Psychology in Russia: State of the Art
Volume 11, Issue 4, 2018

Russian Psychological Society
Lomonosov Moscow State University

Cognitive and Non-Cognitive Predictors of the Unified State Exam Performance of Students from Schools with Regular and Advanced Mathematical Curricula

Ivan A. Voronin[a], Olga N. Ovcharova[a],
Elizaveta M. Bezrukova[b], & Yulia Kovas[b,c,*]

[a] *Psychological Institute of Russian Academy of Education, Moscow, Russia*

[b] *International Centre for Research in Human Development, Tomsk State University, Tomsk, Russia*

[c] *Goldsmiths, University of London, UK*

* Corresponding author. E-mail: y.kovas@gold.ac.uk

Background. Exams such as the SAT, ACT, and GCSE are used to give an account of educational outcomes and provide a unified criterion for university admission. The Unified State Exam (USE) aims to fulfill these functions in Russia.All Russian students take two compulsory USE exams, mathematics and Russian, at the end of their school education.

Objective. Variability in the mathematics and Russian USE scores is vast, both across and within schools. Our study investigated potential sources of this variability.

Design. The sample included 196 students from regular schools (non-selected students) and 306 students from schools with advanced mathematical curriculum (selected students). The mathematical ability (numerical representation, mathematical fluency), intelligence, basic cognitive functions (working memory, reaction time), and mathematical self-efficacy of the students were assessed. We applied structural equation modeling to estimate the proportion of variability in the mathematics and Russian USE scores explained by cognitive predictors and mathematical self-efficacy.

Results. In the whole sample, cognitive predictors and mathematical self-efficacy explained 54% of the variation in the mathematics USE scores and 30% of the variation in the Russian USE scores. These effects diminished after the data were analyzed in two groups separately, suggesting that the associations between predictors and exam scores were to a large extent accounted for by group differences (students from regular and specialized schools).

Conclusion. The students from the schools with an advanced mathematical curriculum exhibited better cognitive performance, appraised their mathematical abilities higher, and achieved higher mathematics and Russian USE scores, compared to the students from regular schools. Within the groups, cognitive and non-cognitive predictors explained a small part of the variation of the mathematics and Russian USE scores.

Keywords: academic achievement, Unified State Exam, mathematical ability, intelligence, number sense.

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2018.0412
http://psychologyinrussia.com

Introduction

Standardized examination is used all over the world to evaluate the level of students' mastery of their school curriculum and to provide a unified criterion for college admission. The ultimate goal of standardized examination is to provide equal access to higher education independent of students' sociodemographic characteristics. Introduced in 2006, the Unified State Exam (USE) now serves as a major educational assessment tool at the end of school. Every year all school students in Russia take mathematics and Russian USEs. Students' exam scores vary immensely (Federal Service for Supervision in Education and Science; Federal Institute of Pedagogical Measurements, 2012). Our study aims to explore the sources of individual differences in the USE performance of students from regular and specialized schools.

One of the strongest predictors of academic performance is intelligence, which represents common variability across the manifold of cognitive abilities and characteristics (Deary, Strand, Smith, & Fernandes, 2007; O'Connell, 2018). Studies of Western standardized exams — the General Certificate of Secondary Education (GCSE, UK), the Scholastic Aptitude Test (SAT, USA), and the American College Testing (ACT) — attribute a large part of the variation of exam scores to general cognitive ability (g). Intelligence explains up to 30% of the variability of GCSE scores (Krapohl et al., 2014; Rimfeld, Dale, & Plomin, 2015) and up to 60% of the variation of SAT and ACT scores (Coyle, 2015; Koenig, Frey, & Detterman, 2008). Generally, the association is tighter for the mathematical exam than the language exam (Deary et al., 2007; Rimfeld et al., 2015).

Although universal cognitive characteristics explain a substantial part of the variation of exam scores, subject-specific abilities are also important. Specific cognitive abilities (verbal and mathematical) are associated with the specific variability (non-g residuals) of SAT and ACT scores (Coyle, Purcell, Snyder, & Kochunov, 2013). SAT and ACT mathematical subtests are positively associated with mathematical ability and negatively with verbal ability. For the verbal subtests, the pattern is the opposite (Coyle, Snyder, Richmond, & Little, 2015; Schult & Sparfeldt, 2016).

Learning complex mathematical concepts is partly associated with a mental representation of quantity. Non-symbolic numerical representation (number sense or an approximate number system), manifesting itself as the ability to intuitively compare quantities, can be detected early in life (Dehaene, 2011; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Tosto et al., 2014). With the start of formal school education, symbolic numerical representation develops, enabling a more precise comparison of quantities and abstract mathematical concepts (Merkley & Ansari, 2016; Siegler & Lortie-Forgues, 2014).

Individual differences in numerical representation systems are among the predictors of mathematical achievement (De Smedt, Noël, Gilmore, & Ansari, 2013; Feigenson, Libertus, & Halberda, 2013). In primary school, number sense can account for as much as 54% of the variation of performance on general curriculum-based mathematical tests (Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013). The relationship between an approximate number system and mathematical performance in preschool years has been shown to be causal (Wang, Odic, Halberda, & Feigenson, 2016). However, by the end of school education, its contribution drops to 10% (Libertus,Odic, & Halberda, 2012; Matthews, Lewis, & Hubbard,

2016). Symbolic numeric representation, on the contrary, is a consistent predictor of mathematical achievement across school grades (Fazio, Bailey, Thompson, & Siegler, 2014). Non-symbolic and symbolic representation systems substantially overlap, but they also contribute some unique variation to the measures of mathematical performance (Lourenco & Bonny, 2017).

Beyond cognitive characteristics and abilities, non-cognitive factors can account for as much as 20% of the variation of academic achievement (Krapohl et al., 2014; Parker, Marsh, Ciarrochi, Marshall, & Abduljabbar, 2014; Seaton, Parker, Marsh, Craven, & Yeung, 2014). A number of studies emphasize self-concept (beliefs about one's own ability) and self-efficacy (beliefs about one's own performance) as important non-cognitive predictors of academic achievement (Bong & Skaalvik, 2003; Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007). Academic self-concept and self-efficacy are domain-specific and reciprocally associated with academic performance in the corresponding domain (Chen, Yeh, Hwang, & Lin, 2013; Malanchini et al., 2017; Marsh & Martin, 2011). Most research shows a modest effect of mathematical self-concept on subsequent mathematical achievement, from about 5% (Möller, Zimmermann, & Köller, 2014) to 20% (Suárez-Álvarez, Fernández-Alonso, & Muñiz, 2014). Reciprocal links between self-concept and achievement over time may also contribute to choice of academic specialization (Parker et al., 2014). Positive academic motivation and self-concept are associated with better academic performance through positive attitudes toward school and engaging in learning experiences (Green et al., 2012; Tosto, Asbury, Mazzocco, Petrill, & Kovas, 2016). On the other hand, lower self-esteem is associated with maladaptive behavior in approaching the task (self-handicapping), which has a detrimental effect on academic performance (Gadbois & Sturgeon, 2011; Schwinger, Wirthwein, Lemmer, & Steinmayr, 2014).

The majority of studies addressing the variability of USE scores have been concerned with the question of the criterion validity of the exams, i.e., the association between exam performance and subsequent academic achievement at university. The USE scores indeed predict individual differences in academic achievement and dropping out of university (Pereyaslavskaya & Pereyaslavskiy, 2014; Poldin, 2011; Zamkov & Pereseckij, 2013). About 20% of the variability of academic achievement at university can be accounted by USE performance (Khavenson & Solovyova, 2014). Therefore, the USE appears to fulfill its selective function to some extent. However, it is not clear whether it can be considered a reliable measure of acquisition of the school curriculum, as the correlations between exam scores and school performance have been found to be highly variable across school subjects (Chernyavskaya & Merkulov, 2015; Saprykina, 2017).

Several studies address the issue of predictors of performance on the USE. The correlation between general cognitive ability and USE performance varies across USE subjects and between years of testing (Kaptsov & Kolesnikova, 2009). The highest correlation between measures of intelligence and USE performance is 0.33 (Matsuta, Bogomaz, & Sudneva, 2014). Intelligence and specific cognitive abilities (verbal, mathematical, and spatial abilities, and abstract reasoning) explain up to 25% of individual differences in exam performance (Kaptsov & Kolesnikova, 2009). USE scores have also been found to be associated with: motivation (14%; Gordeeva & Osin, 2012), emotional intelligence (10%; Dmitrieva & Gelman, 2015), and

personality (agreeableness and neuroticism) (up to 10%; Kochergina, Nye, & Orel, 2013). Family income and investment in additional training have also been found to partially explain USE performance (up to 20%; Prakhov, 2014; Prakhov & Yudkevich, 2012).

Our study aims to add to the data on sources of the variation in the scores of two compulsory exams: the mathematics and Russian USEs. We compare students from regular state schools and from schools with a highly advanced mathematical curriculum to assess whether their mathematics and Russian USE scores are explained by the same set of predictors.

Methods

Objectives

The study aimed to explore the sources of individual differences in USE performance in students from regular and specialized schools.

Sample

We compared two groups of school students. Group 1 (selected students) included 306 students from two schools with a highly enhanced mathematical curriculum. These schools were residential institutions that admit students from all over Russia. To enter these schools at Grade 10, the students passed strict selection criteria based on their previous mathematical achievement, including success in mathematical competitions at regional, federal, and international levels. Group 1 included 102 female and 198 male students (six students did not report their sex), with a mean age of 17.1 years (SD = 0.44 years).

Group 2 (non-selected students) included 196 from regular state schools. These schools do not select their students and use a standard mathematical curriculum. Group 2 included 120 female and 76 male students, with a mean age of 16.8 years (SD = 0.63 years).

Measures

Participants completed a cognitive test battery at the last grade of school (Grade 11). The test battery consisted of measures of non-symbolic numerical representation (Number Sense test, NS), symbolic numerical representation (Number Line, NL; Dot Number, DN), mathematical fluency (Understanding Numbers, UN; Problem Verification, PV), spatial working memory (Corsi Block, CB), and simple reaction time (Reaction Time, RT). The details of the test battery measures are provided in Table 1.

General cognitive ability was assessed using Raven's Standard Progressive Matrices (Raven, 2003). Group 1 performed the standard version of the test, comprising five series of 12 tasks of progressive difficulty. Group 2 took a modified version of the test that included 18 items from the original test and 12 advanced tasks. In both groups, general cognitive ability was represented by the number of correct responses.

The students' perception of their mathematical performance (mathematical self-efficacy, math SE) was assessed using eight questions about their level of con-

Table 1

Summary of the cognitive test battery

	Characteristic	Task	# of items	Variables	Reference
CB – Corsi Block	Spatial working memory	Recall a sequence of boxes	12	Number of correct responses (CB score)	(Pagulayan, Busch, Medina, Bartok, & Krikorian, 2006)
NL – Number Line	Symbolic numerical representation	Locate a number on a 0–1,000 scale	22	Mean deviation from the correct location (NL score)	(Opfer & Siegler, 2007)
RT – Reaction Time	Simple reaction time	Respond to a digit (1, 2, 3, 4)	40	Mean latency of correct response (RT latency)	(Deary, Der, & Ford, 2001)
NS – Number Sense	Non-symbolic numerical representation	Compare the number of mixed yellow and blue dots without counting	150	Number of correct responses (NS score) Mean latency of correct response (NS latency)	(Halberda et al., 2008)
UN – Understanding Number	Mathematical fluency, complex task	Solve a problem from the standard mathematical curriculum	18	Number of correct responses (UN score)	(Kovas, Haworth, Petrill, & Plomin, 2007)
DN – Dot Number	Symbolic numerical representation	Report whether a digit matches the number of dots	32	Number of correct responses (DN score) Mean latency of correct response (DN latency)	–
PV – Problem Verification	Mathematical fluency, simple task	Check a simple arithmetic expression	48	Number of correct responses (PV score) Mean latency of correct response (PV latency)	(Murphy & Mazzocco, 2008)

fidence in solving mathematical tasks (PISA, 2012). Cronbach's alpha of the self-efficacy scale was 0.65 in Group 1 and 0.61 in Group 2.

At the end of the same academic year, students provided their grades for the Unified State Exam (USE) in mathematics and Russian, graded on a scale of 0 to 100. These data were available for 207 students for the math USE and 273 for the Russian USE.

Data preparation

We applied logarithmic transformation to the latency measures (RT, NS, DN, and PV latencies) to eliminate a positive skew of their distributions (Ratcliff, 1993). Then, for each variable we considered the observations deviating more than 1.5 interquartile range from the first and third quartile as outliers and excluded them (Rousseeuw & Hubert, 2011). This procedure was performed within each group and for each variable separately. Upon the data preparation, all the variables exhibited a normal distribution.

Results

Descriptive statistics

Table 2 presents means and standard deviations for cognitive characteristics, intelligence, mathematical self-efficacy, and exam performance for the whole sample and within the groups of selected and non-selected students. We performed a t-test to compare performance across the groups. For latency measures, we present means and standard deviations for both the transformed and original data.

The group of selected students outperformed the non-selected students' group in all cognitive tasks, giving more correct answers in math tasks (UN score, PV score), giving more precise numerical estimations (NS, NL, DN), and achieving a higher spatial working memory score (CB). Also, the selected students' group responded faster on the simple reaction time task (RT) and slower on the tests of numeric estimation (NS) and mathematical problem solving (PV).

The selected students' group achieved higher exam scores for both mathematics and Russian. The difference was not only statistically significant but also large (Cohen's $d = 3.3$ for mathematics and 1.5 for Russian). Compared to the 2011–2012 nationwide USE scores (math: 45–47, Russian: 60–61), Group 2 students performed average and Group 1 performed far above average. Students in Group 1 also evaluated their math performance higher than Group 2 students did (math self-efficacy).

Age and sex differences

The groups of selected and non-selected students were unbalanced in terms of their male/female ratios: there were more male students in the selected group, and more female students in the non-selected groups. We performed a 2×2 ANOVA to test whether there were any sex differences in cognitive characteristics and academic achievement besides the group affiliation. No statistically significant effect of sex was found after adjusting for multiple comparisons, the median effect was 1.7%. The highest effect of 9.3% was seen for the math USE: the average score for males was 2.8 points higher than for females. The sex-by-group interaction was also non-

Table 2
Descriptive statistics

	Total			Group 1 Selected			Group 2 Non–selected			t–test			
	n	**M**	**SD**	**n**	**M**	**SD**	**n**	**M**	**SD**	**t**	**df**	**p**	**Cohen's d**
CB score	446	5.91	2.20	288	6.44	2.07	158	4.92	2.08	7.390	321.520	0.000	0.735
NL score	471	35.10	12.97	289	30.34	8.70	182	42.67	14.92	–10.124	259.405	0.000	–1.074
RT latency [a]	466	–0.44 (0.66)	0.15 (0.14)	291	–0.46 (0.65)	0.14 (0.13)	175	–0.41 (0.69)	0.16 (0.16)	–3.540	306.453	0.000	–0.360
NS score	462	111.24	13.63	279	116.47	8.24	183	103.26	16.15	10.223	244.683	0.000	1.102
NS latency [a]	472	–0.09 (0.90)	0.23 (0.26)	289	–0.03 (0.97)	0.18 (0.21)	183	–0.19 (0.80)	0.27 (0.28)	6.470	247.771	0.000	0.707
UN score	409	14.72	3.10	281	16.30	1.58	128	11.26	2.79	19.109	165.405	0.000	2.482
DN score	453	13.26	6.66	276	14.79	5.47	177	10.88	7.60	5.926	291.605	0.000	0.613
DN latency [a]	468	0.08 (1.08)	0.15 (0.20)	286	0.08 (1.08)	0.15 (0.18)	182	0.09 (1.06)	0.15 (0.22)	–0.598	377.111	0.550	–0.057
PV score	447	41.71	6.06	275	44.70	2.11	172	36.92	7.16	13.868	189.703	0.000	–1.645
PV latency [a]	477	1.03 (2.75)	0.25 (0.88)	293	1.01 (2.80)	0.23 (0.65)	184	1.07 (2.65)	0.27 (1.14)	–2.413	285.568	0.016	–0.249
Raven's	–	–	–	292	55.51 [b]	3.98	192	11.32 [b]	3.48	–	–	–	–
Math SE	460	14.48	2.77	280	15.21	2.25	180	13.36	3.11	6.906	298.072	0.000	0.708
Math USE	207	65.80	20.90	116	81.52	10.98	91	45.76	11.00	23.227	193.243	0.000	3.269
Russian USE	273	76.10	12.70	184	81.13	9.50	89	65.70	12.16	10.518	141.712	0.000	1.483

Note: [a] The untransformed descriptive statistics are presented in parentheses. M = mean, SD = standard deviation.
[b] We did not compare the mean Raven's scores because different versions of the test with different ranges of scores were applied in two groups.

significant, with a maximum effect size of 1.8% (RT correct latency). These small (non-significant) effects of sex and the imbalance of the male/female ratios in the two groups are unlikely to confound the results of further analyses.

We also examined age differences both within and across groups. Overall, the difference in mean age across the selected and non-selected students was small, and the variability of age within the groups was low. Across the whole sample, statistically significant age-related differences were detected in NL, UN, PV, and math SE and USE scores, the correlations being between 0.1 and 0.2. Older participants performed better in these cognitive tasks, yielded higher exam marks, and rated their mathematical skills higher. Within the groups, no statistically significant age effects were found.

Relationships across cognitive characteristics

We studied the relationships between cognitive characteristics for the whole sample (Table 3) and within the groups (Table 4). For the whole sample, the size of the relationships varied between 0.0 and 0.6. NS scores correlated moderately or highly with most cognitive measures in the study, supporting the notion that number sense appears to be an important basic characteristic underlying individual differences in mathematical cognition.

Table 3

Correlations between cognitive abilities for the whole sample

		1	2	3	4	5	6	7	8	9
1	CB score									
2	NL score	–0.233[c]								
3	UN score	–0.186[c]	0.063							
4	RT latency	0.370[c]	–0.306[c]	–0.241[c]						
5	NS score	0.228[c]	–0.242[c]	0.008	0.657[c]					
6	NS latency	0.384[c]	–0.453[c]	–0.112[a]	0.473[c]	0.339[c]				
7	DN score	0.282[c]	–0.304[c]	–0.198[c]	0.421[c]	0.207[c]	0.252[c]			
8	DN latency	0.027	–0.068	0.140[b]	0.190[c]	0.339[c]	0.012	0.099[a]		
9	PV score	0.343[c]	–0.426[c]	–0.223[c]	0.561[c]	0.381[c]	0.609[c]	0.410[c]	0.186[c]	
10	PV latency	0.037	–0.039	0.096[a]	0.345[c]	0.432[c]	0.026	0.040	0.337[c]	0.356[c]

Note:[a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$

To find out whether there is any factor structure underlying relationships across cognitive characteristics, we performed a maximum likelihood factor analysis. The factors were allowed to correlate, so some cognitive measures were loaded by more than one factor. Three factors were extracted in the whole sample and within the groups, explaining 43% of the total variability. The strongest factor (19%) was associated with acquired mathematical knowledge (PV score, UN score), but also numeric representation (NL score). The other two factors represented number

Table 4

Correlations between cognitive abilities in the two groups

			Group 1									
			1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Group 2	**1**	CB score		−0.113	−0.114	0.150[a]	−0.002	0.290[c]	0.192[b]	−0.077	0.186[b]	−0.146[a]
	2	NL score	−0.112		−0.056	−0.087	−0.159[b]	−0.192[b]	−0.150[a]	−0.054	−0.160[b]	0.068
	3	UN score	−0.205[b]	0.058		−0.215[c]	0.242[c]	−0.027	−0.171[b]	0.345[c]	−0.171[b]	0.197[c]
	4	RT latency	0.372[c]	−0.137	−0.183[a]		0.178[b]	0.200[b]	0.376[c]	−0.075	0.217[c]	−0.062
	5	NS score	0.347[c]	−0.062	−0.099	0.809[c]		0.144[a]	0.029	0.290[c]	0.023	0.247[c]
	6	NS latency	0.272[b]	−0.243[b]	−0.109	0.373[c]	0.364[c]		0.087	0.047	0.135[a]	−0.080
	7	DN score	0.262[b]	−0.268[c]	−0.144	0.317[c]	0.209[b]	0.179		−0.216[c]	0.146[a]	−0.194[b]
	8	DN latency	0.161	−0.097	−0.044	0.346[c]	0.387[c]	0.085	0.372[c]		−0.040	0.245[c]
	9	PV score	0.235[b]	−0.228[b]	−0.191[a]	0.463[c]	0.343[c]	0.342[c]	0.396[c]	0.323[c]		−0.090
	10	PV latency	0.194[a]	−0.086	0.053	0.534[c]	0.561[c]	0.215[a]	0.172[a]	0.404[c]	0.616[c]	

Note:[a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$

sense (15%, NS score, NS latency) and response speed (9%, RT latency, DN latency, PV latency). The factors of mathematical knowledge and number sense correlated highly with each other (0.63), but not with the latency factor.

In Group 1, the three-factor structure was less pronounced, as the factors explained 29% of total variability. The first factor comprised all latency measures (11%), the second one represented number sense (10%, NS score, NS latency, also DN score, PV score, and RT latency). The third factor was loaded by working memory and acquired mathematical knowledge (8%, CB score, UN score, PV score, and NL score). The correlations across factors were under 0.22.

In Group 2, the three-factor solution explained as much variability as in the whole sample, 44%. The first factor comprised the measures of number sense (17%, NS latency, NS score), CB score, and UN score. The second factor represented general performance on the tests of mathematical abilities (16%, DN score, PV score, NS score, CB score, NL score, UN score, and RT latency). The third factor comprised the latency measures (11%, PV latency, DN latency, RT latency) and also PV score. The cross-factor correlations were between 0.2 and 0.3.

To summarize, in the whole sample and within the groups of selected and non-selected students, we discovered three main factors regarding the performance on tasks of mathematical abilities: (1) acquired mathematical knowledge, (2) number sense, and (3) response speed. The factor of acquired mathematical knowledge had the most discriminative power in the whole sample and mirrored the differences across the selected and non-selected groups, as it was an explicit criterion of selection for the advanced mathematical schools. In the group of selected students, the factor of acquired mathematical knowledge did not have much discriminative power. In contrast, in the group of non-selected students, acquired mathematical knowledge appeared as part of the broad factor of mathematical performance that comprised all cognitive scores.

Number sense was associated with mathematical performance, especially in non-selected students. In both the selected and non-selected groups, the slower responses on the NS test were associated with more correct responses on this test. Response speed explained the same amount of variation in the selected and non-selected students.

Math self-efficacy, intelligence, and cognitive characteristics

We considered the relationships between cognitive characteristics, intelligence, and mathematical self-efficacy separately. As the students from the two groups completed two different versions of Raven's Progressive Matrices, we did not compute associations between intelligence and cognitive characteristics for the whole sample. Along with the separate cognitive measures, at this stage we included the factor values extracted from the factor analysis. The factor values were computed separately for Group 1 and Group 2. Within the groups, the associations between Raven's scores and cognitive measures were weak and statistically non-significant. Two statistically significant correlations were found in Group 1 (UN score, PV score) and two in Group 2 (PV latency, latency factor). The lack of association between intelligence and cognitive measures in the group of selected students can be accounted for by the restricted range of the Raven's scores. We observed a clear

ceiling effect in the group of selected students, with the average performance on the Raven's test of 55 points out of 60. The whole variation of Raven's scores in Group 1 fits the range between 44 and 60 points.

The associations between math SE and cognitive characteristics for the whole sample ranged between 0.1 and 0.3 (statistically significant for all score indices and PV latency). Within the groups, math SE correlated with NS latency in Group 1 and with NS score in Group 2, with effects under 0.16. Most likely, the association between math SE and cognitive characteristics resulted from group differences: selected students gave a higher appraisal of their mathematical ability and performed better on cognitive tests.

Cognitive abilities and mathematical self-efficacy as predictors of exam performance

The correlations between exam performance and cognitive abilities, and between intelligence and math SE are reported in Table 5. For the whole sample, the exam performance (both mathematics and Russian) were most closely associated with measures of mathematical fluency (UN and PV scores) and to the lesser extent with numerical representation (NS and NL score), working memory (CB score and CB latency), and response speed (RT, NS, and PV latency), as well as math SE. Within the groups, these associations were mostly not significant. Among cognitive measures, the highest association was between PV latency and the Russian USE in Group 2 ($r = 0.391$, $p = 0.000$).

Table 5

Correlations between exam performance (USE) and cognitive abilities, intelligence, and mathematical self-efficacy

	Mathematics USE			**Russian USE**		
	All	**Group 1**	**Group 2**	**All**	**Group 1**	**Group 2**
CB score	0.279 [c]	0.077	–0.030	0.135 [a]	0.001	–0.134
NL score	–0.368 [c]	0.061	–0.083	–0.255 [c]	0.004	–0.036
UN score	–0.172 [a]	–0.228 [a]	–0.109	0.040	0.104	0.020
RT latency	0.377 [c]	0.031	0.086	0.199 [b]	–0.045	0.006
NS score	0.323 [c]	–0.051	0.142	0.241 [c]	–0.030	0.163
NS latency	0.623 [c]	–0.001	0.149	0.398 [c]	0.048	–0.104
DN score	0.303 [c]	0.131	0.053	0.117	–0.019	–0.047
DN latency	–0.044	0.022	0.011	0.079	0.071	0.124
PV score	0.517 [c]	–0.056	0.046	0.365 [c]	–0.048	0.072
PV latency	–0.122	–0.077	0.124	0.192 [b]	0.107	0.391 [c]
Raven's	–	0.095	0.001	–	–0.026	0.158
Math SE	0.274 [c]	–0.035	0.132	0.165 [b]	–0.001	–0.032

Note: [a] $p < 0.05$, [b] $p < 0.01$, [c] $p < 0.001$

To estimate the proportion of variability of the USE scores accounted for by the cognitive predictors and math self-efficacy, we applied structural equation modeling (Loehlin, 2004). The model included either mathematics or Russian USE scores regressed using cognitive measures (CB score, NS score, UN score, RT latency, NS score and latency, DN score and latency, PV score and latency) and math SE as independent variables. The predictors were allowed to covariate. Five modifications of this model were used to address the research questions:

(1) The homogeneity model did not differentiate selected and non-selected students, assuming that all students had come from the same population.
(2) The heterogeneity model assumed that two groups of school students had come from two different populations. The associations across variables were allowed to differ across the groups.
(3) Model A — **a**symmetric relationships constrained: the model assumed that cognitive variables in two groups were associated with exam performance in the same way. The covariation across predictors was allowed to differ between the groups.
(4) Model S — **s**ymmetric relationships constrained: the model assumed that the covariation across predictors was equal in the two groups. The regression coefficients were allowed to differ between the groups.
(5) Model AS — **a**symmetric and **s**ymmetric relationships constrained: the model assumed that both regression and covariation relationships were equal in the two groups.

The models were compared using a chi-squared test (Table 6). A statistically significant difference in model fit implied that the model with fewer parameters did not explain the data well and could not be accepted. In contrast, when the fit indexes of two models were close, the model with fewer the parameters was preferred. Models A, S and AS were compared against the heterogeneity model. Models S and AS explained data less well, indicating that the structure of covariation across cognitive predictors differed between the two groups. Model A yielded the same fit as the heterogeneity model, meaning that no difference in regression paths was detected between the groups.

For the whole sample, cognitive characteristics explained 52% of the variability of the mathematics USE scores and 30% of the variability of the Russian USE scores. Significant predictors were measures of mathematical fluency (UN and PV scores) (Table 7). Other predictors had smaller effects and did not reach statistical significance, except for latency in the PV task as a predictor of Russian USE scores.

Within the groups, all predictors explained 10% and 16% of the mathematics USE scores and 2% and 35% of the Russian USE scores. In Group 1, the only statistically significant predictor of performance on the mathematics USE was RT latency, while no predictors were found to explain a sufficient amount of variation in the Russian USE scores. In Group 2, PV latency was the only notable predictor of Russian USE scores (29% of the variation).

Table 6

Model fit statistics

	Model	Comparison	ep	–2LL	df	AIC	BIC	RMSEA	ΔLL	Δdf	p
					Mathematics USE						
1	Heterogeneity	Saturated	180	11980.851	4971	2038.851	–18931.810	0.000	0.000	0	1.000
2	Model A	Heterogeneity	169	11991.612	4982	2027.612	–18989.454	0.000	10.760	11	0.464
3	Model S	Heterogeneity	114	12641.225	5037	2567.225	–18681.863	0.134	660.374	66	0.000
4	Model AS	Heterogeneity	102	12651.926	5049	2553.926	–18745.786	0.123	671.075	78	0.000
5	Homogeneity	Saturated	90	13283.215	5061	3161.215	–18189.120	0.000	0.000	0	1.000
					Russian USE						
1	Heterogeneity	Saturated	180	12318.670	5037	2244.670	–19004.419	0.000	0.000	0	1.000
2	Model A	Heterogeneity	169	12332.325	5048	2236.325	–19059.168	0.022	13.655	11	0.253
3	Model S	Heterogeneity	114	12978.323	5103	2772.323	–18755.193	0.134	659.653	66	0.000
4	Model AS	Heterogeneity	102	12998.001	5115	2768.001	–18810.138	0.124	679.331	78	0.000
5	Homogeneity	Saturated	90	13523.200	5127	3269.200	–18359.563	0.000	0.000	0	1.000

*Note:. ep = number of estimated parameters, -2LL = -2 * log-likelihood, df = degrees of freedom, AIC = Akaike information criterion, BIC = Bayesian information criterion, RMSEA = root mean square error of approximation, ΔLL = difference in log-likelihood, Δdf = difference in degrees of freedom*

Model A = asymmetric (regression) relationships constrained, Model S = symmetric (covariate) relationships constrained, Model AS = both asymmetric and symmetric relationships constrained

Table 7

Estimates of regression paths

	Mathematics USE									Russian USE								
	Homogeneity			Group 1			Group 2			Homogeneity			Group 1			Group 2		
	estEst	lCI	uCI	Est	lCI	uCI	Est	lCI	uCI	Est	lCI	uCI	Est	lCI	uCI	Est	lCI	uCI
CB score	–0.04	–0.16	0.08	0.05	–0.15	0.25	–0.17	–0.41	0.10	–0.04	–0.16	0.08	–0.01	–0.17	0.16	–0.17	–0.38	0.07
NL score	0.01	–0.11	0.14	0.06	–0.14	0.25	–0.02	–0.25	0.22	0.00	–0.13	0.14	0.01	–0.15	0.17	–0.01	–0.23	0.19
UN score	**0.50**	**0.34**	**0.65**	0.02	–0.16	0.20	0.17	–0.13	0.44	**0.30**	**0.12**	**0.48**	0.05	–0.10	0.21	–0.11	–0.35	0.14
RT latency	–0.07	–0.18	0.04	**–0.26**	**–0.44**	**–0.06**	–0.17	–0.38	0.06	0.03	–0.09	0.14	0.05	–0.11	0.22	–0.07	–0.27	0.14
NS score	–0.02	–0.20	0.15	–0.09	–0.31	0.15	–0.09	–0.50	0.31	–0.03	–0.22	0.15	–0.03	–0.21	0.15	–0.15	–0.55	0.23
NS latency	0.08	–0.06	0.23	–0.02	–0.21	0.17	0.19	–0.19	0.55	0.06	–0.08	0.21	–0.05	–0.20	0.10	0.27	–0.10	0.61
DN score	0.04	–0.09	0.17	0.15	–0.06	0.36	0.06	–0.21	0.31	0.02	–0.11	0.16	0.04	–0.14	0.21	0.07	–0.17	0.30
DN latency	0.00	–0.11	0.10	0.17	–0.02	0.35	0.00	–0.23	0.23	0.01	–0.11	0.12	0.05	–0.12	0.22	–0.01	–0.22	0.20
PV score	**0.20**	**0.03**	**0.36**	–0.11	–0.30	0.08	–0.12	–0.41	0.18	**0.24**	**0.05**	**0.41**	–0.04	–0.20	0.12	0.02	–0.24	0.28
PV latency	–0.06	–0.19	0.08	–0.04	–0.25	0.17	0.28	–0.09	0.59	**0.15**	**0.02**	**0.28**	0.09	–0.08	0.25	0.54	0.23	0.78
Math SE	0.09	–0.02	0.19	–0.11	–0.31	0.11	0.12	–0.08	0.30	0.05	–0.07	0.16	–0.02	–0.18	0.14	0.01	–0.17	0.18

Note. Est = estimate, lCI, uCI = lower and upper bounds of the 95% confidence interval, respectively. The statistically significant estimates are in bold.

Discussion

The Unified State Exam is a major milestone in the life of any school student in Russia. However, the origin of individual differences in USE performance remains largely unexplored. The current study addressed the question of the predictive power of intelligence, mathematical abilities, and mathematical self-efficacy in respect to USE performance. We compared students from regular state schools (non-selected students) and students who had passed strict selection and undertaken an advanced mathematical curriculum at a specialized school. Structural equation modeling was applied to assess to what extent cognitive abilities and mathematical skills explained the variation in exam performance. We also compared the structure of the relationships across mathematical abilities and cognitive characteristics in selected and non-selected students.

The average math and Russian USE scores in the group of the students from regular state schools corresponded well to the national statistics for USE scores (Federal Service for Supervision in Education and Science; Federal Institute of Pedagogical Measurements, 2012). The students from specialized schools scored high above the national average on both the math and Russian USEs. Schools in Russia differ in their formal status (e.g., gymnasia, lyceums, specialized schools). These schools often receive more funding, are better equipped, and provide advanced curricula. The students from such schools achieved higher average USE scores (Popova & Sheina, 2017; Sobkin, Adamchuk, Kolomiets, Likhanov, & Ivanova, 2010); for most school subjects, the difference between regular and advanced schools lies within 10 points on a 100-point scale.

In our study, the students from specialized schools displayed higher levels of taught and untaught cognitive characteristics and mathematical abilities. The design of our study did not allow us to find out to what extent these differences arose from the curricular instruction provided by specialized schools or from selection (entry to specialized schools based on high academic performance). Some results are more consistent with the selection explanation. For example, the group difference in non-symbolic and symbolic representation in our study is unlikely to have originated from instruction, as the numeric representations stop evolving by the end of primary school (Ashcraft & Moore, 2012; Dehaene, 2011; Friso-van den Bos et al., 2015).

There is an ongoing discussion whether the effect of cognitive ability on academic outcomes is uniform through the whole range of ability levels. One view is that the share of intelligence in the variation of specific cognitive abilities decreases at the top end of the ability distribution (Reynolds & Keith, 2007; te Nijenhuis & Hartmann, 2006). Our results do not support this hypothesis, consistent with much research that also does not support this view (Coyle, 2015; Coyle, Snyder, Pillow, & Kochunov, 2011; Karwowski & Gralewski, 2013; Robertson, Smeets, Lubinski, & Benbow, 2010). In the present study, cognitive predictors were associated with exam performance uniformly across the groups of selected and non-selected students, with mostly negligible effects in both groups. In contrast, the factor structure of cognitive characteristics differed across the groups: in the group of selected students, the associations among cognitive predictors were overall weaker. This result is likely due to the restricted variance in the selected groups, including ceiling ef-

fects: the whole range of the actual intelligence scores fits in the top quarter of the Raven's scale.

Within the groups, cognitive characteristics and mathematical abilities were only modestly related to USE performance. This is inconsistent with much research into academic achievement and exam performance in different countries. This research has demonstrated substantial links between academic achievement and general cognitive ability (Coyle & Pillow, 2008; Deary et al., 2007; Roth et al., 2015), and specific cognitive abilities (Coyle et al., 2013; Krapohl et al., 2014), including numerical representations (Halberda, Mazzocco, & Feigenson, 2008; Matthews et al., 2016). The reasons for this inconsistency are not clear. Further research is needed to explain why so little variance in USE scores was explained by the measures that have been shown to be more predictive in samples from other countries.

Although our study included a number of predictors, many other factors may contribute to the observed variation in USE scores. For example, research has suggested an important role for executive functions and working memory. Executive functions — updating, shifting and inhibition — along with working memory bring together basic cognitive functions that take part in information processing and perform cognitive and behavioral regulation (Baddeley, 2012; Banich, 2009; Best & Miller, 2010; Miyake & Friedman, 2012). These functions have been linked to learning in general (Alloway & Alloway, 2010; St Clair-Thompson & Gathercole, 2006) and to mathematical ability (Cragg & Gilmore, 2014; Mazzocco & Kover, 2007; Meltzer, 2018). Effortful control and executive functions have also been linked to academic achievement and the quality of teacher-student interaction at school (Liew, 2012).

In our study, the Corsi block test was the only measure providing any insight into the relationship between working memory and exam performance. We did not find an association between this measure and USE scores, in either group. This lack of association may be explained by the fact that the Corsi block task does not capture the executive component of working memory (Kessels, van den Berg, Ruis, & Brands, 2008; Vandierendonck, Kemps, Fastame, & Szmalec, 2004). It is possible that individual differences in working memory and executive functions explain part of the variance in exam performance and the tight association between mathematics and Russian USE scores within the groups.

Another group of factors that may explain variation in USE scores concerns emotional processes, such as mathematical and test anxiety. Mathematical anxiety emerges specifically in situations when a person has to deal with mathematical problems, either in a classroom or in real life (Maloney & Beilock, 2012), impairing learning and academic success (Ashcraft, 2002; Ma & Xu, 2004; Sherman & Wither, 2003; Wang et al., 2014). University students with higher levels of mathematical anxiety have been shown to have less precise representations of numerical magnitude (Núñez-Peña & Suárez-Pellicioni, 2014). Test anxiety may impair cognitive control and thereby undermine cognitive performance, unless compensatory strategies are used (Eysenck, Derakshan, Santos, & Calvo, 2007; Putwain, 2008, 2018; Putwain, Connors, & Symes, 2010).

Another group of factors, relevant to the USE, involves students' socioeconomic background. Students' academic success is associated with the socioeconomic

status of their family (Schoon, 2010; Sirin, 2005). Performance on standardized exam tests, like the SAT, is also associated with socioeconomic background (Zwick, 2004). One study revealed a 15-point difference in math USE scores in students from rich and poor families, controlling for school performance (Prakhov and Yudkevich, 2012). The students from affluent families had higher USE scores, even though at school they achieved the same marks as students from poor families. To boost children's performance on specific exams, many Russian families, both poor and wealthy, hire private tutors or send their children to preparatory courses. One study (Prakhov and Yudkevich, 2012) demonstrated that students from wealthier families achieved higher USE scores, even with the same amount of additional training. Research has demonstrated a 20% overall effect of extracurricular training (Prakhov, 2014).

The Unified State Exam in Russia aims to provide equal opportunities to all school students in Russia, in terms of entry to higher education and occupations. However, it is not yet clear whether this goal is fully achieved (Uvarov & Yastrebov, 2014). Further research is needed to gain a better understanding of factors that lead to the observed wide variation in exam performance. As many factors are interdependent, it remains a challenge to estimate the independent contribution of each factor, as well as their interactive effects. Large-scale population-based studies are needed to achieve this goal.

Conclusion

Students from schools with an advanced mathematical curriculum, in comparison with students from regular schools, performed better on cognitive tasks, appraised their mathematical abilities higher, and achieved higher exam scores, both on mathematics and Russian USEs. Cognitive measures, including acquired mathematical knowledge, number sense, and response speed differed on average between the two groups, but they did not explain much variation in USE scores within the groups.

Limitations

The present study combined the data of 496 school students; however, only 207 of them reported math USE scores and only 273 reported Russian USE scores. This limits the statistical power of analyses involving USE scores in our study, potentially leading to undetected small effects. In addition, the pattern of missing data may not be random in terms of why students did not report their exam scores. Another constraint of the study is the limited number of cognitive and non-cognitive characteristics that were used as predictors of exam performance. Future research may address other potential predictors, such as executive functions, test and math anxiety, and family wealth. In addition, further research is needed into potential cohort effects to follow up on updates of the USE. For example, the advanced form of the mathematics USE was introduced for better differentiation of high-performing students.

Acknowledgments

Funding: This work was supported by the Russian Ministry for Education and Science, project No.25.8663.2017/9.1

References

Alloway, T.P., & Alloway, R.G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Ashcraft, M.H., & Moore, A.M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*(2), 246–267. https://doi.org/10.1016/j.jecp.2011.08.005

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology, 63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Banich, M.T. (2009). Executive Function: The Search for an Integrated Account. *Current Directions in Psychological Science, 18*(2), 89–94. https://doi.org/10.1111/j.1467-8721.2009.01615.x

Best, J.R., & Miller, P.H. (2010). A Developmental Perspective on Executive Function: Development of Executive Functions. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bong, M., & Skaalvik, E.M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review, 15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S.S.J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences*, 23, 172–178. https://doi.org/10.1016/j.lindif.2012.07.021

Chernyavskaya, V.S., & Merkulov, D.S. (2015). Vzaimosvjaz' rezul'tatov EGJe i uspeshnosti uchebnoj dejatel'nosti vypusknikov shkoly [School leavers' Unified State Exam result and school success level of relationship]. *Sovremennye issledovanija social'nyh problem* [Contemporary research on social issues], *50*(6), 290–300. https://doi.org/10.12731/2218-7405-2015-6-25

Coyle, T.R. (2015). Relations among general intelligence (g), aptitude tests, and GPA: Linear effects dominate. *Intelligence*, 53, 16–22. https://doi.org/10.1016/j.intell.2015.08.005

Coyle, T.R., & Pillow, D.R. (2008). SAT and ACT predict college GPA after removing g. *Intelligence, 36*(6), 719–729. https://doi.org/10.1016/j.intell.2008.05.001

Coyle, T.R., Purcell, J.M., Snyder, A.C., & Kochunov, P. (2013). Non-g residuals of the SAT and ACT predict specific abilities. *Intelligence, 41*(2), 114–120. https://doi.org/10.1016/j.intell.2012.12.001

Coyle, T.R., Snyder, A.C., Richmond, M.C., & Little, M. (2015). SAT non-g residuals predict course specific GPAs: Support for investment theory. *Intelligence*, 51, 57–66. https://doi.org/10.1016/j.intell.2015.05.003

Coyle, T.R., Snyder, A., Pillow, D., & Kochunov, P. (2011). SAT predicts GPA better for high ability subjects: Implications for Spearman's Law of Diminishing Returns. *Personality and Individual Differences, 50*(4), 470–474. https://doi.org/10.1016/j.paid.2010.11.009

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Deary, I.J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*(1), 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Dehaene, S. (2011). *The number sense: how the mind creates mathematics* (Rev. and updated ed). New York: Oxford University Press.

Dmitrieva, E.S., & Gelman, V.Y. (2015). Dinamika izmenenij vzaimosvjazi jemocional'nogo intellekta s rezul'tatami EGJe v hode adaptacii shkol'noj sistemy k ego vvedeniju [Dynamics of the relationship of emotional intelligence with the results of the state exam during the adaptation of school system to its introduction]. *Jeksperimental'naja psihologija* [Experimental psychology], *8*(2), 139–150. https://doi.org/10.17759/exppsy.2015080210

Fazio, L.K., Bailey, D.H., Thompson, C.A., & Siegler, R.S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. Journal of Experimental Child Psychology, 123, 53–72. https://doi.org/10.1016/j.jecp.2014.01.013

Federal Service for Supervision in Education and Science. Federal Institute of Pedagogical Measurements. (2012). *Itogovyj analiticheskij otchet o rezul'tatah Edinogo gosudarstvennogo jekzamena 2012 goda* [Final analytical report on the results of the Unified State Exam in 2012], 31. Moscow.

Feigenson, L., Libertus, M.E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, *7*(2), 74–79. https://doi.org/10.1111/cdep.12019

Friso-van den Bos, I., Kroesbergen, E.H., Van Luit, J.E.H., Xenidou-Dervou, I., Jonkman, L.M., Van der Schoot, M., & Van Lieshout, E.C.D.M. (2015). Longitudinal development of number line estimation and mathematics performance in primary school children. *Journal of Experimental Child Psychology,* 134, 12–29. https://doi.org/10.1016/j.jecp.2015.02.002

Gadbois, S.A., & Sturgeon, R.D. (2011). Academic self-handicapping: Relationships with learning specific and general self-perceptions and academic performance over time: Academic self-handicapping. *British Journal of Educational Psychology, 81*(2), 207–222. https://doi.org/10.1348/000709910X522186

Gordeeva, T.O., & Osin, E.N. (2012). Osobennosti motivacii dostizhenija i uchebnoj motivacii studentov, demonstrirujushhih raznye tipy akademicheskih dostizhenij (EGJe, pobedy v olimpiadah, akademicheskaja uspevaemost') [Differences in achievement motivation and learning motivation in students exhibiting different types of academic attainment (Unified State Examination (USE) scores, academic competition results, academic records)]. *Psikhologicheskie Issledovaniya* [Psychological Studies], *5*(24).

Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence, 35*(5), 1111–1122. https://doi.org/10.1016/j.adolescence.2012.02.016

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences, 109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665–668. https://doi.org/10.1038/nature07246

Kaptsov, A., & Kolesnikova, E. (2009). Vzaimosvjaz' urovnja intellekta studentov i rezul'tatov sdachi EGJe [The interrelation of IQ of students and results of delivery USE]. *Vestnik Samarskoj gosudarstvennoj akademii. Serija "Psihologija"* [Bulletin of Samara State Academy. Serias "Psycholohy"], *6*(2), 87–94.

Karwowski, M., & Gralewski, J. (2013). Threshold hypothesis: Fact or artifact? *Thinking Skills and Creativity*, 8, 25–33. https://doi.org/10.1016/j.tsc.2012.05.003

Kessels, R.P. C., van den Berg, E., Ruis, C., & Brands, A.M.A. (2008). The Backward Span of the Corsi Block-Tapping Task and Its Association With the WAIS-III Digit Span. *Assessment, 15*(4), 426–434. https://doi.org/10.1177/1073191108315611

Khavenson, T., & Solovyova, A. (2014). Svjaz' rezul'tatov Edinogo gosudarstvennogo jekzamena i uspevaemosti v vuze [Studying the Relation between the Unified State Exam Points and Higher Education Performance]. *Voprosy obrazovanija* [Issues of education], 1, 176–199.

Kochergina, E.V., Nye, J., & Orel, E.A. (2013). Faktory Bol'shoj pjaterki kak psihologicheskie prediktory akademicheskoj uspevaemosti studentov vuzov [The Big Five traits as predictors of academic achievements in university students*]. Psihologicheskie issledovanija* [Psychological Studies], *6*(27), 4.

Koenig, K.A., Frey, M.C., & Detterman, D.K. (2008). ACT and general cognitive ability. *Intelligence, 36*(2), 153–160. https://doi.org/10.1016/j.intell.2007.03.005

Krapohl, E., Rimfeld, K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.-B., … Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences, 111*(42), 15273–15278. https://doi.org/10.1073/pnas.1408777111

Libertus, M.E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica, 141*(3), 373–379. https://doi.org/10.1016/j.actpsy.2012.09.009

Liew, J. (2012). Effortful Control, Executive Functions, and Education: Bringing Self-Regulatory and Social-Emotional Competencies to the Table: Self-Regulation and Education. *Child Development Perspectives, 6*(2), 105–111. https://doi.org/10.1111/j.1750-8606.2011.00196.x

Lourenco, S.F., & Bonny, J.W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Developmental Science, 20*(4), e12418. https://doi.org/10.1111/desc.12418

Malanchini, M., Wang, Z., Voronin, I., Schenker, V.J., Plomin, R., Petrill, S.A., & Kovas, Y. (2017). Reading self-perceived ability, enjoyment and achievement: A genetically informative study of their reciprocal links over time. *Developmental Psychology, 53*(4), 698–712. https://doi.org/10.1037/dev0000209

Marsh, H.W., & Martin, A.J. (2011). Academic self-concept and academic achievement: Relations and causal ordering: Academic self-concept. *British Journal of Educational Psychology, 81*(1), 59–77. https://doi.org/10.1348/000709910X503501

Marsh, H.W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The Big-Fish-Little-Pond Effect: Persistent Negative Effects of Selective High Schools on Self-Concept After Graduation. *American Educational Research Journal, 44*(3), 631–669. https://doi.org/10.3102/0002831207306728

Matsuta, V.V., Bogomaz, S.A., & Sudneva, O.Y. (2014). Rol' intellektual'nyh i lichnostnyh faktorov v dostizhenii vysokoj rezul'tativnosti EGJe po matematike [The role of intellectual and personal factors in achieving high performance in the USE in mathematics]. *Sibirskij psihologicheskij zhurnal* [Psychological journal of Syberia], 52, 52–66.

Matthews, P. G., Lewis, M. R., & Hubbard, E. M. (2016). Individual Differences in Nonsymbolic Ratio Processing Predict Symbolic Math Performance. *Psychological Science, 27*(2), 191–202. https://doi.org/10.1177/0956797615617799

Merkley, R., & Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: evidence from brain and behavior. *Current Opinion in Behavioral Sciences*, 10, 14–20. https://doi.org/10.1016/j.cobeha.2016.04.006

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology, 84*(4), 591–611. https://doi.org/10.1111/bjep.12047

O'Connell, M. (2018). The power of cognitive ability in explaining educational test performance, relative to other ostensible contenders. *Intelligence*, 66, 122–127. https://doi.org/10.1016/j.intell.2017.11.011

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*(1), 29–48. https://doi.org/10.1080/01443410.2013.797339

Pereyaslavskaya, L., & Pereyaslavskiy, V. (2014). Issledovanie korreljacij mezhdu rezul'tatami EGJe po matematike abiturientov i ih uspevaemost'ju v vuze [The study of correlation between the result of exams in mathematics and students´ academic performance in higher education]. *Vestnik associacii vuzov turizma i servisa* [Bulletin of the association of universities of tourism and service], *8*(4), 49–55. https://doi.org/10.12737/6475

Poldin, O. (2011). Prognozirovanie uspevaemosti v vuze po rezul'tatam EGJe [Predicting success in college on the basis of the results of unified national exam]. *Prikladnaja jekonometrika*, *21*(1), 56–69.

Popova, E., & Sheina, M. (2017). Ucheba v sil'noj shkole — garantija vysokih akademicheskih rezul'tatov v vuze? [Does Studying in a Strong School Guarantee Good College Performance?]. *Voprosy obrazovaniya* [Issues of education], 1, 128–157. https://doi.org/10.17323/1814-9545-2017-1-128-157

Prakhov, I.A. (2014). Vlijanie investicij v dopolnitel'nuju podgotovku na rezul'taty EGJe [Effects of Investments in Preparation Courses on the USE Scores]. *Voprosy obrazovanija* [Issues of education], 3, 74–99. https://doi.org/10.17323/1814-9545-2014-3-74-99

Prakhov, I.A., & Yudkevich, M.M. (2012). Vlijanie dohoda domohozjajstv na rezul'taty EGJe i vybor vuza [The role of household income in the USE performance and the choice of university]. *Voprosy obrazovanija* [Issues of education], 1, 126–147. https://doi.org/10.17323/1814-9545-2012-1-126-147

Ratcliff, R. (1993). Methods for Dealing With Reaction Time Outliers. *Psychological Bulletin, 114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Raven, J. (2003). Raven Progressive Matrices. In R.S. McCallum (Ed.), Handbook of Nonverbal Assessment (pp. 223–237). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-0153-4_11

Reynolds, M.R., & Keith, T.Z. (2007). Spearman's law of diminishing returns in hierarchical models of intelligence for children and adolescents. *Intelligence, 35*(3), 267–281. https://doi.org/10.1016/j.intell.2006.08.002

Rimfeld, K., Dale, P.S., & Plomin, R. (2015). How specific is second language-learning ability? A twin study exploring the contributions of first language achievement and intelligence to second language achievement. *Translational Psychiatry, 5*(9), e638–e638. https://doi.org/10.1038/tp.2015.128

Robertson, K.F., Smeets, S., Lubinski, D., & Benbow, C.P. (2010). Beyond the Threshold Hypothesis: Even Among the Gifted and Top Math/Science Graduate Students, Cognitive Abilities, Vocational Interests, and Lifestyle Preferences Matter for Career Choice, Performance, and Persistence. *Current Directions in Psychological Science, 19*(6), 346–351. https://doi.org/10.1177/0963721410391442

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence,* 53, 118–137. https://doi.org/10.1016/j.intell.2015.09.002

Rousseeuw, P.J., & Hubert, M. (2011). Robust statistics for outlier detection: Robust statistics for outlier detection. Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery, 1*(1), 73–79. https://doi.org/10.1002/widm.2

Saprykina, T.A. (2017). O perehode shkola-vuz prediktory uspevaemosti studentov-pervokursnikov [The transition "school-university": Predictors of academic performance of first-year students]. *Vysshee obrazovanie v Rossii* [Higher education in Russia], 6, 77–87.

Sasanguie, D., Göbel, S.M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology, 114*(3), 418–431. https://doi.org/10.1016/j.jecp.2012.10.012

Schult, J., & Sparfeldt, J.R. (2016). Do non-g factors of cognitive ability tests align with specific academic achievements? A combined bifactor modeling approach. *Intelligence, 59*, 96–102. https://doi.org/10.1016/j.intell.2016.08.004

Schwinger, M., Wirthwein, L., Lemmer, G., & Steinmayr, R. (2014). Academic self-handicapping and achievement: A meta-analysis. *Journal of Educational Psychology, 106*(3), 744–761. https://doi.org/10.1037/a0035832

Seaton, M., Parker, P., Marsh, H.W., Craven, R.G., & Yeung, A.S. (2014). The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology, 34*(1), 49–72. https://doi.org/10.1080/01443410.2013.825232

Siegler, R.S., & Lortie-Forgues, H. (2014). An Integrative Theory of Numerical Development. *Child Development Perspectives, 8*(3), 144–150. https://doi.org/10.1111/cdep.12077

Sobkin, V.S., Adamchuk, D., Kolomiets, Y.O., Likhanov, I.D., & Ivanova, A.I. (2010). Sociologicheskoe issledovanie rezul'tatov EGJe [Sociological study of USE performance]. In V. S. Sobkin (Ed.), *Psihologija obrazovanija. Trudy po psihologii obrazovanija* [Educational psychology. Works on educational psychology] (Vol. 14, pp. 9–30). Moscow, Institut sociologii obrazovanija RAO.

St Clair-Thompson, H.L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Suárez-Álvarez, J., Fernández-Alonso, R., & Muñiz, J. (2014). Self-concept, motivation, expectations, and socioeconomic level as predictors of academic performance in mathematics. *Learning and Individual Differences*, 30, 118–123. https://doi.org/10.1016/j.lindif.2013.10.019

te Nijenhuis, J., & Hartmann, P. (2006). Spearman's "Law of Diminishing Returns" in samples of Dutch and immigrant children and adults. *Intelligence, 34*(5), 437–447. https://doi.org/10.1016/j.intell.2006.02.002

Tosto, M.G., Asbury, K., Mazzocco, M.M.M., Petrill, S.A., & Kovas, Y. (2016). From classroom environment to mathematics achievement: The mediating role of self-perceived ability and subject interest. *Learning and Individual Differences*, 50, 260–269. https://doi.org/10.1016/j.lindif.2016.07.009

Tosto, M.G., Petrill, S.A., Halberda, J., Trzaskowski, M., Tikhomirova, T.N., Bogdanova, O. Y., … Kovas, Y. (2014). Why do we differ in number sense? Evidence from a genetically sensitive investigation. *Intelligence*, 43, 35–46. https://doi.org/10.1016/j.intell.2013.12.007

Vandierendonck, A., Kemps, E., Fastame, M.C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95*(1), 57–79. https://doi.org/10.1348/000712604322779460

Wang, J. (Jenny), Odic, D., Halberda, J., & Feigenson, L. (2016). Changing the precision of preschoolers' approximate number system representations changes their symbolic math per-

formance. *Journal of Experimental Child Psychology*, 147, 82–99. https://doi.org/10.1016/j.jecp.2016.03.002

Zamkov, O. O., & Pereseckij, A. A. (2013). EGJe i akademicheskie uspehi studentov bakalavriata MIJeF NIU VShJe [USE and academic achievement in bacchalaureate students of ICEF NRU HSE]. *Prikladnaja jekonometrika*, *30*(2), 93–114.

Original manuscript received August 07, 2018
Revised manuscript accepted October 30, 2018
First published online December 30, 2018