

Patterns of First-Graders' Development at the Start of Schooling: Cluster Approach Based on the Results of iPIPS Project

E. Kardanova, A. Ivanova, P. Sergomanov, T. Kanonir, I. Antipkina, D. Kaiky

Received in
September 2017

Elena Kardanova

Candidate of Sciences in Mathematical Physics, Associate Professor, Head of the Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. Email: ekardanova@hse.ru

Alina Ivanova

Junior Researcher, Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. Email: aeivanova@hse.ru

Pavel Sergomanov

Candidate of Sciences in Psychology, Associate Professor, Director of the Center for Leadership in Education, Institute of Education, National Research University Higher School of Economics. Email: psergomanov@hse.ru

Tatyana Kanonir

PhD, Associate Professor, Co-Head of the Master's Program "Measurements in Psychology and Education", Institute of Education, National Research University Higher School of Economics. Email: tkanonire@hse.ru

Inna Antipkina

Junior Researcher, Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. Email: [ian-
tipkina@hse.ru](mailto:ian-
tipkina@hse.ru)

Diana Kaiky

Research Assistant, Center for Monitoring the Quality in Education, Institute of Education, National Research University Higher School of Economics. Email: dkaiky@hse.ru

Address: 20 Myasnitskaya St, 101000 Moscow, Russian Federation.

Abstract. The first year at school is crucial for determining further academic achievement. Information on the range of first-graders' abilities and needs and prediction of their educational trajectories allow the education system to facilitate adaptation to school dramatically and increase the efficiency of teaching approaches. The article presents the results of a survey of 7,778 first-graders enrolled in schools in four Russian cities—Moscow, Naberezhnye Chelny, Sevastopol and Tambov—in 2015. Cluster analysis of the data on children's cognitive skills (manifested in mathematical and reading literacy) and non-cognitive (personal, social and emotional) development produced four groups of first-graders with typical patterns of development assessed at school entry. Combining cognitive and non-cognitive indicators creates an additional opportunity for understanding the peculiarities of child development in elementary school and allows for building a "gallery" of four typical first-grader profiles. The findings can be used in helping teachers choose and customize education programs and other means of supporting children during adaptation.

Keywords: first graders, iPIPS, baseline assessment, progress measurement, cognitive development, non-cognitive development, cluster analysis.

DOI: 10.17323/1814-9545-2018-1-8-37

The research was supported by a grant from the Russian Science Foundation (Project No. 16-18-10401 "Predicting Teaching Effectiveness in Elementary School").
Translated from Russian by I. Zhuchkova.

Starting school is a critical period in a child's life. Academic achievement and overall wellbeing at school will largely depend on how well the child and the school adapt to each other [Margetts 2009, Alexander, Entwisle, Dauber 1993, Domitrovich et al. 2017; Zuckerman, Polivanova 2012]. In order to provide relevant support to children at this crucial point, it is necessary to understand what skills and pre-developed patterns of cognitive and non-cognitive development first-graders bring to school, on the one hand, and what the school can offer to enhance their academic achievement and support their overall wellbeing, on the other hand.

For the first time in the history of Russian education, the Federal State Education Standard of Primary Education includes legally defined personal outcomes as a goal of general education. Of course, personal, social, and emotional skills as important aspects of development and later achievement in life have been objects of measurement for quite a while [Durlak et al. 2011; Poropat 2009; OECD2015]. However, bringing the issue to the legislative field added fuel to the debate over competent (valid and reliable) assessment of personal, social, and emotional skills and displayed the shortage of modern instruments that would be structured intricately enough (to fit the structure of the object measured), have a high predictive validity, and, most valuably, be effective in helping teachers choose and customize education programs and other means of supporting children's adaptation to school. In fact, teachers themselves need reliable and accurate diagnostic tools to adapt to the development peculiarities of their students.

Once the cognitive, personal, social, and emotional skills have been assessed, there will be enough information to understand the specific aspects of children's situations, their capabilities and needs. Research shows that cognitive and non-cognitive development (personal and socioemotional skills) at the early stages of education is a powerful predictor of further academic performance and success in adult life [Kautz et al. 2014]. This is not to say that preschool development, child's adaptation skills and parenting practices should be underestimated.

Evaluating preschool and elementary school-aged children is fraught with particular difficulties. First of all, children at this age cannot stay focused for a long time. Second, many of them cannot read when they first come to school, so general questioning techniques will hardly be applicable here. Third, any evaluation should take into account the size of vocabulary and the overall level of language development to make sure the child understands the instructions and tasks [Merrell, Tymms 2016]. That is why the content and forms of child assessment at this period of development have certain limitations [Slentz 2008].

There are not so many school-entry assessments with established psychometric properties in global psychodiagnostics that would be suitable for large-scale testing. A proportion of the educational community is convinced that children should be assessed through ob-

servations and adult surveys, whereas others believe that even the youngest can and should be interviewed. For instance, the well-known *Early Development Instrument* (EDI) represents a questionnaire completed by preschool teachers that measures the level of child development across various domains [Janus et al. 2007]. The *Early Childhood Environment Rating Scales* (ECERS) and the *School-Age Care Environment Rating Scales* (SACERS) are based on observations and structured assessments of environment, time management, student-teacher interactions, and education quality [Harms, Clifford, Cryer 2015; Harms 2013]. These are indirect assessment tools which describe the environment of child development using the observations made by the adults involved.

This article is based on a study that uses the *international Performance Indicators in Primary School (iPIPS)*, which allow for evaluating cognitive (basic reading and mathematical literacy) and non-cognitive (personal, social, emotional) development of children at school entry. Unlike the questionnaires mentioned above, the iPIPS were designed specifically to assess first-graders directly. Assessors work with children in individual one-to-one sessions, which allows them to measure the level of their knowledge and skills at school entry as well as the progress they make by the end of their first year at school [Ivanova, Nisskaya 2015]. The instrument offers computer-delivered interactive assessment using an adaptive algorithm so that assessment is not too difficult for each individual child and does not get them tired or demotivated. The iPIPS also include a questionnaire for teachers and parents on how children grow and develop.

The iPIPS have been widely used abroad for both practical and research purposes, allowing teachers to get to know their first-graders better. Experts use the iPIPS to collect important statistics to study various educational situations. In particular, the instrument has been applied to measure the role of preschool education in later attainment [Tymms, Merrell, Henderson 1997], identify children's individual and group progress [Tymms, Merrell, Henderson 2000], compare first-graders' progress across countries [Copping et al. 2016; Tymms, Merrell, Wildy 2015], and identify the high-risk groups [Tymms et al. 2012].

Russian researchers have used iPIPS to explore regional differences in first-graders' skills measured at school entry [Ivanova et al. 2016], analyze progress in reading literacy [Antipkina, Kuznetsova, Kardanova 2017], assess the role of phonological awareness [Kuzmina, Ivanova, Antipkina 2017], and monitor the socioemotional development of children and their behavioral patterns [Orel, Ponomareva 2018; 2016]. This article presents a large-scale iPIPS-based empirical study to describe a 2015 cohort of first-graders, analyze the patterns of combinations of cognitive and non-cognitive (socioemotional) skills observed in first-graders at school entry, and create a "gallery" of four typical first-grader profiles.

1. Criteria for identifying levels in cognitive and non-cognitive development

In assessing cognitive skills at school entry, the focus should be placed on the skills that are directly related to the first-grade curriculum and capable of predicting later achievement. Greg J. Duncan and his colleagues [Duncan et al. 2007] used six longitudinal studies conducted on representative samples in different countries to demonstrate that mathematical and reading literacy measured at school entry are the best predictors of academic performance at the end of elementary school. Those studies found no differences related to gender or socioeconomic status. Canadian researchers later reproduced the data analysis strategy used by Duncan and his colleagues on a sample of Quebec children and obtained very similar results, confirming the high prognostic value of initial reading and mathematical skills for progress by the middle of elementary school [Pagani et al. 2010]. They also added fine motor skills as a predictor to the model and revealed, in contrast to Duncan's findings, noticeable gender differences in its significance. A number of minor studies also confirmed the prognostic role of early reading [Müller, Brady 2001] and mathematical [Manfra et al. 2017; Jordan et al. 2009] skills, the assessment of which should take into account the socioeconomic status and other relevant contexts such as ethnicity, type and location of school, etc.

Ample research has confirmed over and over again that personal and socioemotional skills affect various aspects of life at any age [OECD2015; Durlak et al. 2011]. The lack of such skills poses risks to child development [Domitrovich et al. 2017]. Important and distinctive features of all such skills include: (a) conceptual independence from cognitive competencies, (b) overall benefit to an individual in cases where such characteristics are distinctly manifested, (c) relative temporal stability provided there is no external interference, (d) possibility of changing as a result of interference, and (e) situational manifestations [Duckworth, Yeager 2015].

Psychopedagogical studies involving elementary pupils stress the importance of a comprehensive approach to assessments [Merrell, Tymms 2011]. Measuring the cognitive and non-cognitive development of first-graders provides for an integrated evaluation of children's skills, increasing the opportunities for analysis and interpretation of results and expanding the choice of support tools.

Russian researchers have studied some of the aspects of readiness for school and adaptation of first-graders, in particular their physiological readiness [Paranicheva, Tyurina 2012; Gritsinskaya et al. 2003], overall development and ability to follow the rules [Kovaleva et al. 2012; Salnikova, Tkachenko 2012], motivation for learning [Gani, Gani 2009], and adaptation as a challenge in life [Gagay, Grineva 2013]. However, with some minor exceptions [Kovaleva et al. 2012], most of such studies use samples of no more than a few hundred children.

This paper investigates (i) the level of development of cognitive and non-cognitive skills in Russian children at school entry and (ii) the

possibility of identifying empirically distinct groups (clusters) of children with similar profiles of cognitive and non-cognitive development.

2. Method

2.1. Description of the instrument and the research procedure

Cognitive and non-cognitive development of children was evaluated using the iPIPS instrument originating from Durham University, Great Britain [Tymms 1999]. The instrument allows for assessing children at school entry and monitoring their individual progress made during the first school year. This article only uses the baseline assessment results.

The iPIPS were selected as a research method for specific reasons. First, they are well in line with the latest achievements in global research on assessments. Second, this is a high-quality, world-recognized standardized instrument with established psychometric properties and validity. Third, they use a special measurement technique to assess individual progress made by a child throughout the first year of schooling. Finally, the iPIPS are designed as a computer adaptive assessment, which makes it possible to evaluate each individual child as carefully and accurately as possible, avoiding bias.

The National Research University Higher School of Economics (HSE) in cooperation with Durham University developed a Russian version of the iPIPS in 2013–2014 [Hawker, Kardanova 2015; Ivanova, Nisskaya 2015]. Given that the school starting age differs in Great Britain and Russia (Russian first-graders are on average two years older), the instrument had to undergo an essential adaptation. In particular, new tasks were designed to match the age and cultural contexts of the development of Russian children. The resulting Russian version of the iPIPS has been successfully applied in Russian schools.

The iPIPS instrument is unique in that it offers a holistic approach, assessing not only cognitive but also socioemotional development of a child. In addition, it makes use of the contextual information on the conditions of children's preschool life and development, their families, and adopted child-rearing practices. Special attention is paid to school teaching methods.

The cognitive development module consists of the following measures:

- Handwriting (assessment of writing skills)
- Vocabulary (passive vocabulary and knowledge of high-frequency words)
- Phonological awareness
 - Word repetition (familiar and unfamiliar words as well as non-words)
 - Rhyming (supported by pictures of rhyming words)
- Ideas about reading:
 - Text structure (knowledge of capital and lowercase letters, the notions of the beginning and end of a sentence, periods, etc.)
 - Letter knowledge

- Word reading (written word recognition)
- Short story reading (text decoding)
- Reading comprehension
- Ideas about mathematics:
 - Simple counting (counting objects in a picture)
 - Simple adding and subtracting (supported by pictures)
 - Number knowledge
 - Mathematical problems (logical problems, problems with and without symbols, word problems, contextual problems)

First-graders are not required to possess knowledge in any of the domains listed above—it is the school's objective to develop them relevant skills in these. In reality, however, many children have already developed some reading, counting and writing skills by the time they first come to school, so it becomes vital to know the level of each child's development at school entry in order to make any justified assessment of their progress.

The assessment is designed as an exploratory game to ensure as safe and comfortable environment for a child as possible. An interviewer, whose role may be played by a pre-briefed school teacher, psychologist or counselor, assists each individual child in the computer adaptive assessment, which takes about 20–30 minutes (depending on the child's level of development). The child follows the software instructions voiced by a professional speaker and performs a series of entertaining tasks, while the interviewer records the answers. The versatility of the tasks prevents children from getting bored or tired. The semi-adaptive algorithm selects items that match each individual child's abilities.

Non-cognitive development is assessed using a questionnaire completed by teachers from their knowledge of children gained through day-to-day interaction and observation. Teachers are asked to evaluate every child using a five-point scale in all items. A descriptor is provided for each point on the scale so that teachers could choose a suitable example and decide which descriptor provides the closest match for their observation.

The following personal and socioemotional skills are assessed by the iPIPS: adjusting to school environment; independence and self-help skills; confidence and participation in group activities; ability to concentrate on the task under the teacher's guidance and independently; ability to think through one's decisions and avoid acting impulsively; ability to follow the rules and behave in accordance with the established code; communication skills (level of speech development and such socio-cognitive aspects of communication as ability to ask questions, hear out what others have to say, and wait for one's turn to speak); ability to interact with adults (to approach adults confidently and fearlessly and behave accordingly and naturally) and peers (to establish and maintain friendly relationships); awareness of the fact

Table 1. Sample description

City	No. of schools	No. of classes	No. of students	Proportion of students from schools of higher status	Percentage of girls
Moscow	16	140	3,173	N*	48%
Naberezhnye Chelny	41	94	2,379	18%	52%
Sevastopol	22	59	1,283	38%	49%
Tambov	5	37	943	39%	49%

*The assessment in Moscow involved school complexes to which the notion of "advanced type" is inapplicable.

that the way other people live may differ from what is accepted in the child's family, and having respect for such differences. These aspects of socioemotional development were selected due to their empirically-proved significance for school adaptation, friendship ties, and academic achievement [Merrell, Tymss, Buckley 2015; Spence 1987].

2.2. Sample

The empirical basis of the research consists of the iPIPS data obtained in 2015 on an extensive sample of first-graders in four large cities of Russia: Moscow, Naberezhnye Chelny, Tambov, and Sevastopol¹.

The iPIPS instrument may be used either to analyze system performance or as a means of assessing the individual progress of school-children. In cases where the goal of the research is to analyze the education system at the level of a city or region, the sample should be large enough and representative of the respective city or region. If, however, research is aimed at thorough assessment of individual students in particular schools, the sample should embrace all first-graders at those schools and does not have to be representative.

Table 1 describes the sample of this study, which involved schools of different types (regular schools and higher-status schools such as gymnasiums, and specialized schools) located in different parts of the aforementioned cities. In Sevastopol and Naberezhnye Chelny, the sample unit was a class of students which was picked arbitrarily from the first-grade cohort of a selected school in a few municipal districts. In Moscow and Tambov, the sample unit was a school, so all the first-graders of a particular school were assessed.

¹ We are grateful to the Department of Education of Moscow, the Ministry of Education and Science of the Republic of Tatarstan, the Republican Center for Monitoring of Education Quality of the Republic of Tatarstan, the Department of Education of Sevastopol, the Department of Education and Science of Tambov Oblast, and the Institute of Advanced Educator Training of Tambov Oblast for their assistance in the research.

Even though the final sample features a great diversity of school types, it is not representative in every sense of the word (this is dwelled upon in the discussion part of the article).

2.3. Analysis Item Response Theory (IRT) was applied to build scales for the basic cognitive and non-cognitive skills. In particular, one-parameter dichotomous Rasch model [Wright, Stone 1979] was used to translate children's raw scores into assessments of their skills on cognitive scales (in mathematics, reading, phonological awareness, and vocabulary), while the non-cognitive components were assessed using the Rating Scale Model (RSM), an extension of the dichotomous Rasch model for Likert-type scales [Wright, Masters 1982]. Prior to assessment, psychometric characteristics of students were measured by conducting psychometric item analysis, dimensional analysis, and reliability analysis, and then building scales using Rasch models. Win-steps software was used to perform psychometric item analysis and assess item and ability parameters [Linacre 2011].

k-means clustering was used to group children based on their levels of cognitive and non-cognitive development. A correlation matrix of all the variables was constructed prior to clustering. In the cluster analysis, data is grouped based on response patterns. This particular study aimed at identifying groups of first-graders with similar levels of cognitive or non-cognitive development.

k-means clustering is one of the most intuitive and popular methods of response pattern assessment and grouping, yet its numerous limitations should be born in mind [Jain 2010]. For instance, unlike hierarchical cluster analysis, *k*-means clustering implies determining the optimal number of clusters prior to analysis. There are various ways of determining the number of clusters of data analyzed, like logical representation [Ibid.]. We expected four or five clusters to be concealed in the data: one cluster would reveal high performance in all the scales, low performance in all the scales in another, and two or three more clusters would presumably represent various combinations of scores reflecting dominant development of cognitive or non-cognitive skills. In order to test this hypothesis, attempts were made to divide the sample into three, four, and five clusters. Each of the solutions was discussed in terms of reasonable interpretability and balanced clustering.

Cluster analysis has some disadvantages, in particular it always provides a statistical classification even if there are no objective grounds for classifying. For this reason, stability of the resulting cluster solutions had to be verified in order to prove the internal validity of clustering results. First of all, the sample was randomly divided into two groups. Each of them was subjected to *k*-means clustering, and the results were compared to the master sample. In addition to that, temporal clustering stability and the stability of clusters for samples in other cities were verified using the iPIPS-2014 data. All the new clus-

ter solutions were found to be similar to the solution obtained for the master sample.

The resulting clusters (groups of students) were described using the following indicators:

- 1) Sociodemographic characteristics (gender, mother's education, number of books at home, preschool attendance)
- 2) Average score in phonological literacy and vocabulary

Chi-squared test and Cramer criterion were used to assess correlations between these parameters and the resulting groups of children. Differences among the clusters were also verified using standard statistical procedures for each indicator. Z-tests were applied to measure differences among the clusters in gender, mother's education, number of books at home (more or fewer than 100), and preschool attendance (nominal variables). Analysis of variance (ANOVA) was used to analyze the differences in vocabulary and phonological awareness among the groups.

3. Results

3.1. Assessment of cognitive development

Before going on to the secondary data analysis, it is necessary to make sure that the data is of high quality. For this purpose, psychometric analysis of initial assessment results was performed, which showed that all cognitive iPIPS items and the assessment as a whole have good characteristics. All the scales (mathematics, reading, phonological awareness, and vocabulary) are essentially unidimensional, and all the items comply with the model used. Cronbach's alpha is high for all the scales: 0.92 for mathematics, 0.97 for reading, 0.78 for phonological awareness, and 0.84 for vocabulary². It thus follows that the scales can be used to measure cognitive skills of children at school entry.

Every child was introduced to four baseline assessments: in mathematics, reading, phonological awareness, and vocabulary. Table 2 presents descriptive statistics of students' results expressed in the logit scale³. Mean item complexity is set to zero for each scale as a reference point.

Below is a brief interpretation of low, medium and high scores in each of the scales.

Mathematics. Children who gained the minimum score in mathematics can name numbers up to ten and cope with the simplest calcu-

² For more details on the results of the psychometric analysis of the math scale, see [Ivanova et al. 2016].

³ Logit is a special unit of measurement used in modern Item Response Theory.

Table 2. Results of assessing the cognitive development of children at school entry

Indicator	Mean	Standard deviation	Minimum	Maximum
Mathematics	-0.29	1.83	-7.44	7.24
Reading	0.94	2.41	-7.35	7.14
Phonological awareness	1.48	1.52	-4.84	4.33
Vocabulary	0.79	1.79	-5.41	5.36

lations supported by object pictures. Children in the mid-scale can name one- and two-digit numbers (with occasional mistakes), do simple calculations, solve easy logical and word problems, and do sums with mathematical symbols that do not cross the next ten. Children with the highest scores in mathematics can name multidigit numbers, do fairly complex additions and subtractions with two-digit numbers, and solve “real-life” problems.

Reading. A minimal score means that a child does not possess the basic reading skills, including letter knowledge. Average scores show that children know letters and can read but cannot always understand what they have just read. Children who score the highest have mastered the basic reading skills quite well, they can read a few successive texts without mistakes and comprehend what they have read.

Vocabulary. Children with the minimum scores in the vocabulary scale know and can recognize relatively simple, frequently used words. Average scores indicate that a child knows and can recognize less common and more complex words. Finally, the top scorers have a pretty sizeable vocabulary and are able to recognize some narrowly-specialized and uncommon words, such as *saxophone* or *silhouette*.

Phonological awareness. Although the phonological awareness scale is represented by relatively few tests of two types, it is good at differentiating students. A minimum score in this scale means that a child is only able to repeat simple and well-known words aloud. Those who score in the middle are able to repeat non-words, understand and recognize some unsophisticated rhymes. Children scoring maximum can repeat complex or unfamiliar words and non-words and rhyme words easily.

3.2. Assessment of non-cognitive development

Psychometric analysis shows that the non-cognitive development questionnaire includes two scales, conventionally referred to as “confidence” and “classroom behavior”. Both are one-dimensional, with tasks in line with the model used, good psychometric characteris-

Table 3. **Results of assessing the non-cognitive development of children at school entry**

Scale	Mean	Standard deviation	Minimum	Maximum
Confidence	1.69	1.70	-5.34	5.59
Classroom behavior	0.98	2.33	-6.17	6.16

tics, and adequate response categories. Cronbach's alpha is 0.84 for confidence and 0.88 for classroom behavior. Therefore, both scales have good psychometric properties and can be used for assessment⁴.

The classroom behavior scale describes behavioral skills: the ability to focus on the task and follow the school rules and schedules, and the level of cultural awareness, i. e. understanding that other people may have different lifestyles that should be respected. Minimum scores in this scale demonstrate that a child gets distracted a lot both under the teacher's guidance and when working by oneself, violates the established classroom rules, and acts impulsively. The highest score is associated with ability to focus for quite a while (about 15 minutes), staying within the established rules, and understanding that there are diverse cultural traditions that may be different from what is accepted in the child's family.

The confidence scale describes children's independence and autonomy skills, social skills for maintaining relationships with other people, age-mates and adults, at school and in broader social contexts. Children who score low on this scale feel uncomfortable in the school environment, miss their parents, need to ask for help with buttons or using the toilet, find it difficult to make friends at school, address adults inadequately or hesitate to approach them. Maximum scores are gained by children who are well-adapted, independent, possess necessary self-care skills, and communicate adequately with both adults and peers.

Scores in both scales, confidence and classroom behavior, have been obtained for every child. Table 3 contains the descriptive statistics of students' results expressed in the logit scale. Mean item complexity is set to zero for each scale as a reference point.

3.3. Preliminary analysis of children's differences in cognitive and non-cognitive development

The next step was to analyze how children's school-entry profiles varied depending on a number of basic sociodemographic and institutional factors: gender, parental (mother's) education, number of books at home, and preschool attendance. Tables 4–7 present the ef-

⁴ For more details on the results of psychometric analysis of the non-cognitive development questionnaire scales, including questionnaire size analysis, see [Orel et al. 2016].

Table 4. Results of girls and boys

Indicator	No.	Mean		Standard deviation		Effect size
		Girls	Boys	Girls	Boys	
Mathematics	7,753	-0.15	0.14	0.90	1.07	-0.29
Reading	7,753	0.08	-0.07	0.98	1.01	0.15
Vocabulary	7,753	-0.09	0.09	1.00	0.99	-0.18
Phonological awareness	7,753	0.09	-0.09	0.99	1.00	0.19
Confidence	6,233	0.21	-0.20	0.98	0.98	0.41
Classroom behavior	6,233	0.29	-0.28	0.92	0.99	0.58

Table 5. Correlations between children's results and their mothers' education

Indicator	Mean		Standard deviation		Effect size
	No college degree	College degree	No college degree	College degree	
Mathematics	-0.25	0.14	0.97	0.98	0.39
Reading	-0.26	0.14	1.02	0.94	0.40
Vocabulary	-0.22	0.13	1.03	0.96	0.35
Phonological awareness	-0.24	0.13	0.96	0.99	0.38
Confidence	-0.13	0.08	1.01	0.99	0.21
Classroom behavior	-0.15	0.09	1.02	0.99	0.24

fects of belonging to a particular group, identified based on the factors specified, on the child's performance in every scale. The effect size is estimated as the proportion of standardized mean difference in the respective variable for two groups to the overall standard deviation for the group. The effect size of 0.2 is considered as insignificant (small), 0.5–0.8 as medium, and above 0.8 as large [Cohen 1988].

Table 4 breaks the figures down by gender. Dispersion is on average slightly higher among boys, which means that boys are more likely to score extremely high or extremely low. The gender gap is particularly wide in mathematics.

The effect of gender on reading, phonological awareness and vocabulary is insignificant. However, even though the effect is small in mathematics, it still proves that girls fall behind noticeably, by almost one third of a standard deviation. The medium-sized effect of gender in non-cognitive domains shows that teachers consider both behavior and confidence to be better developed in girls than in boys. The dif-

Table 6. Children's results depending on the number of books at home

Indicator	Mean		Standard deviation		Effect size
	Fewer than 100 books	100 books or more	Fewer than 100 books	100 books or more	
Mathematics	-0.10	0.18	0.98	0.99	0.28
Reading	-0.11	0.19	0.98	0.97	0.30
Vocabulary	-0.14	0.26	1.00	0.95	0.40
Phonological awareness	-0.10	0.18	0.98	0.99	0.28
Confidence	-0.03	0.07	0.99	1.00	0.10
Classroom behavior	-0.04	0.10	0.99	1.02	0.14

Table 7. Results of children with and without preschool education experience

Indicator	Mean		Standard deviation		Effect size
	Preschool attendance	No preschool attendance	Preschool attendance	No preschool attendance	
Mathematics	-0.02	0.00	1.05	0.98	0.02
Reading	-0.02	-0.01	1.13	0.97	0.02
Vocabulary	-0.02	0.00	1.10	0.99	0.02
Phonological awareness	-0.07	0.00	1.03	0.99	0.07
Confidence	-0.06	0.01	1.00	1.00	0.07
Classroom behavior	0.17	-0.02	0.97	1.01	-0.18

ference between assessments in the non-cognitive scales equals 0.5 of a standard deviation or more.

Table 5 demonstrates how students' performance correlates with whether or not their mothers have a college degree.

The effect of parental education is small yet consistent in all the domains. Although it is the smallest on confidence, children of college-educated mothers are assessed on average by 1/5 of a standard deviation higher than those whose mothers have no college degree. The effect is the strongest in the cognitive scales: reading, mathematics, vocabulary, and phonological awareness. Children from well-educated families outdo their classmates by more than 1/3 of a standard deviation.

Similar patterns are observed for the number of books at home (Table 6). Possession of educational resources in the form of large home libraries has a significant positive effect on all the indicators of

cognitive and non-cognitive development, correlating the strongest with reading literacy and vocabulary.

Preschool attendance appears to have no statistically significant effect (Table 7). Perhaps this is due to the great difference in the size of the respective subsamples: about 90 percent of children attended a kindergarten during the last year before school. However, the sample of those who did not attend a preschool institution was rather large (741 students).

As we can see, primary data analysis revealed quite evident discrepancies in the levels of cognitive and non-cognitive development of children. A deeper analysis will allow for identifying student groups statistically, based on the characteristics revealed.

3.4. Student clusters

Indicators of both cognitive (scores in mathematics and reading) and non-cognitive (performance in confidence and classroom behavior) development were selected for clustering. Table 8 presents the results of correlation analysis that demonstrate the main correlations among the variables used in further analysis.

In the process of analysis, attempts were made to separate children into three, four, five, and six clusters based on combinations of the indicators of their cognitive and non-cognitive development. Each of the solutions obtained was analyzed from the viewpoint of reasonable interpretability and balanced clustering, and the four-cluster solution proved to be the optimal one. The cluster analysis results are given in Table 9.

In Table 10, the resulting clusters are broken down by children's sociodemographic characteristics as well as their phonological awareness and vocabulary.

Brief profile descriptions of the identified student groups can be provided based on the results of cluster and descriptive analyses.

Cluster 1: High levels of cognitive and non-cognitive development

Students of this group perform well in all the domains. They out-run the sample mean by almost one standard deviation in mathematics and reading, and their performance in non-cognitive development is even better in standard deviation units (confidence 1.20, classroom behavior 1.08).

Around 90 percent attended a kindergarten during the year before school. Children in this group score the highest in phonological awareness and vocabulary. The cluster also features the highest proportion of college-educated mothers (74%), and 43 percent of the families have more than 100 books at home.

Cluster 2: Average and high levels of cognitive development and a lower-than-average level of non-cognitive development

Children in this group have average and even quite good scores in reading and mathematics (0.61 and 0.53 in SD units, respectively)

Table 8. Mutual correlations among the variables

Indicator	Mathematics	Reading	Confidence	Classroom behavior
Mathematics	1			
Reading	0.58**	1		
Confidence	0.29**	0.30**	1	
Classroom behavior	0.25**	0.29**	0.67**	1

Table 9. Cluster analysis results

Indicator	Cluster			
	1 (N=1,224)	2 (N=1,666)	3 (N=1,790)	4 (N=1,535)
Mathematics	0.93	0.61	-0.42	-0.90
Reading	0.87	0.53	-0.29	-0.92
Confidence	1.20	-0.42	0.38	-0.93
Classroom behavior	1.08	-0.39	0.45	-0.96
% of the sample	20%	27%	29%	25%

Table 10. Descriptive characteristics of the clusters

Variable	Cluster				Cramer criterion
	1	2	3	4	
Phonological awareness*	54**	53	49	44	0.30
Vocabulary*	54	53	48	46	0.20
Gender (percentage of boys)	41%	65%	36%	62%	0.25
Mother's education (percentage of college-educated mothers)	74% ₂	71% ₁	63%	47%	0.21
Number of books at home (percentage of families with more than 100 books at home)	43%	36% ₃	32% ₂	26%	0.12
Preschool attendance (percentage of children who attended a kindergarten during the last preschool year)	87% ₂₃₄	89% ₁₃₄	89% ₁₂₄	88% ₁₂₃	0.03

* For ease of comparison, the scores in phonological awareness and vocabulary were translated from the logit scale to 100-point scales with a mean of 50 and a standard deviation of 10.

** All differences among the clusters in every domain are statistically significant except those with subindex numbers indicating the clusters which differ insignificantly from the current one.

but score lower than average in confidence and classroom behavior (-0.39 and -0.42 in SD units, respectively). They also perform fairly well in phonological awareness and demonstrate extensive passive vocabularies. The cluster features a high percentage of boys, who account for about 65 percent. The socioeconomic status of such children is rather high: more than one in every three families has a large home library, and almost 70% of the mothers have college degrees. Nearly all the children in the group attended a kindergarten before they came to school.

Cluster 3: A lower-than-average level of cognitive development and an average/high level of non-cognitive development

Non-cognitive skills are developed quite well in this group: the children communicate easily, interact confidently, and behave in accordance with the rules and standards (0.32 and 0.45 in SD units, respectively). However, their reading and mathematical literacies are often lower than the sample mean (-0.29 and -0.42 in SD units).

The cluster is represented mostly by girls (64%). Vocabulary and phonological awareness appear to be developed statistically worse than in the previous two groups. The socioeconomic status is somewhat lower, too. Meanwhile, nearly all the children in this cluster had a preschool education experience.

Cluster 4: Low levels of cognitive and non-cognitive development

The group includes children whose levels of cognitive and non-cognitive skills are almost one SD unit lower than the sample mean: -0.9 and lower in all the four indicators. Nearly 90 percent of the students attended a kindergarten. Boys account for 62 percent of the cluster. Levels of vocabulary and phonological awareness are much lower in this cluster than in the other three. The socioeconomic status is also the lowest: only 25 percent of the families have more than 100 books at home, and over 50 percent of the mothers have no college degree.

Figures 1–4 display the relative position of the clusters across the four scales used for clustering.

4. Results and discussion

This study had a few interrelated purposes and relevant implications, both for research and practice.

First, it sought to demonstrate that direct independent assessment of children's skills at school entry is possible and teachers can use the iPIPS to build an accurate evidence-based picture of every child's cognitive and non-cognitive development. The core value of the instrument is evident in that it allows the teacher to assess the child's level of development: not only does it construct individual "profiles" of children but it also explains trajectories of child development, making predictions and inviting the teacher to help children in their cognitive and non-cognitive development. Importantly, the instrument

Figure 1. **Scores in mathematics**

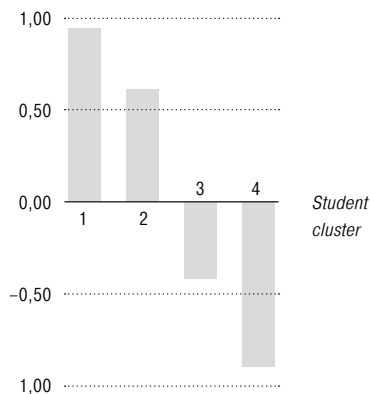


Figure 2. **Scores in reading**

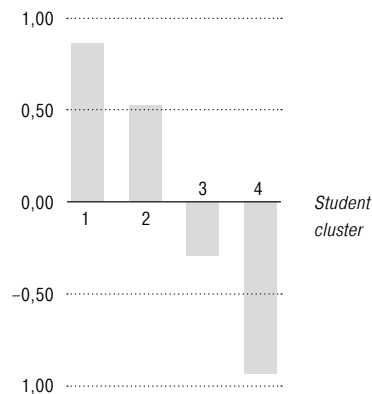


Figure 3. **Scores in confidence**

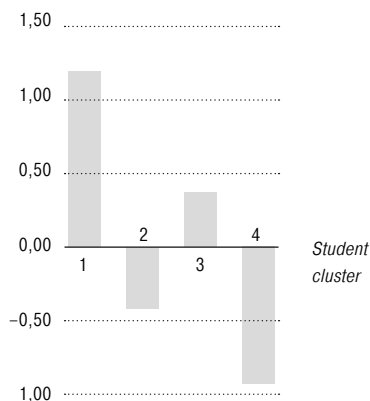
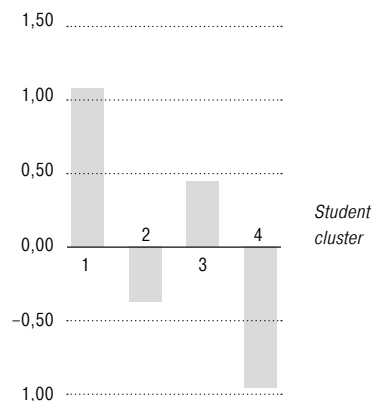


Figure 4. **Scores in classroom behavior**



allows for assessing not only the circumstances and components that are controllable by the teacher but also those beyond such control. For instance, the teacher cannot influence parental education in any way but is able to influence children's interactions with classmates and their individual cognitive needs. The teacher may change the conditions of child development, thus allowing for academic progress as well as adaptation of the child to the school and of the school, i. e. teaching strategies, to the child.

Second, the study forms the idea not only about children's development trajectories but also about the types of such trajectories, i. e. the groups (clusters) of children with similar characteristics. Such clustering, on the one hand, describes the development

of a first-grader cohort, and on the other hand it allows for identifying the common objectives in helping children with their cognitive and non-cognitive skills, thus facilitating the design of customized education programs. Cohort characteristics that will be obtained in replication studies during the coming years will make it possible to find out the differences among cohorts or generations of children enrolling to school.

The iPIPS provide teachers with multidimensional assessments of cognitive and non-cognitive development of children at school entry and at the end of the first year. These assessments can be used for evidence-based discussions on the real individual progress of every student as well as for the design of strategies to promote cognitive and non-cognitive development with allowances made for the new understanding of the child's situation represented as a personal history and trajectory of development.

The article used the findings of a large-scale survey of first-graders in four regions of the Russian Federation. Analysis revealed four groups (clusters) of first-graders that differ in the levels of cognitive and non-cognitive development and sociodemographic characteristics. The bottom scorers in both cognitive and non-cognitive skills are mostly represented by boys with non-college-educated mothers. Contrariwise, children with the highest levels of cognitive and non-cognitive development have well-educated parents, the largest home libraries, and the best scores in phonological awareness and vocabulary at the beginning of the school year. These findings confirm the results obtained by other researchers [Hindman et al. 2010].

However, the conducted study has a couple of limitations. The sample is key for generalizations. In this study, the sample consisted entirely of urban first-graders. Students from suburban or rural schools might well demonstrate different patterns of development; a separate study is required to find out the peculiar aspects of their cognitive and non-cognitive skills. Besides, although the sample in this study includes schools of different types from four cities, the sample unit is inconsistent, being either a class or a school in different cities. The total sample is thus non-representative, so the findings cannot be yet generalized for other regions or the country as a whole.

k-means clustering is used to identify the groups (clusters) of first-graders. One of the disadvantages of this method, and of cluster analysis as such, is its ultimately "statistical" nature, which means that clusters may be formed from the available quantitative data even if there is no real "theoretical" basis for classification. Stability of the resulting cluster solutions was verified using subsamples and findings from previous years in order to overcome this limitation.

Analysis is based on the results of a baseline assessment at the beginning of the school year, i. e. clusters describe children at school entry. This data is going to serve as the basis for preliminary recommendations on the best interaction strategies to be applied to children

from different groups. The recommendations for school teachers and principals are intended to be extended and improved in the future using the data on individual progress and the assistance strategies deployed by teachers.

References

- Alexander K. L., Entwisle D. R., Dauber S. L. (1993) First-Grade Classroom Behavior: Its Short- and Long-Term Consequences for School Performance. *Child Development*, vol. 64, no 3, pp. 801–814.
- Antipkina I., Kuznetsova M., Kardanova E. (2017) Chto sposobstvuet i chto meshaet progressu detey v chtenii [What Factors Help and Hinder Children's Progress in Reading?]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 2, pp. 206–233. DOI: 10.17323/1814-9545-2017-2-206-233
- Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Copping L. T., Cramman H., Gott S., Gray H., Tymms P. (2016) Name Writing Ability not Length of Name is Predictive of Future Academic Attainment. *Educational Research*, vol. 58, no 3, pp. 237–246.
- Domitrovich C. E., Durlak J. A., Staley K. C., Weissberg R. P. (2017) Social-Emotional Competence: An Essential Factor for Promoting Positive Adjustment and Reducing Risk in School Children. *Child Development*, vol. 88, no 2, pp. 408–416.
- Duckworth A. L., Yeager D. S. (2015) Measurement Matters: Assessing Personal Qualities Other than Cognitive Ability for Educational Purposes. *Educational Researcher*, vol. 44, no 4, pp. 237–251.
- Duncan G. J., Dowsett C. J., Claessens A., Magnuson K., Huston A. C., Klebanov P., Japel C. (2007) School Readiness and Later Achievement. *Developmental Psychology*, vol. 43, no 6, pp. 1428–1446.
- Durlak J. A., Weissberg R. P., Dymnicki A. B., Taylor R. D., Schellinger K. B. (2011) The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development*, vol. 82, no 1, pp. 405–432.
- Gagai V., Grineva K. (2013) Osobennosti shkolnoy adaptatsii pervoklassnikov kak trudnoy zhiznennoy situatsii dlya detey i roditel'ey [Features of the First-Graders' School Adaptation As a Difficult Vital Situation for Children and Parents]. *Herald of Chelyabinsk State Pedagogical University*, no 2, pp. 33–44.
- Gani V., Gani S. (2009) Razvitiye uchebnoy motivatsii pervoklassnikov pri razlichnykh stilyakh pedagogicheskogo rukovodstva [Developing Learning Motivation in First-Graders: The Effect of Different Teaching Styles]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 1, pp. 188–198. DOI: 10.17323/1814-9545-2009-1-188-198
- Grytsinskaya V., Gordiets A., Galaktionova M., Savchenko A., Manchouk V. (2003) Kharakteristika adaptatsionnykh vozmozhnostey pervoklassnikov [The Characteristic of the First-Formers Adaptive Abilities]. *Siberian Medical Journal*, no 38(3), pp. 75–78.
- Harms T., Vineberg Jacobs E., Romano White D. (2013) *School-Age Care Environment Rating Scale (SACERS)*. New York: Teachers College.
- Harms T., Clifford R. M., Cryer D. (2015) *Early Childhood Environment Rating Scale*. New York: Teachers College.
- Hawker D., Kardanova E. (2014) Startovaya diagnostika detey na vkhode v nachalnuyu shkolu i otsenka ikh progressa v techenie pervogo goda obucheniya: mezhdunarodnoe issledovanie iPIPS [First-Grade Baseline Evaluation and First-Grade Progress Monitoring: The International Performance Indicators in Primary Schools (iPIPS)]. Proceedings of the Conference «Trends

- in *Education Development. What Is an Effective School or Kindergarten?*» (February 19–20, 2014, Moscow), Moscow: Delo Publishing House, Russian Presidential Academy of National Economy and Public Administration, pp. 311–320.
- 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*, vol. 25, no 2, pp. 235–250.
- Ivanova A., Kardanova E., Merrell C., Tymss P., Hawker D. (2016) Checking the Possibility of Equating a Mathematics Assessment between Russia, Scotland and England for Children Starting School. *Assessment in Education: Principles, Policy and Practice* <https://doi.org/10.1080/0969594X.2016.1231110>
- Ivanova A., Kuznetsova M., Semenov S., Fedorova T. (2016) Faktory, opredelyayushchie gotovnost pervoklassnikov k shkole: vyvavlenie regionalnykh osobennostey [School Readiness of First-Graders and Its Factors: Identifying Region-Specific Characteristics]. *Voprosy obrazovaniya / Educational Studies Moscow*, no 4, pp. 84–105. DOI: 10.17323/1814-9545-2016-4-84-105
- Ivanova A., Nisskaya A. (2015) Startovaya diagnostika detey na vkhode v nachalnuyu shkolu: mezhdunarodnoe issledovanie iPIPS [Start Diagnosis of Children at the Beginning of Elementary School and the Mark of their Progress in the First Year of Studying]. *School Technologies*, no 2, pp. 161–168.
- Jain A. K. (2010) Data Clustering: 50 Years beyond K-Means. *Pattern Recognition Letters*, vol. 31, no 8, pp. 651–666.
- Janus M., Brinkman S., Duku E., Hertzman C., Santos R., Sayers M., Schroeder J., Walsh C. (2007) *The Early Development Instrument: A Population-Based Measure for Communities*. Hamilton, ON: Offord Centre for Child Studies, McMaster University.
- Jordan N. C., Kaplan D., Ramineni C., Locuniak M. N. (2009) Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes. *Developmental Psychology*, vol. 45, no 3, pp. 850–867.
- Kautz T., Heckman J. J., Diris R., Weel B. T., Borghans L. (2014) *Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success*. National Bureau of Economic Research Paper No 8696.
- Kovaleva G., Danilenko O., Ermakova I., Nurminskaya N., Gaponova N., Davydova Y. (2012) O pervoklassnikakh: po rezultatam issledovaniy gotovnosti pervoklassnikov k obucheniyu v shkole [On First-Graders: Based on Research Findings about First-Grader Readiness for School]. *Munitsipalnoe obrazovanie: innovatsii i eksperiment*, no 5, pp. 30–37.
- Kuzmina Y. V., Ivanova A. E., Antipkina I. V. (2017) *Direct and Indirect Effects of Phonological Ability and Vocabulary Knowledge on Math Performance in Elementary School*. Higher School of Economics Research Paper No WP BRP 76/PSY/2017. Moscow: HSE.
- Linacre J. M. (2011) *A User's Guide to WINSTEPS: Rasch Model Computer Programs*. Chicago: MESA.
- Manfra L., Squires C., Dinehart L. H. B., Bleiker C., Hartman S. C., Suzanne C., Winsler A. (2017) Preschool Writing and Premathematics Predict Grade 3 Achievement for Low-Income, Ethnically Diverse Children. *The Journal of Educational Research*, vol. 110, no 5, pp. 528–537. DOI: 10.1080/00220671.2016.1145095
- Margetts K. (2009) Early Transition and Adjustment and Children's Adjustment after Six Years of Schooling. *European Early Childhood Education Research Journal*, vol. 17, no 3, pp. 309–324.
- Merrell C., Tymms P. (2011) Changes in Children's Cognitive Development at the Start of School in England 2001–2008. *Oxford Review of Education*, vol. 37, no 3, pp. 333–345.

- Merrell C., Tymms P. (2016) Assessing Young Children: Problems and Solutions. *Understanding What Works in Oral Reading Assessments: Recommendations from Donors, Implementers and Practitioners*. Montreal, Canada: UNESCO Institute for Statistics, pp. 126–133.
- Müller K., Brady S. (2001) Correlates of Early Reading Performance in a Transparent Orthography. *Reading and Writing*, vol. 14, no 7–8, pp. 757–799.
- OECD (2015) *Skills for Social Progress: The Power of Social and Emotional Skills*. Paris: OECD.
- Orel E., Brun I. V., Kardanova E., Ivanova A. (2016) *Noncognitive Development of First Graders and Their Cognitive Performance*. Higher School of Economics Paper No WP BRP 57/PSY/2016. Moscow: HSE.
- Orel E., Ponomareva A. (2016) *Children with Behavioral Problems in the First Grade of Russian School: Similarities and Differences*. Higher School of Economics Paper No WP BRP 66/PSY/2016. Moscow: HSE.
- Orel E., Ponomareva A. (2018) Patterny sotsialno-emotsionalnogo razvitiya pervoklassnika na vkhode v shkolu [Patterns of Socio-Emotional Development of First-Graders as They Come to School]. *Psychology. Journal of the Higher School of Economics*, no 1 (in print).
- Pagani L. S., Fitzpatrick C., Archambault I., Janosz M. (2010) School Readiness and Later Achievement: A French Canadian Replication and Extension. *Developmental Psychology*, vol. 46, no 5, pp. 984–994.
- Paranicheva T., Tyurina E. (2012) Funktsionalnaya gotovnost k shkole detey 6–7 let [Functional School Readiness in 6–7-Year-Old Children]. *Novye issledovaniya*, no 1(30), pp. 135–145.
- Poropat A. E. (2009) A Meta-Analysis of the Five-Factor Model of Personality and Academic Performance. *Psychological Bulletin*, vol. 135, no 2, pp. 322–338.
- Salnikova S., Tkachenko M. (2012) Uroven sformirovannosti universalnykh uchebnykh deystviy u pervoklassnikov: startovaya diagnostika [The Level of Formedness of Universal Learning Activities among First-Graders: Beginning of the Year Assessment]. *Ekspерiment i innovatsii v shkole*, no 2, pp. 3–10.
- Slentz K. L. (2008) *A Guide to Assessment in Early Childhood: Infancy to Age Eight*. Olympia, WA: Washington State Office of Superintendent of Public Instruction.
- Spence S. H. (1987) The Relationship between Social—Cognitive Skills and Peer Sociometric Status. *British Journal of Developmental Psychology*, vol. 5, no 4, pp. 347–356.
- Tsukerman G., Polivanova K. (2012) *Vvedenie v shkolnuyu zhizn: programma adaptatsii detey k shkolnoy zhizni: posobie dlya uchiteley* [Introduction to School Life: School Orientation Program for Children: Teacher's Guide], Moscow: Vita-Press.
- Tymms P. (1999) Baseline Assessment, Value-Added and the Prediction of Reading. *Journal of Research in Reading*, vol. 22, no 1, pp. 27–36.
- Tymms P., Merrell C., Buckley H. (2015) *Children's Development at the Start of School in Scotland and the Progress Made During their First School Year: An Analysis of PIPS Baseline and Follow-Up Assessment Data*. Edinburgh, UK: The Scottish Government.
- Tymms P., Merrell C., Henderson B. (1997) The First Year at School: A Quantitative Investigation of the Attainment and Progress of Pupils. *Educational Research and Evaluation*, vol. 3, no 2, pp. 101–118.
- Tymms P., Merrell C., Henderson B. (2000) Baseline Assessment and Progress during the First Three Years at School. *Educational Research and Evaluation: An International Journal on Theory and Practice*, vol. 6, no 2, pp. 105–129.
- Tymms P., Merrell C., Henderson B., Albone S., Jones P. (2012) Learning Difficulties in the Primary School Years: Predictability from On-Entry Baseline

- Assessment. *Online Educational Research Journal*. <http://community.dur.ac.uk/p.b.tymms/oerj/>
- Tymms P., Merrell C., Wildy H. (2015) The Progress of Pupils in their First School Year across Classes and Educational Systems. *British Educational Research Journal*, vol. 41, no 3, pp. 365–380.
- Wright B. D., Masters G. N. (1982) *Rating Scale Analysis: Rasch Measurement*. Chicago, IL: MESA.
- Wright B. D., Stone M. H. (1979) *Best Test Design. Rasch Measurement*. Chicago, IL: MESA.