I., and Clark, T. G. (2007). Personality and smoking status: A meta-analysis. *Nicotine and Tobacco Research*, 9(3), 405-413. doi: 10.1080/14622200701188851

Papachristou, H. et al. (2013). Higher levels of trait impulsiveness and a less effective response inhibition are linked to more intense cue-elicited craving for alcohol in alcohol-dependent patients. *Psychopharmacology*, 228(4), 641-649. doi: 10.1007/s00213-013-3063-3

Pianta, R. C., & Nimetz, S. L. (1991). Relationships between children and teachers: Associations with classroom and home behavior. *Journal of Applied Developmental Psychology*, 12, 379–393.

Smith, E. V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement*, 3:2, 205-231.

Sutin, A. R. et al. (2011). Personality and obesity across the adult life span. *Journal of Personality and Social Psychology*, 101(3), 579-592. doi: 10.1037/a0024286

Tymms, P. (1999a). *Baseline Assessment and Monitoring in Primary Schools: Achievements, Attitudes and Value-added Indicators*. London, David Fulton Publishers.

Tymms, P. B. (1999b). Baseline assessment, value-added and the prediction of reading. *Journal of Research in Reading*, 22(1), 27-36.

Tymms, P., Brien, D., Merrell, C., Collins, J. and Jones, P. (2003). Young deaf children and the prediction of reading and mathematics. *Journal of Early Childhood Research*, 1(2), 197-212.

Tymms, P., Merrell, C., Henderson, B., Albone, S., and Jones, P. (2012). Learning Difficulties in the Primary School Years: Predictability from On-Entry Baseline Assessment. *Online Educational Research Journal*.

Versteeg, H. et al. (2012). Type D personality and health status in cardiovascular disease populations: a meta-analysis of prospective studies. *European Journal of Preventive Cardiology*, 19(6), 1373-1380.

Wright B.D., Masters G.N. *Rating Scale Analysis*. Rasch Measurement. – Chicago: Mesa Press, 1979, 206 p.

Wright B.D., Stone M.N. *Best Test Design*. Rasch Measurement. – Chicago: Mesa Press, 1979, 223 p.

Brun Irina, Center of Education Quality Monitoring, Institute of education, National Research University Higher School of Economics, research assistant, E-mail: ibrun@hse.ru

Ivanova Alina, Center of Education Quality Monitoring, Institute of education, National Research University Higher School of Economics, junior research fellow, E-mail: aeivanova@hse.ru

Kardanova Elena, Center of Education Quality Monitoring, Institute of education, National Research University Higher School of Economics, head of Center of Education Quality Monitoring, PhD, E-mail: ekardanova@hse.ru

Orel Ekaterina, Center of Education Quality Monitoring, Institute of education, National Research University Higher School of Economics, research fellow, E-mail: eorel@hse.ru

Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.

© Brun, Ivanova, Kardanova, Orel 2016