PSYCHOLOGY IN RUSSIA: STATE OF THE ART

Volume 11 Issue 4 2018

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

Russian Psychological Society
Lomonosov Moscow State University

EDUCATION AND HUMAN DEVELOPMENT

Effortful Control in Primary Schoolchildren: Links with Personality, Problem Behaviour, Academic Achievement, and Subjective Well-Being

Olga S. Kornienko[a,b*], Evgeniya N. Petrenko[a], Irina V. Leto[a], Natalija A. Fedorova[c], and Helena R. Slobodskaya[a,b]

[a] *Institute of Physiology and Basic Medicine, Novosibirsk, Russia*

[b] *Novosibirsk State University, Novosibirsk, Russia*

[c] *Geraskov Children's Clinical Hospital No. 4, Novosibirsk, Russia*

[*] Corresponding author. E-mail: olfa@mail.ru

Background: Effortful control is a core aspect of self-regulation and refers to the ability to voluntarily regulate behaviour and attention, measured by temperament questionnaires. Although the Temperament in Middle Childhood Questionnaire is widely used in different countries, this measure has not been fully explored. Most research on the links of effortful control with personality and important outcomes has been carried out in Western nations; the possibility of extending these findings to other cultures requires study.

Objective: To examine effortful control and its relations to personality and well-being in a community sample of primary schoolchildren in Russia.

Design: Parents of 7–10-year-olds (N = 614) completed the abbreviated Effortful Control scale of the TMCQ, the Inventory of Child Individual Differences–Short version, and the Strengths and Difficulties Questionnaire (SDQ); teachers provided SDQ data and school grades; children completed the Students' Life Satisfaction Scale.

Results: The findings supported a four-factor structure of Effortful Control, including Attention Focusing, Inhibitory Control, Activation Control, and Low-Intensity Pleasure. Effortful Control was associated with the personality traits of Conscientiousness, Agreeableness, and Openness, and also with Positive Emotions and low Neuroticism. Effortful Control was also associated with academic achievement, subjective well-being, and lower levels of externalising and internalising problems. Structural modelling showed that Attentional Control contributed to problem behaviour and subjective well-being; Inhibitory Control contributed to externalising problems; and Activation Control contributed to academic achievement.

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

doi: 10.11621/pir.2018.0401
http://psychologyinrussia.com

Conclusion: Effortful Control and its components were strongly related to higher- and lower-order personality traits. The findings confirmed the important role of effortful control in the academic success and well-being of Russian primary schoolchildren.

Keywords: effortful control, middle childhood, personality, well-being, problem behaviour, academic achievement, life satisfaction

Introduction

Effortful control is an integral part of temperament in children, together with Negative Affectivity and Positive Affectivity. It refers to the ability to choose a course of action under conditions of internal conflict, to detect errors, and to plan (Rothbart, 2007). Effortful control is also defined as the ability to voluntarily regulate behaviour and attention, as seen in the inhibition of a dominant response and activation of a subdominant response (Rothbart, Ahadi, & Evans, 2000). This construct emerged initially from psychometric studies of parent reports (Rothbart, 2007), showing that effortful control develops throughout the early years along with the maturation of attentional and inhibitory mechanisms (Rothbart, Sheese, Rueda, & Posner, 2011). Factor-analytic studies provided evidence that the content of the Effortful Control factor is largely similar across ages; the differences mostly reflect changes in the behavioural repertoire (Putnam, Ellis, & Rothbart, 2001). In all ages after early childhood, effortful control includes traits reflecting the capacity to focus and to shift attention, and the inhibitory control trait, reflecting an ability to suppress inappropriate actions or responses. In children, but not in adolescents and adults, effortful control includes the trait of low-intensity pleasure, reflecting enjoyment related to low stimulus intensity, rate, novelty, and incongruity; and the perceptual sensitivity trait, which encompasses the detection of subtle stimuli from the external environment. Activation control, reflecting the capacity to perform an action when there is a strong tendency to avoid it, is included in the effortful control measure from middle childhood onwards.

Effortful Control in middle childhood includes five components — Attention Focusing, Inhibitory Control, Activation Control, Low-Intensity Pleasure, and Perceptual Sensitivity — measured by the Temperament in Middle Childhood Questionnaire (TMCQ, Simonds, Kieras, Rueda, & Rothbart, 2007; Simonds & Rothbart, 2004). Although the TMCQ is widely used in different countries, this measure has not been fully explored, and has dubious psychometric properties. A recent study of a large sample of 9-year-old North American children (Kotelnikova, Olino, Klein, Mackrell, & Hayden, 2016) demonstrated that the structure of the TMCQ differed from the original three-factor model developed by Rothbart and colleagues (Rothbart, 2007; Rothbart et al., 2000); it also showed that Effortful Control items did not form five facets as posited by Simonds and Rothbart (2004). Thus, the structure of Effortful Control in middle childhood requires further investigation.

The first aim of this study was to investigate the psychometric properties of the Russian version of the TMCQ Effortful Control scale. We examined the structure of Effortful Control in primary schoolchildren, and also whether gender and effects in the Russian cultural context are similar to those found in other cultures. A meta-analysis of gender differences in child temperament revealed a

large difference in effortful control favouring girls and moderate differences in inhibitory control and perceptual sensitivity (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Because most evidence on the development of effortful control comes from experimental studies of attention and inhibitory control (Rothbart et al., 2011), our analysis of age differences in parent-report measures was primarily exploratory.

Traditionally, research on individual differences has conceptualized childhood traits as temperament, and adult traits as personality. Work with parental and teacher descriptions of children during recent decades has demonstrated the robustness of the Five Factor Theory of personality development (Caspi & Shiner, 2006; Halverson et al., 2003; Kohnstamm, Halverson, Mervielde, & Havill, 1998). It has been also recognized that both temperament and personality traits are manifestations of the same basic dimensions (Caspi & Shiner, 2006; Shiner & DeYoung, 2013). For example, Digman and Shmelyov (1996) demonstrated the structural similarity of temperament and the Big Five personality factors in a sample of Russian primary schoolchildren. Self-regulatory behaviours encompassed by effortful control are consistently related to the personality domain of Conscientiousness (Halverson et al., 2003; Shiner & DeYoung, 2013). Effortful control is also related to the domain of Agreeableness, which reflects self-regulation in positive social interactions (Caspi & Shiner, 2006). However, little is known about how different aspects of effortful control relate to higher- and lower-order personality traits. Therefore, our second aim was to examine the relationships of effortful control and its components with personality in middle childhood.

The biological basis of temperament and personality appear to be shared across cultures, but outcomes may vary depending on cultural context and personal experience (Rothbart, 2007). There is increasing evidence that children's effortful control predicts important life outcomes. In middle childhood, these include problem behaviours, academic achievement, and life satisfaction, which are also considered to be core domains of child well-being (O'Hare & Gutierrez, 2012; Rees et al., 2012). In childhood, problem behaviours are widespread; they may interfere with everyday life and prevent children from accomplishing important developmental tasks. Research has shown that most childhood problems fall into two main groups: The behavioural or externalising problems include destructiveness, aggression, hyperactivity, and antisocial behaviours; the emotional or internalising problems involve worries, fears, depressive features, and psychosomatic symptoms (Goodman & Scott, 2012). Many studies have established strong links between a lack of effortful control and children's problem behaviour; links with externalising problems tend to be stronger than those with internalising problems (Rothbart, 2007).

There is consistent evidence that children's effortful control is associated with academic outcomes such as self-regulated learning and mathematics and literacy skills (Eisenberg, Valiente, & Eggum, 2010). Academic trajectories tend to be established early in preschool or elementary school (McClelland, Acock, & Morrison, 2006), and the transition from preschool to elementary school poses greater demands for effortful control, requiring children to focus their attention for a longer time, to inhibit inappropriate behaviour in class, and to do home-

work. However, only a few studies have examined the contribution of different components of effortful control to academic achievement (Oberle & Schonert-Reichl, 2013; Sánchez-Pérez, Fuentes, Pina, López-López, & González-Salinas, 2015).

A longitudinal study of a cohort of 1,000 people from birth to age 32 showed that childhood self-control predicts physical and mental health, as well as life satisfaction in adulthood (Moffitt et al., 2011; Moffitt, Poulton, & Caspi, 2013). Life satisfaction, a term which is often used interchangeably with subjective well-being and happiness, is a subjective evaluation of overall quality of life (Proctor, Linley, & Maltby, 2009). While early research on child development was largely focused on important outcomes and children's quality of life from an adult perspective, contemporary studies emphasize the importance of measuring children's subjective well-being and/or life satisfaction (Rees et al., 2012). However, very little is known about the association between effortful control and subjective well-being in childhood.

Most research on the links between effortful control and important life outcomes has been carried out in Western nations; the possibility of extending these findings to other cultures requires study. Therefore, the third aim of this study was to examine the relations between effortful control and well-being, including problem behaviour, subjective well-being, and academic achievement, in a community sample of Russian primary schoolchildren. Because in Russia formal schooling typically starts around age 7, this study focuses on middle childhood (7–10 years). Drawing on prior research, we anticipated significant associations between a lack of effortful control and problem behaviour, specifically with externalising problems. We expected that effortful control would be related to academic achievement in primary school, although a priori predictions regarding components of effortful control could not be specified, since the results of previous studies have been inconclusive. Addressing the links between effortful control and children's subjective well-being, our study was largely exploratory.

Method

Participants and Procedures

The sample consisted of 614 children from 7 to 10 years old (55% female); the average age was 8.6 ($SD = 1.1$). Participants were from two Siberian cities, Novosibirsk (65%) and Novokuznetsk (28%), and from nearby rural regions. Most of the children (68%) lived with both biological parents, 20% with a single mother, 10% with a mother and a stepfather, and the rest with other caregivers; 53% had one or more siblings. Regarding education, 8% of the mothers and 12 % of the fathers had 10 years of schooling or less; 27% of the mothers and 29% of the fathers had secondary vocational education; 65% of the mothers and 59% of the fathers had university education. Twenty-two percent of the mothers and 36% of the fathers were unskilled or manual workers; 48% of the mothers and 35% of the fathers had professional occupations; and 13% of the mothers and 26% of the fathers had administrative occupations. Seventeen percent of the mothers and 3% of the fathers were unemployed.

Participants were recruited through schools, social websites, and in person. Parents were asked to complete the Inventory of Child Individual Differences (ICID), the Strengths and Difficulties Questionnaire (SDQ), the Effortful Control (EC) scale, and a brief demographic questionnaire. Parents of 69 children also completed the EC scale 12 months later. Teachers completed the SDQ and provided information on school grades for 437 children. The Students' Life Satisfaction Scale was completed by 330 children. All questionnaires were administered as paper and pencil instruments. Informed consent was obtained from all individual participants included in the study.

Measures

Effortful Control. The Effortful Control scale of the Temperament in Middle Childhood Questionnaire (TMCQ, Simonds et al., 2007; Simonds & Rothbart, 2004) consists of 48 items relating to five subscales. These include Activation Control (example items: "Has a hard time working on an assignment s/he finds boring" (reversed scored) and "Can make him/herself do homework, even when s/he wants to play"), Attention Focusing (example items: "Needs to be told to pay attention" and "Is easily distracted when listening to a story", both reversed scored), Inhibitory Control (example items: "Can stop him/herself from doing things too quickly" and "Has an easy time waiting to open a present"), Low-Intensity Pleasure (example items: "Enjoys looking at books" and "Likes the sound of poems"), and Perceptual Sensitivity (example items: "Notices things others don't notice" and "Notices when parents are wearing new clothing"). Parents were asked to rate their child, using a 5-point scale ranging from "almost always untrue of your child" to "almost always true of your child", with an additional option of "not applicable". More information about TMCQ is available on Mary Rothbart's Temperament Questionnaires website. We deleted items related to situations where parents were often absent, such as child behaviour in class (e.g., "Needs to be told by the teacher to pay attention") and with peers (e.g., "Can apologize or shake hands after a fight"). The remaining 19 items were translated from English to Russian and subsequently back-translated. The original and the back-translated copies were compared, and the Russian translation was revised on the basis of the observed discrepancies.

Personality. The Inventory of Child Individual Differences–Short version (ICID-S; Slobodskaya & Zupančič, 2010) is an age- and culture-decentred measure of child personality. The ICID-S includes 62 items rated on a 7-point scale ranging from "much less than the average child" to "much more than the average child". The items represent 15 mid-level traits that form the five higher-order factors. Extraversion is comprised of the scales Activity, Positive Emotions, and Sociable. Disagreeableness includes Antagonism, Strong Willed, and reversed Considerate. Conscientiousness is comprised of Achievement, Compliant, Organized, and reversed Distractible. Neuroticism includes Fearful, Shy, and Negative Affect. Openness is comprised of Intelligent and Open to Experience. The Russian version has been validated, supporting good reliability of the scales (Slobodskaya & Zupančič, 2010). In the present study, alphas for five higher-order ICID-S scales ranged from

.83 to .90 with a mean of .86; alphas for 15 mid-level ICID-S scales ranged from .73 to .85 with a mean of .79.

Problem behaviour. The Strengths and Difficulties Questionnaire (SDQ, Goodman, 2001) is a widely used measure for assessing common emotional and behavioural problems in children. The SDQ 20 items on difficulties are divided into four scales: Emotional Symptoms, Peer Problems, Conduct Problems, and Hyperactivity; all are summed to generate a Total Difficulties score. Each subscale consists of five statements, which describe the characteristic behaviour of the child in the last six months. Respondents marked each statement as "not true", "somewhat true", or "certainly true". The Internalising Problems scale includes Emotional Symptoms and Peer Problems; the Externalising Problems scale includes Conduct Problems and Hyperactivity. The Russian version has been validated in a stratified random sample (Goodman, Slobodskaya, & Knyazev, 2005). In the present study, we used parent- and teacher-versions of the SDQ, because information from more than one source can help to obtain more robust findings. Alphas for the SDQ scales ranged from .68 to .88.

Academic achievement. Academic achievement was evaluated by an internationally recognized measure, grade point average (GPA), the mean of the Russian, math, and literature grades for the year. These three disciplines are core academic subjects in primary school. It should be noted that the restricted range of grades is a limitation, however, high internal consistency reliability ($\alpha = .85$) provides justification for its use.

Subjective well-being. Subjective child well-being was evaluated by the Russian version of the Students' Life Satisfaction Scale (SLSS, Huebner, 1991), which consists of seven statements rated on a 6-point scale ranging from "strongly disagree" to "strongly agree". The statements are overall evaluations of the young person's life, for example: "My life is going well". The Russian version has been validated, supporting good reliability of the scale (Leto, Petrenko, & Slobodskaya, 2018). In the present study, alpha was .74.

Results

Structure of the Effortful Control Scale

We tested an a priori five-factor structure of Effortful Control using confirmatory factor analysis with AMOS 17.0 software (Arbuckle, 2008). This model did not, however, show acceptable fit: $\chi^2 = 479.6$, $df = 141$, $p < .001$, CFI = .74, RMSEA = .074. An examination of factor loadings and modification indices led to the exclusion of the Perceptual Sensitivity subscale; one item was deleted from the Inhibitory Control subscale and two items from the Activation Control subscale. The final model of Effortful Control (EC) comprised four latent variables — Attention Focusing (AF), Inhibitory Control (IC), Activation Control (AC), and Low-Intensity Pleasure (LIP) — with one additional cross-loading of Item 5, "Can make him/herself do homework, even when s/he wants to play" on the IC factor (*Figure 1*).

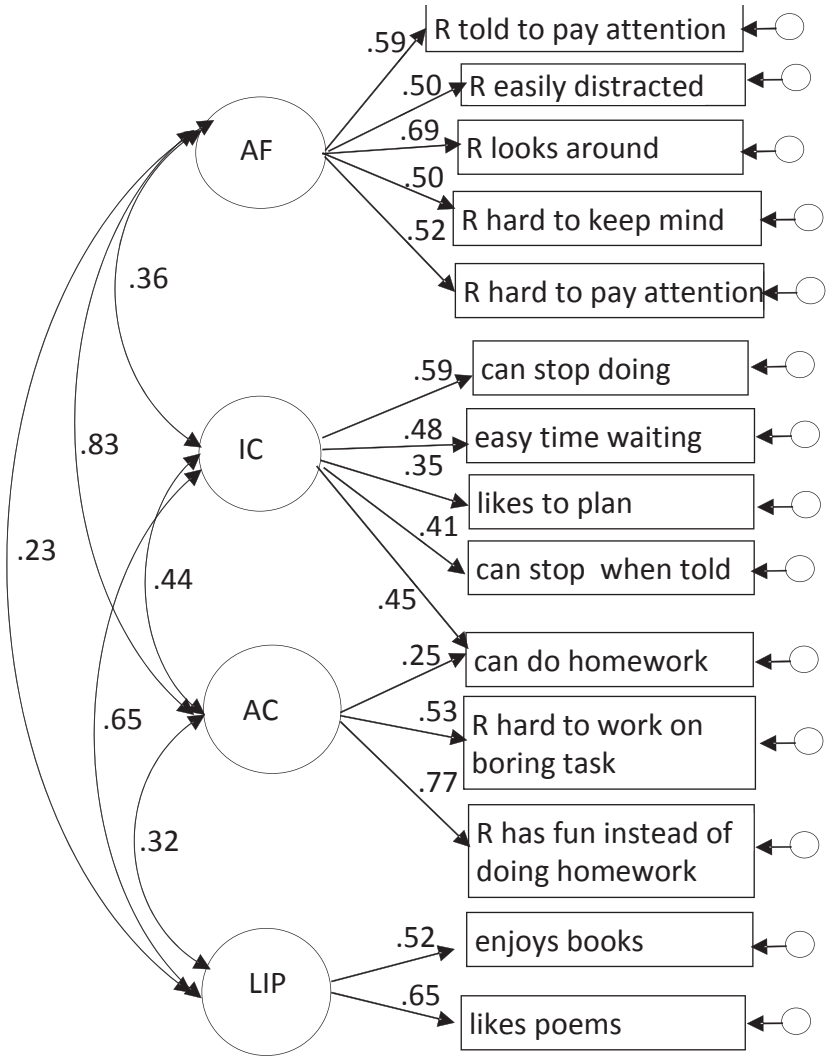


Figure 1. Structural model of the Effortful Control scale in Russian primary schoolchildren (N = 614).

Note. Values are standardised parameter estimates. AF — Attention Focusing, IC — Inhibitory Control, AC — Activation Control, LIP — Low-Intensity Pleasure. R = reversed.

The fit of the final model was satisfactory: $\chi^2 = 178.65$, $df = 70$, $p < .001$, CFI = .919, RMSEA = .050. Cronbach's alpha coefficients for the 14-item EC scale and four subscales, AF, IC, AC, and LIP, were .76, .70, .59, .60 and .50, respectively. One-year test–retest correlations, obtained for 69 children whose parents completed the Effortful Control scale one year later, were .56, .45, .44 and .64 for EC, AF, IC and LIP, respectively; all *p*s < .001, and $r = .40$ for AC, $p < .01$.

To assess the effect of demographic variables, 2 (gender) × 4 (ages 7, 8, 9, and 10) analyses of variance were performed. The effect size was estimated by eta

Table 1

Descriptive Statistics and Gender Differences for Effortful Control Measures

	Total sample (*N* = 614)	**Girls (*N* = 338)**	**Boys (*N* = 276)**		
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*[a]	η^2
Effortful Control	3.18 (.55)	3.26 (.54)	3.09 (.55)	15.47***	.025
Attention Focusing	2.97 (.75)	3.06 (.74)	2.86 (.75)	10.65***	.017
Inhibitory Control	3.34 (.71)	3.37 (.67)	3.31 (.74)	1.03	.002
Activation Control	3.06 (.84)	3.16 (.81)	2.94 (.87)	10.44**	.017
Low-Intensity Pleasure	3.72 (.96)	3.86 (.93)	3.54 (.97)	18.06***	.029

Note. [a]*Degrees of freedom were (1,613). **p < .01; ***p < .001.*

squared (η^2). Results showed that age effect and gender-by-age interactions were not significant. Table 1 shows descriptive statistics along with gender differences. Gender differences were significant for the Effortful Control scale and all subscales except Inhibitory Control, with girls scoring higher than boys. Gender effects accounted for 2.5% of the variance or less.

Relations Between Effortful Control and Personality Traits

Table 2 shows the correlation coefficients between Effortful Control and personality traits. Effortful Control and its subscales had high to medium positive associations with the Conscientiousness domain. Negative correlations with the Disagreeableness domain were medium, whereas negative correlations with the Neuroticism domain were of low to medium magnitude. Effortful Control and its subscales were positively correlated with the Openness domain; correlations with the trait Intelligent were mostly higher than those with trait Open to Experience. The correlations between Effortful Control and traits from the Extraversion domain were mostly insignificant; however, the trait of Positive Emotions had low to medium correlations with Effortful Control and its subscales.

Relationship of Effortful Control to Problem Behaviour, Academic Achievement, and Subjective Well-Being

Table 3 shows bivariate relationships between parent-reported effortful control and problem behaviours reported by parents and teachers. All significant correlations were negative; for both parent and teacher reports, correlations for externalising problems were higher than for internalising problems. Correlations of Effortful Control measures with teacher SDQ scales were lower than those with parent SDQ

Table 2

Correlations Between the Effortful Control and ICID-S Scales

	EC	AF	IC	AC	LIP
Extraversion	*.04*	*–.00*	*.00*	.10	*.07*
Activity	*–.09*	–.11	*–.06*	*.02*	–.03
Sociable	.01	–.02	*–.02*	*.06*	*.04*
Positive emotions	.18	.13	*.09*	*.15*	.18
Disagreeableness	–.38	–.26	–.28	–.36	–.21
Considerate	.23	.14	.18	.20	.19
Antagonism	–.33	–.24	–.24	–.30	–.16
Strong Willed	–.36	–.25	–.26	–.36	–.16
Conscientiousness	**.62**	**.49**	**.47**	**.51**	.24
Achievement	**.54**	.39	**.44**	**.46**	.22
Organized	**.53**	**.40**	**.41**	**.45**	.21
Compliant	**.48**	.35	.39	.38	.21
Distractible	**–.55**	**–.52**	–.32	**–.44**	–.17
Neuroticism	–.26	–.20	–.16	–.29	–.09
Fearful	–.14	–.13	*–.07*	–.17	–.01
Shy	–.11	–.12	*–.00*	–.18	–.02
Negative Affect	–.30	–.19	–.26	–.27	–.16
Openness	.32	.29	.16	.26	.20
Open to experience	.18	.14	*.09*	.12	.19
Intelligent	.37	.36	.18	.34	.17

Note. EC — Effortful Control, AF — Attention Focusing, IC — Inhibitory Control, AC — Activation Control, LIP — Low-Intensity Pleasure. All correlations greater than |.15| are significant at $p < .001$; correlations greater than |.09| are significant at $p < .05$; correlations greater than |.40| are in bold; non-significant correlations are in italics.

scales. Table 3 also shows that, with the exception of Inhibitory Control, all Effortful Control measures were positively correlated with children's subjective well-being and academic achievement.

Next, we tested structural models in which effortful control was expected to predict child well-being as indexed by the latent constructs of problem behaviour, academic achievement, and subjective well-being. We used structural equation modelling (SEM) because this approach allowed us to explore several outcomes simultaneously. Given that each child at the same time has distinct levels of academic achievement, subjective well-being, and problem behaviour, it would be most appropriate to analyse the contribution of effortful control to these components of child well-being in the same model. In addition, the advantage of SEM is that it estimates how well the model fits the empirical data.

Table 3

Correlations Between Effortful Control and Well-Being

SDQ scales[a]	EC	AF	IC	AC	LIP
Internalising Problems	–.28*** (–.11*)	–.30*** (–.12*)	–.15*** (–.01)	–.27*** (–.13**)	–.03 (–.02)
Emotional Problems	–.30*** (–.09)	–.33*** (–.12**)	–.15*** (–.03)	–.25*** (–.08)	–.06 (.01)
Peer Problems	–.17*** (–.09)	–.16*** (–.08)	–.10* (.01)	–.20*** (–.14**)	.01 (–.05)
Externalising Problems	**–.65***** (–.31***)	**–.53***** (–.29***)	**–.50***** (–.18***)	**–.50***** (–.27***)	–.23*** (–.11*)
Conduct Problems	**–.43***** (–.20***)	–.31*** (–.17***)	–.35*** (–.11*)	–.36*** (–.17***)	–.19*** (–.09)
Hyperactivity-Inattention	**–.64***** (–.34***)	**–.54***** (–.32***)	**–.48***** (–.20***)	**–.48***** (–.29***)	–.21*** (–.10*)
Total Difficulties	**–.57***** (–.26***)	**–.50***** (–.26***)	**–.40***** (–.13**)	**–.47***** (–.25***)	–.16*** (–.08)
Grade point average	.28***	.29***	.09	.30***	.12*
Subjective well-being	.18**	.24***	.04	.18**	.03

*Note. EC — Effortful Control, AF — Attention Focusing, IC — Inhibitory Control, AC — Activation Control, LIP — Low-Intensity Pleasure. [a]Parent-ratings, teacher-ratings (in parenthesis). *p < .05; **p < .01; ***p < .001.*

In Model 1, the Effortful Control construct was measured by Attention Focusing (AF), Inhibitory Control (IC), Activation Control (AC), and Low-Intensity Pleasure (LIP) scales. The small negative error variance for Activation Control, equal to -.084 (*S.E.* = .092, $p = .359$), was constrained to be zero. The model with one error covariance, Inhibitory Control with Low-Intensity Pleasure, showed excellent fit: $\chi^2 = 1.43$, $df = 2$, $p = .49$, CFI = 1.00, RMSEA = .000. In Model 2, Effortful Control was measured by four latent constructs, AF, IC, AC, and LIP, depicted in Figure 1.

The measurement model for problem behaviour included two latent constructs: Internalising Problems were indicated by parent- and teacher-reported emotional symptoms and peer problems; Externalising Problems were indicated by parent- and teacher-reported conduct problems and hyperactivity. The bivariate correlations between parent-reported and teacher-reported Externalising problems, Internalising problems, and Total difficulties were .50, .23 and .41, respectively. We used correlated errors among similar informants to model method variance. The fit indices indicated good model fit: $\chi^2 = 24.85$, $df = 9$, $p = .003$, CFI = .983, RMSEA = .054.

Academic achievement was measured as a latent variable with three indicators (Russian, math, and literature grades for the year). The subjective well-being construct was measured by self-reported items of the SLSS; the model with one error

covariance between two negatively worded items also showed good fit: $\chi^2 = 29.65$, $df = 13$, $p = .005$, CFI = .971, RMSEA = .046.

The structural models included paths from effortful control to internalising problems, externalising problems, subjective well-being, and academic achievement. Model 1 failed to yield a good fit ($\chi^2 = 545.88$, $df = 189$, $p < .001$, CFI = .847, RMSEA = .056). Therefore, correlated error was allowed between outcome latent constructs, among their indicators and between Inhibitory Control and Externalising Problems. The fit of the final Model 1 (Figure 2) was adequate: $\chi^2 = 474.40$, $df = 187$, $p < .001$, CFI = .908, RMSEA = .050. Higher effortful control was related to better academic achievement and subjective well-being, and fewer externalising and internalising problems, accounting for 11%, 2%, 29%, and 16% of the variance, respectively.

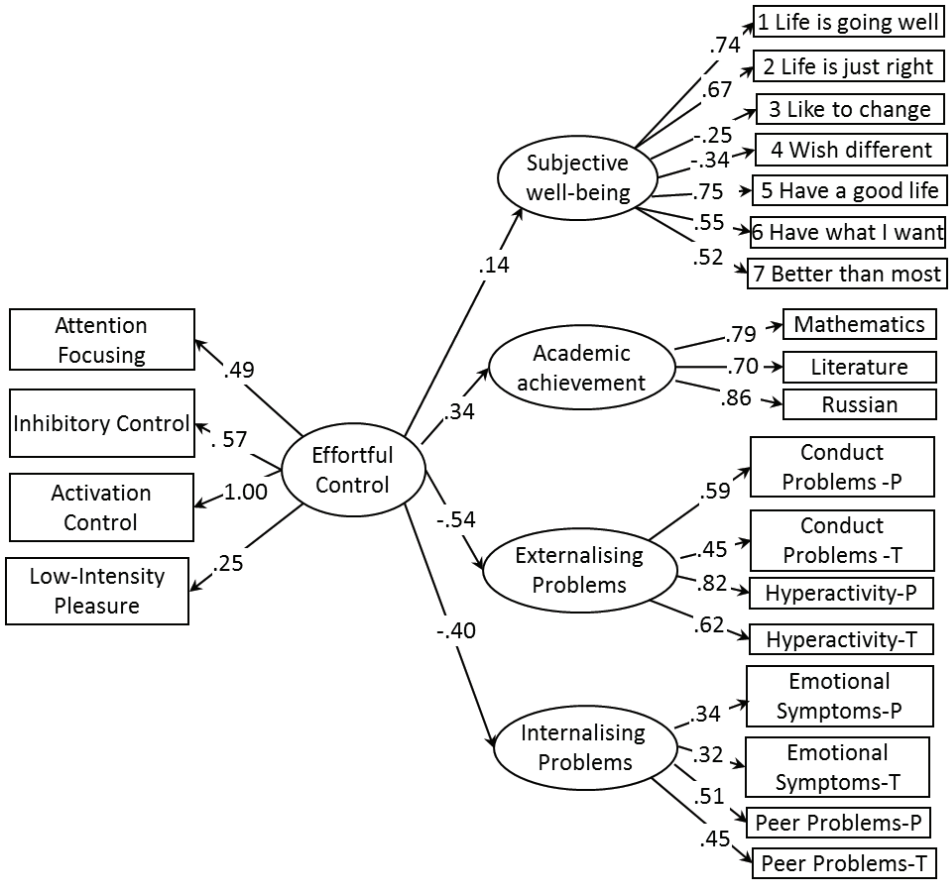


Figure 2. Structural model of effortful control predicting problem behaviour (P = parent reports, T = teacher reports), academic achievement, and subjective well-being (self-reports).

Note. The coefficients are standardised loadings. For clarity, pathways that were not significant at $p < .05$ and correlated errors are not shown.

Model 2, with insignificant paths removed (Figure 3), provided satisfactory fit to the data: $\chi 2 = 757.50$, df = 384, $p < .001$, CFI = .903, RMSEA = .040. Higher attention focusing was related to better subjective well-being and fewer externalising and internalising problems; inhibitory control was negatively related to externalising problems, whereas activation control was positively associated with academic achievement. Together, the model explained 22% of the variance for academic achievement, 8% for subjective well-being, 69% for externalising problems, and 72% for internalising problems.

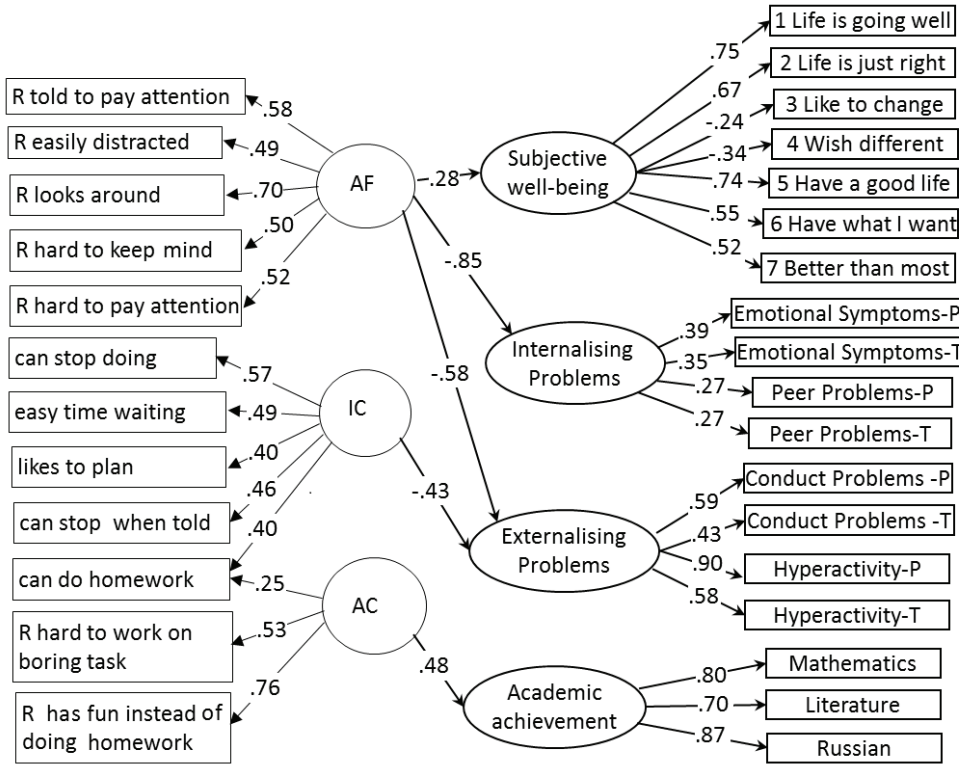


Figure 3. Structural model of Attention Focusing (AF), Inhibitory Control (IC), and Activation Control (AC) predicting problem behaviour (P = parent reports, T = teacher reports), academic achievement, and subjective well-being (self-reports).

Note. Values are standardised estimates ($p < .001$). For clarity, covariances between AF, IC, and AC are not shown.

Discussion

The findings of this study supported the four-factor structure of the abbreviated TMCQ Effortful Control scale from the Temperament in Middle Childhood Questionnaire (TMCQ, Simonds & Rothbart, 2004), including Attention Focusing, Inhibitory Control, Activation Control, and Low-Intensity Pleasure. The Russian version of the abbreviated TMCQ Effortful Control scale showed good convergent, discriminant, and concurrent validity, and adequate internal consistency and test–

retest reliability. Thus, this instrument has satisfactory psychometric properties and may be used as a brief measure of child effortful control in large-scale studies, allowing cross-cultural comparisons and greater generalization of findings. Higher parent ratings for girls on effortful control and its components are largely consistent with the existing research on gender differences in temperament (Else-Quest et al., 2006). However, in contrast to Else-Quest et al.'s (2006) meta-analysis and recent TMCQ findings with Spanish children (Ato, Galián, & Fernández-Vilar, 2014), there were no gender differences in inhibitory control in this study. Whether this discrepancy reflects cultural differences in gender differentiation or cultural bias in parent reports, remains to be investigated.

Showing that effortful control and its components were strongly related to personality traits, our results confirmed and extended previous findings (Halverson et al., 2003) to lower-order traits. The strongest links were found for conscientiousness and all traits from this domain, supporting a conceptual link between effortful control and conscientiousness (Eisenberg, Duckworth, Spinrad, & Valiente, 2014) in the Russian cultural context. Relations between effortful control and agreeableness are also in line with previous findings from other samples. Positive association with the openness domain has been found in adolescents (Tackett, Kushner, De Fruyt, & Mervielde, 2013) and may be due to the high covariation between conscientiousness and intellect at earlier ages (Tackett et al., 2012). Negative association between effortful control and negative affect have been repeatedly supported by other studies with Western and non-Western samples. The evidence on the relationship between effortful control and positive emotions is controversial (Eisenberg, Eggum, Vaughan, & Edwards, 2010); however, positive correlation found in the present study, along with the link with subjective well-being, is consistent with the role of effortful control in subsequent life satisfaction (Moffitt et al., 2013). To our knowledge, this is the first study to show a positive association between effortful control and subjective well-being in childhood.

This study is also the first to examine simultaneously the effect of effortful control on three domains of child well-being — mental health, subjective well-being, and academic achievement — in structural equation models. The findings confirmed that higher effortful control was associated with lower levels of internalising and externalising problems (Eisenberg et al., 2010). In line with previous research, the links with externalising problems were somewhat stronger than with internalising problems (Caspi, Henry, McGee, Moffitt, & Silva, 1995; Zhou, Lengua, & Wang, 2009). This consistency in the pattern of results obtained in different parts of the world (North America, New Zealand, China, and Russia) suggests that the relationship between effortful control and two major groupings of problem behaviour — externalising and internalising — may be universal. Future studies using neuroimaging techniques should investigate how individual differences in the activity of brain regions involved in emotion and behaviour regulation are related to individual differences in questionnaire measure of effortful control that emerged from factor analytic studies of voluntary regulatory behaviours.

Our findings also confirmed the important role of effortful control in academic success, suggesting that activation control makes a major contribution. In our study, inhibitory control was not related to academic achievement in primary schoolchil-

dren, consistent with the findings on math achievement of Spanish primary schoolchildren (Sánchez-Pérez et al., 2015). However, another study found that inhibitory control significantly predicted math achievement in Canadian 9–11-year-olds (Oberle & Schonert-Reichl, 2013). More research is needed to clarify the role of each specific component of effortful control in child well-being and to understand the neural mechanisms involved. Our findings suggest that attentional control is important for problem behaviour and subjective well-being, whereas inhibitory control only contributed to externalising problems, and activation control only contributed to academic achievement.

Conclusion

This study showed that effortful control and its components were strongly related to personality traits and extended previous findings to lower-order traits. Our findings also confirmed the important role of effortful control in academic success and well-being of Russian primary schoolchildren. The findings in this study provided a strong basis for further investigation of the relations between effortful control on the one hand, and personality and well-being on the other, in primary schoolchildren from different cultures. It is also worth mentioning that interventions targeting effortful control may have major implications for child well-being by reducing problem behaviours and promoting academic achievement and life satisfaction. Overall, this study suggests that community-based programs developed in the West may improve child developmental outcomes in other cultural contexts.

Limitations

The present study has several limitations that should be addressed in future research. Firstly, the cross-sectional design did not allow consideration of causal influences; the findings should be supported by longitudinal data. Secondly, although reasonably large and diverse, the sample was not representative; to increase the validity of the findings, more elaborated sampling methods are needed. Thirdly, although three-informant data provided support for the links of effortful control with problem behaviour, subjective well-being, and academic achievement, the relationships of effortful control with personality traits might be partly due to shared method variance; the study would benefit from the inclusion of experimental and neurobiological measures.

Acknowledgements

The work on this manuscript was supported by the Russian Scientific Foundation # 16-18-00003 (data collection and analyses) and the Russian Foundation for Basic Research # 16-06-00022 (preparation of this manuscript). We thank the colleagues, participating parents, teachers, and the administrative authorities of the schools.

References

Arbuckle, J.L. (2008). Amos 17.0 User's Guide. Chicago, IL: SPSS Inc.

Ato, E., Galián, M.D., & Fernández-Vilar, M.A. (2014). Gender as predictor of social rejection: The mediating/moderating role of effortful control and parenting. *Anales de Psicología, 30*, 1069–1078. http://dx.doi.org/10.6018/analesps.30.3.193171

Caspi, A., & Shiner, R.L. (2006). Personality development. In W. Damon, & R. Lerner (Series Eds.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology, Vol. 3. Social, emotional, and personality development* (6th edition, pp. 300–365). New York: Wiley.

Caspi, A., Henry, B., McGee, R.O., Moffitt, T.E., & Silva, P.A. (1995). Temperamental origins of child and adolescent behavior problems: From age three to age fifteen. *Child Development, 66,* 55–68. http://dx.doi.org/10.1111/j.1467-8624.1995.tb00855.x

Digman, J.M., & Shmelyov, A.G. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology, 71*, 341. http://dx.doi.org/10.1037/0022-3514.71.2.341

Eisenberg, N., Duckworth, A.L., Spinrad, T.L., & Valiente, C. (2014). Conscientiousness: Origins in childhood? *Developmental Psychology, 50*, 1331–1349. http://dx.doi.org/10.1037/a0030977

Eisenberg, N., Eggum, N., Vaughan, J., & Edwards, A. (2010). Relations of self-regulatory/control capacities to maladjustment, social competence, and emotionality. In R. Hoyle (Ed.). *Handbook of personality and self-regulation* (pp. 21–46). Oxford, England: Wiley-Blackwell. https://doi.org/10.1002/9781444318111.ch2

Eisenberg, N., Valiente, C., & Eggum, N.D. (2010). Self-regulation and school readiness. *Early Education and Development, 21*(5), 681–698. http://dx.doi.org/10.1080/10409289.2010.497451

Else-Quest, N.M., Hyde, J.S., Goldsmith, H.H., & Van Hulle, C.A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin, 132,* 33–72. http://dx.doi.org/10.1037/0033-2909.132.1.33

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 1337–1345. http://dx.doi.org/10.1097/00004583-200111000-00015

Goodman, R., & Scott, S. (2012). *Child and adolescent psychiatry*. 3nd edition. Wiley-Blackwell. https://doi.org/10.1002/9781118340899

Goodman, R., Slobodskaya, H.R., & Knyazev, G.G. (2005). Russian child mental health: A cross-sectional study of prevalence and risk factors. *European Child and Adolescent Psychiatry, 14,* 28–33. http://dx.doi.org/10.1007/s00787-005-0420-8

Halverson, C.F., Havill, V.L., Deal, J., Baker, S.R., Victor, J.B., Pavlopoulos, V., … Wen, L. (2003). Personality structure as derived from parental ratings of free descriptions of children: The Inventory of Child Individual Differences. *Journal of Personality, 71,* 995–1026. http://dx.doi.org/10.1111/1467-6494.7106005

Huebner, E.S. (1991). Initial development of the Student's Life Satisfaction Scale. *School Psychology International, 12,* 231–240. http://dx.doi.org/10.1177/0143034391123010

Kohnstamm, G. A., Halverson, C. F., Jr., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality: developmental antecedents of the Big Five*? Mahwah, NJ: Lawrence Erlbaum Associates.

Kotelnikova, Y., Olino, T.M., Klein, D.N., Mackrell, S.V.M., & Hayden, E.P. (2016). Higher- and lower-order factor analyses of the Temperament in Middle Childhood Questionnaire. *Assessment, 24(8), 1–12.* http://dx.doi.org/10.1177/1073191116639376

Leto, I.V., Petrenko, E.N., & Slobodskaya, H.R. (2018). Life satisfaction in Russian primary schoolchildren: links with personality and family environment. Journal of Happiness Studies. https://doi.org/10.1007/s10902-018-0036-6

Mary Rothbart's Temperament Questionnaires. Retrieved September 22, 2018. Available from https://research.bowdoin.edu/rothbart-temperament-questionnaires/

McClelland, M.M., Acock, A.C., & Morrison, F.J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly, 21*, 471–490. http://dx.doi.org/10.1016/j.ecresq.2006.09.003

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*, 2693–2698. http://dx.doi.org/10.1073/pnas.1010076108

Moffitt, T.E., Poulton, R., & Caspi, A. (2013). Lifelong impact of early self-control. *American Scientist, 101*(5), 352–359. http://dx.doi.org/10.1511/2013.104.352

Oberle, E., & Schonert-Reichl, K.A. (2013). Relations among peer acceptance, inhibitory control, and math achievement in early adolescence. *Journal of Applied Developmental Psychology, 34,* 45–51. http://dx.doi.org/10.1016/j.appdev.2012.09.003

O'Hare, W.P., & Gutierrez, F. (2012). The use of domains in constructing a comprehensive composite index of child well-being. *Child Indicators Research, 5*, 609–629. https://doi.org/10.1007/s12187-012-9138-6

Proctor, C. L., Linley, P. A., & Maltby, J. (2009). Youth life satisfaction: A review of the literature. *Journal of Happiness Studies, 10*, 583–630. http://dx.doi.org/10.1007/s10902-008-9110-9

Putnam, S.P., Ellis, L.K., & Rothbart, M.K. (2001). The structure of temperament from infancy through adolescence. In A. Eliasz & A. Angleitner (Eds.), *Advances/proceedings in research on temperament* (pp. 165–182). Lengerich, Germany: Pabst Scientist.

Rees, G., Goswami, H., Pople, L., Bradshaw, J., Keung, A., & Main, G. (2012). *The good childhood report 2012: A review of our children's well-being*. London: The Children's Society.

Rothbart, M.K. (2007). Temperament, development, and personality. *Current Directions in Psychological Science, 16,* 207–212. http://dx.doi.org/10.1111/j.1467-8721.2007.00505.x

Rothbart, M.K., Ahadi, S.A., & Evans, D.E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 55,* 958–966. http://dx.doi.org/10.1037/0022-3514.78.1.122

Rothbart, M.K., Sheese, B.E., Rueda, M.R., & Posner, M.I. (2011). Developing mechanisms of self-regulation in early life. *Emotion Review*, *3,* 207–213. http://dx.doi.org/10.1177/1754073910387943

Sánchez-Pérez, N., Fuentes, L.J., Pina, V., López-López, J.A., & González-Salinas, C. (2015). How do different components of Effortful Control contribute to children's mathematics achievement? *Frontiers in Psychology: Educational Psychology, 6, 1383.* http://dx.doi.org/10.3389/fpsyg.2015.01383

Simonds, J., Kieras, J., Rueda, M., & Rothbart, M. (2007). Effortful control, executive attention, and emotional regulation in 7–10-year-old children. *Cognitive Development*, *22*, 474–488. http://dx.doi.org/10.1016/j.cogdev.2007.08.009

Simonds, J., & Rothbart, M.K. (2004). The Temperament in Middle Childhood Questionnaire (TMCQ): A computerized self-report measure of temperament for ages 7–10. Poster presented at the Occasional Temperament Conference, Athens, GA.

Shiner, R.L., & DeYoung, C.G. (2013). The structure of temperament and personality traits: A developmental perspective. In P. Zelazo (Ed.), *Oxford handbook of developmental psychology* (pp. 113–141). New York: Oxford University Press.

Slobodskaya, H.R., & Zupančič, M. (2010). Development and validation of the Inventory of Child Individual Differences–Short Version in two Slavic countries. *Studia Psychologica, 52,* 23–39.

Tackett, J.L., Slobodskaya, H.R., Mar, R.A., Deal, J., Halverson, C.F., Baker, S.R., ... Besevegis, E. (2012). The hierarchical structure of childhood personality in five countries: Continuity from early childhood to early adolescence. *Journal of Personality, 80*, 847–879. https://doi.org/10.1111/j.1467-6494.2011.00748.x

Tackett, J.L., Kushner, S.C., De Fruyt, F., & Mervielde, I. (2013). Delineating personality traits in childhood and adolescence: associations across measures, temperament, and behavioral problems. *Assessment, 20*, 738–751. http://dx.doi.org/10.1177/1073191113509686

Zhou, Q., Lengua, L.J., & Wang, Y. (2009). The relations of temperament reactivity and effortful control to children's adjustment problems in China and the United States. *Developmental Psychology, 45*, 724–739. http://dx.doi.org/10.1037/a0013776

Original manuscript received January 12, 2018
Revised manuscript accepted September 13, 2018
First published online December 20, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

Academic Motivation of Elementary School Children in Two Educational Approaches — Innovative and Traditional

Tamara O. Gordeeva[a,c*], Oleg A. Sychev[b], Diana V. Pshenichnuk[a], Anastasia N. Sidneva[a]

[a] *Lomonosov Moscow State University, Moscow, Russia*

[b] *Shukshin Altai State Humanities Pedagogical University, Biysk, Russia*

[c] *National Research University "Higher School of Economics", Moscow, Russia*

* Corresponding author. E-mail: tamgordeeva@gmail.com

The school must teach how to think.
—Ewald Il'enkov

Background. While the current literature provides valuable insight into how school climate perceptions and student motivation impact academic achievement, research examining the mediating effects of motivation in the linking of an innovative educational system, school climate, and achievement is limited. This study considers the potential of the El'konin–Davydov system of developmental education as a basis for educational innovation. Self-determination theory is applied as a useful theoretical framework that allows for consideration of both the intensity and the quality of academic motivation.

Objective. The study examines a model that illustrates the role of intrinsic and different types of extrinsic motivation in linking the El'konin–Davydov system of developmental education (DE) and school climate to the academic achievement of elementary schoolchildren.

Design. Participants were 345 third and fourth graders drawn from four public schools in Moscow, with some (N = 192, 2 schools) educated in the traditional system and others (N = 153, 2 schools) in one that follows the DE system. A cross-sectional design was implemented.

Results. Students in the DE system showed significantly lower external motivation for all three subscales (Parents, Teachers, General) and perceived school climate more favorably. Structural equation modeling showed that the hypothesized model fit the data well, supporting the hypothesis that student external motivation plays a mediating role in linking educational system (innovative vs. traditional) with academic achievement. Students' autonomous motivation was shown to play a mediating role in linking positive perceptions of school climate with academic achievement.

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

doi: 10.11621/pir.2018.0402
http://psychologyinrussia.com

Conclusion. The elementary school students from developmental education classes compared to their peers from traditional education classes demonstrate more positive profile of academic motivation including lower external motivation, more positive attitude towards school and study; however, the two groups do not differ in the level of intrinsic, identified, and introjected motivations.

Keywords: Intrinsic motivation, extrinsic motivation, self-determination theory, El'konin–Davydov developmental educational system, school climate, academic achievement, elementary schoolchildren

Introduction

A body of research has been dedicated to identifying the factors, both environmental and personal, that influence student academic achievement. Two significant lines of research in educational psychology are the study of (a) characteristics of school climate and educational system, and (b) motivational processes, such as intrinsic motivation, persistence, and academic self-efficacy. Studies have found that both school climate and student motivation are consistently associated with academic achievement (Gustafsson & Nilsen, 2016; Hardre & Reeve, 2003; Roeser & Eccles, 1998; Scherer & Nilsen, 2016). However, while current research provides valuable insight into how school climate and innovative educational systems impact academic achievement, less is known about the mechanisms that link innovative educational systems based on psychological theories with academic achievement, and in particular the mediating role of motivation in this process. The present study applies self-determination theory (SDT) to investigate the role of different types of motivation (e.g., intrinsic and identified versus external) in the link between learning environment (educational system and school climate) and academic achievement.

A Differentiated Approach to Learning Motivation: Self-Determination Theory (SDT)

According to SDT, motivation refers to the reasons that underlie behavior; motivation differs not only in quantity, but also in quality (types of motivation) (Ryan & Deci, 2000b). Applied to education, this refers to the reasons that students engage in learning activity (Ryan & Deci, 2000a). SDT distinguishes among intrinsic motivation and different types of extrinsic motivation, which vary in terms of their self-determination—i.e., the extent to which behavior originates from the self. Intrinsic motivation refers to engaging in an activity for its own sake, for the interest or enjoyment derived from mastering optimal challenges. By contrast, extrinsic motivation refers to engaging in an activity for instrumental reasons such as receiving rewards, avoiding punishments, boosting one's self-worth, or reaching personally valued goals.

According to SDT, there are four subtypes of extrinsic motivation that vary in the degree of self-determination; these four subtypes can be placed along a continuum of relative autonomy — external, introjected, identified, and integrated (Ryan & Connell, 1989).

External motivation is completely noninternalized (see Fig. 1 from Sheldon & Gordeeva, 2018), the least autonomous form of motivation. It refers to doing an

activity because one feels controlled by external contingencies involving threats of punishments, demands and criticism or, alternatively, approval and material rewards bestowed by others. Next in terms of degree of relative autonomy, *introjected motivation* refers to the regulation of behavior by internal forces, such as feelings of obligation and contingent self-esteem. The individual somewhat endorses the reasons for learning, but still in a controlled way. Within the introjected subtype of extrinsic motivation, positive introjected and negative introjected motivations can be reliably distinguished: the former is related to feelings of self-worth and pride, and the latter to feelings of shame, guilt, and obligation (Assor, Vansteenkiste, & Kaplan, 2009; Sheldon, Osin, Gordeeva, Suchkov, & Sychev, 2017). The next motivational type, *identified motivation,* is a self-determined form of extrinsic motivation, because the behavior originates from the self in a non-contingent way. It is observed when students identify with the reasons for performing a behavior, or when they personally value it and find it important. Finally, *integrated motivation* refers to the most internalized type of extrinsic motivation. This type requires the individual to have formed an identity such that s/he can identify with the importance of a behavior and assimilate that identification with other aspects of a coherent sense of self. Since the hierarchical organization of the self is developed only by the end of adolescence (Harter, 2003), we did not assess integrated motivation in the present study.

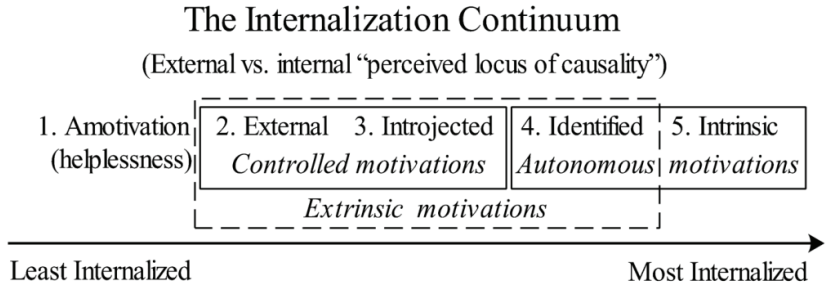


Figure 1. The Internalization Continuum (from Sheldon & Gordeeva, 2018)

In the educational domain, these types of motivation have been found to lead to a number of important cognitive, behavioral, and emotional outcomes. Students who express more autonomous types of motivation (intrinsic and identified types of motivation) are more persistent and cognitively involved in their tasks, display more positive coping styles and higher quality learning, whereas students who are motivated in a more controlled fashion (introjected and external types of motivation) expend less effort, are more easily distracted, experience more anxiety and other negative emotions, and have lower grades (Guay, Ratelle, & Chanal, 2008; Ryan & Connell, 1989; Ryan & Deci, 2017; Gordeeva, 2014).

Developmental Education and School Climate as External Factors of Academic Attainment

Both educational system and school climate are considered to have a significant impact on students' achievement and persistence. Definitions of school climate usually include students' perceptions about school, relationships with teachers and

classmates, and values and expectations that help feeling emotionally, socially, and physically safe (Cohen et al., 2009). Educational system is defined as a set of pedagogical principles and strategies that relate to the goals, content, methods, forms of education (teaching), as well as the grading system; its main goal is to improve instructional quality.

Developmental Education (DE). Based on L.S. Vygotsky's idea about "education that leads development" (Vygotsky, 1956/2017) and Davydov's theory of developmental education (Davydov, 1996), this practical approach to education was proposed by Russian psychologists V.V. Davydov, D.B. El'konin and V.V. Repkin in the late 1950s and was further developed by G.A. Zuckerman (2011, 2014). DE has received official state approval and is currently used in some Russian schools (Vorontsov, Zaslavsky, L'vovsky, Chudinova, & El'konin, 2013). Compared to the traditional educational system, DE involves changes in the goals and content of learning, as well as innovative methods, different forms of learning, and a different grading system.

The main goal of the Elkonin–Davydov "system of developmental education", as it is called, is to encourage students' learning initiative, independent critical thinking, and "the development of reflective abilities … and learning potential that requires transcending the limits of one's knowledge and skills in search of new ways of acting" (Zuckerman, 2014, p. 198). In particular, this system implies a change in the content of education, which means the inclusion in such content of theoretical knowledge and the corresponding general methods of action. This change presumably leads to the formation of theoretical thinking. Theoretical knowledge is knowledge about the genesis of the concept being learned (Davydov, 2008). Theoretical knowledge allows children to understand not only how to do something, but also why to perform a task this way and not another.

DE criticizes traditional educational system for developing in children the ability to act deliberately in accordance with given rules and instructions, while in the DE system the teacher deliberately does not provide ready-made knowledge for the students, but makes an effort to help them search independently for new means and methods of action (Zuckerman, 2014). Within this educational system, continuing the ideas of J. Bruner's "discovery learning" (Bruner, 1960), different problem-based methods of teaching are used. In particular, various means of educating children about how to work with a contradiction found within a problem are utilized, while entertaining different points of view about the same phenomenon. As to the forms of teaching, DE emphasizes the students' collaboration toward the formation of learning cooperation. Zuckerman (2014) considers the "molecule" of developmental education to be the mutual (i.e., interactive) actions of the child (or group of children) and the adult, which intersect at the point of the socially designed tool, sign, or symbol that is to be mastered.

According to Zuckerman, the general principles of cultivating learning skills in DE are as follows. First, new concepts are introduced when schoolchildren come across a contradiction between their knowledge and a new fact. Second, the introduction and concretization of each concept take place in the form of a discussion specially "sharpened" by the teacher, so that different points of view existing in the classroom are presented in a contrasting way. Third, judgments made by a student, teacher, or textbook author are considered to be an hypothesis until evidence for it

is provided, and the methods for that are specifically taught (Zuckerman, 2005; see examples in Zuckerman, 2014).

Finally, the DE system differs from the traditional one in that the traditional (teacher-guided, public, and comparative) grading system does not apply. From the very beginning of the learning activity, children are taught to evaluate their own achievements based on a system of different criteria jointly discussed and selected with the teacher. Additionally, as the DE system suggests, if the child is working at the limit of his or her abilities, the teacher should proceed from the fact that the pupil deserves the highest evaluation, regardless of what the achievements of other children were.

A number of studies by Davydov, Zuckerman and their colleagues showed that DE facilitates better theoretical thinking, problem-solving skills, the ability to learn (Davydov, 2008; Zuckerman, 2005), and a higher level of readiness for school education (Nisskaya, 2018). However, the results were based on rather small samples and questionable measures. For example, Zuckerman (2005) has shown that high school students from DE classes compared to their peers in modern traditional classes demonstrate higher ability to learn, which means the ability to acquire knowledge independently and use intellectual tips. However, the sample of participants from DE classes was small and there were no differences in intellectual achievements measured by math PISA tests. Even more importantly, less is known about motivational and self-regulatory factors that stimulate the positive intellectual outcomes of DE students, and in particular the level of intrinsic and different types of extrinsic motivation that may support them. It is also unclear whether these results apply already to elementary school students studied in DE classes.

In Voronkova's study (2003) it was shown that 8th graders from DE classes had significantly higher learning motivation than their peers from high school lyceum classes with a pre-selection but standard educational program. Based both on this study and the theoretical principles of the theory of developmental education (see Selevko, 2005), we assumed that the level of intrinsic motivation of children from DE classes would be higher than that of children from classes with traditional education.

School climate. In the present study, we consider school climate as a relevant dimension of the learning environment which, together with the educational system, may shape student motivation related to academic achievement. Empirical evidence suggests that children's perception of a positive school climate is linked to better academic performance and learning outcomes (Jia et al., 2009; Niehaus, Rudasill, & Rakes, 2012; Reyes et al., 2012); to well-being, including greater life satisfaction and less likelihood of depression; and to fewer behavioral and emotional problems (Scrimin, Moscardino, Altoé, & Mason 2017). Differences in school climate have helped to explain why, despite large class sizes, students in China demonstrate higher academic achievement and less disruptive and aggressive behaviors than those at many schools in the United States (Yang et al., 2013).

Relationships with teachers are often considered to be the most salient and important components of school climate. In classrooms with a positive climate, teachers are responsive and caring in relation to their students' academic and emotional needs; they create a warm and harmonious atmosphere, respect and highly value

the students' perspectives. Relationships with teachers, when measured as positive teacher regard and perceptions of teacher support, have been found to correlate positively with student grades (Roeser & Eccles, 1998). Similarly, students' perceptions of a caring and supportive relationship (Battistich, Solomon, Kim, Watson, & Schaps, 1995) were found to positively relate to math, science, and writing performance for elementary and middle school students.

A recent large-sample study of high school students showed that intrinsic motivation mediates the link between perceived school climate (in particular, the teacher–student relationship) and achievement in reading and math (Fan & Williams, 2018). We hoped to replicate these findings using a sample of elementary school children, utilizing the SDT perspective on motivation. To the best of our knowledge, no attention has been given to the link between an innovative educational system and controlled types of motivation, and the role of these types of motivation in the link between the educational system and academic achievement. Furthermore, despite the fact that the system of developmental education has been found to be effective (Zuckerman, 2005), there is a lack of studies that show the psychological mechanisms that mediate its impact on students' educational attainments.

Three research questions guided the present study:

1. Do elementary school students in DE classes show higher intrinsic and identified (i.e., autonomous) motivations than students in traditional classes?
2. Do elementary school students in DE classes show lower introjected and external (i.e., controlled) motivations than students in traditional classes, and a higher relative autonomy index (RAI), which testifies to the dominance of autonomous over controlled forms of motivation?
3. Are autonomous and controlled motivations associated with the type of educational system, perceived school climate, and academic achievement of elementary school students?

Our hypotheses were the following:

1. On the basis of the literature suggesting that (a) DE increases problem-solving skills and academic achievement, and (b) DE by itself is an activity aimed to develop the child's potential and this goal should be the subject of reflection by the child (Zuckerman, 2005), we anticipated that students in DE would report higher autonomous and lower controlled motivations.
2. Building on our own research as well as the work by Fan and Williams (2018), we expected that the developmental education system would have an impact on student motivation, which in turn would be associated with higher grades.
3. Given that motivation has been found to be linked both to the school's emotional climate and to children's adjustment to school, sometimes mediating both variables (e.g., Reyes et al., 2012), we hypothesized that perceived school climate would be associated with autonomous motivation, which in turn would be associated with academic achievement.

Method

Participants and Procedure

Participants were 345 third and fourth graders drawn from four public schools in Moscow, with some (N = 192) educated in the traditional system and others (N = 153) in an innovative one (DE). The total sample comprised 186 boys and 158 girls (one child did not specify gender); age $M = 9.7$, $SD = 0.69$, age range 8–12 years. The questionnaires were administered to students in group settings during regular class hours. The research was introduced as "a study of children's views on life and study". Parental consent to participate was obtained for all students. At the end of the school year, we obtained information on the academic achievement of 259 pupils (final marks for four main subjects) from the school administrations.

Measures

Academic motivation. For the purposes of this study, we created Academic Motivation Questionnaire (AMQ) based on the Ryan and Connell (1989) SRQ measure. This questionnaire addresses the reasons that children do their schoolwork and homework, and try to answer questions in class; it measures intrinsic, identified, introjected, and external regulations (motivations) for these various school-related activities. Each subscale had 4 items. An expert committee of two professors and two doctoral students created a set of items, which included new items and several translated items previously used in SRQ. Partly following Vallerand's ideas (Vallerand et al., 1993) and our own model of intrinsic motivation (Gordeeva, 2014), we developed two intrinsic motivation subscales, one to measure motivation to learn (a sample item is, "I do my homework because I like knowing new things") and another for self-development motivation (a sample item is, "Because I like to improve myself"). Following previous research (Sheldon et al., 2017), we developed two introjected subscales, positive and negative. Sample items are, "Because I'll be proud of myself if I do it" (positive introjection) and "Because I'll feel ashamed of myself if I don't" (negative introjection). Finally, to capture more precisely the variety of external pressures that induce children to learn, in addition to a general external motivation scale, scales were created to assess students' perceptions of external pressures coming from parents and teachers. Example items are, "I have no choice, I have to work in class" (General), "Because my parents demand that I study well" (Parents), "Because my teacher will be glad if I study well" (Teachers). CFA for the questionnaire is $\chi^2 = 997.62$; df = 436; $p < 0.001$; CFI = 0.841; NNFI = 0.820; RMSEA = 0.061; 90% CI for RMSEA: 0.056–0.066; PCLOSE = 0.000; N = 345. Reliability coefficients (Cronbach's α) for all scales are presented in table 1.

Educational system. The educational system (innovative vs. traditional) was coded as 1 and 0, respectively, for classes that used the El'konin–Davydov (DE) system (2 schools, 8 classes, N = 153) and the traditional educational system (2 schools, 9 classes, N = 192). The specificity of the DE educational system compared to the traditional one is described above.

School climate. We assessed school climate via the Perception of School subscale from the Multidimensional Students' Life Satisfaction Scale (MSLSS) (Huebner, 1994, Russian adaptation —Sychev, Gordeeva, Osin, & Sidneva, 2019) and by creating a scale to measure relationships with teachers. Both Perception of School

and Relationships with Teachers subscales had six items, with items rated on 5-point Likert scales. Example items are "I do not really like my school" (School), "I like to listen to my teachers" (Teachers). Cronbach's alphas were 0.87 and 0.89, respectively.

A nonverbal measure of general attitude towards school and study was also employed (Andrews & Withey, 1976). It depicts seven faces, ranging from most happy to least happy. The child was asked to decide which face best reflects his/her attitude towards school and study.

Academic achievement. The end of the year's grades in four principal school subjects (math, native language [Russian], foreign language [English], and reading/literature) were collected from school records. We were able to compare academic achievement in two types of schools, because even a school that does not give the children grades in the traditional sense (DE) is obliged to submit grades to the school administration at the end of the year. The four grades were averaged to form an overall academic achievement indicator (pairwise correlations between four subjects ranged from 0.51 to 0.77). All schools except one used traditional scale of marks ranging from 2 to 5 and in one school 100-point scale was used. We transformed 100-point marks into traditional scale using the rules recommended by the school administration before statistical analysis (the system they actually use presenting students' academic achievement at the end of the school year).

Results

First, we conducted a correlation analysis of the scales of the extended version of the AMQ. Correlations of scales, the mean values, and standard deviations are presented in Table 1. As we expected, the two intrinsic types of motivation and identified motivation were highly correlated ($r > 0.7$), which justifies combining them to represent the construct of autonomous motivation. The three subscales of external motivation (Teachers, Parents, and General) were also strongly correlated, although the values of coefficients were somewhat lower ($0.63 > r > 0.44$). In accordance with SDT, there was no positive association between intrinsic and external (as least autonomous form of extrinsic) motivation, while introjected motivations, which is middle of the self-determination continuum, showed positive correlations with both intrinsic and external types of motivation. On the whole, the correlation analysis demonstrated that in spite of the large number of scales, their correlations support well the simplex pattern postulated in SDT (Ryan & Connell, 1989). Accordingly, an unweighted relative autonomy index (RAI) was calculated following Sheldon et al. (2017), showing the relative dominance of autonomous motivation over controlled motivation. In particular, RAI index was calculated as sum of motivation to learn, self-development, identified and positive introjected motivation minus sum of negative introjected and three external types of motivation.

Given that the sample comprised students from four schools we started by comparing children from schools that used the same educational system by all the measured variables. Using Student's *t*-test we discovered that students studying in different schools but using the DE innovative system were similar to each other (no statistically significant differences were found). Relationships with teachers (as a school climate indicator) was the only scale to show weak, but significant differ-

Table 1

Means, standard deviations and correlations by scales of academic motivation

	Number of items	Mean	SD	1	2	3	4	5	6	7	8
1. Intrinsic M.	4	2.81	0.85	(0.77)							
2. Self-development M.	4	3.22	0.76	0.77***	(0.81)						
3. Identified M.	4	3.19	0.77	0.75***	0.79***	(0.78)					
4. Positive introjected M.	4	2.84	0.85	0.36***	0.40***	0.42***	(0.70)				
5. Negative introjected M.	4	2.89	0.90	0.36***	0.43***	0.43***	0.63***	(0.74)			
6. External M. (Parents)	4	2.89	0.79	0.02	0.10	0.08	0.38***	0.40***	(0.66)		
7. External M. (Teachers)	4	2.85	0.80	-0.12*	-0.03	-0.07	0.29***	0.29***	0.59***	(0.70)	
8. External M. (General)	4	2.78	0.92	-0.12*	-0.04	-0.08	0.16**	0.30***	0.44***	0.63***	(0.65)
9. RAI	32	0.67	3.22	0.68***	0.61***	0.63***	0.17**	-0.07	-0.48***	-0.63***	-0.65***

*Note. * p < .05, ** p < .01, *** p < .001, two-tailed, N = 345. On the diagonal in parentheses are values of Cronbach's α. M — motivation, RAI — relative autonomy index, numbers of variables in column names correspond to the numbers in row names.*

ence at $p < .05$. For students studying in the two schools employing the traditional system we found two significant differences on positive introjected motivation and Perception of school scales. However, these differences were also rather weak and significant at $p < .05$. The small magnitude and low number of differences allowed us to merge the students into two groups, on the basis of the respective educational system, for convenience of further analysis.

T-tests conducted upon the academic motivation scales comparing innovative and traditional educational systems revealed that students in the DE system showed significantly lower external motivation for all three subscales (Parents, Teachers, General) and a higher relative autonomy index (see Table 2). The most significant difference between students educated in the two systems was in parental control. For this scale, the Cohen's *d* effect size equal to 0.60 demonstrated a medium effect of educational system, while the effect size for the two other extrinsic motivation scales (Teachers and General) was small. No differences were found for intrinsic, identified, and introjected motivation as well as for academic achievement.

T-tests conducted upon indicators of school climate comparing the DE and traditional educational systems revealed that students in the innovative system had better attitudes towards school and studies and a slightly more positive perception of school (marginally significant), which was demonstrated both via self-report and nonverbal measures (see Table 2).

Table 2

Differences in motivation, indicators of school climate and academic achievement between students in the system of developmental education (DE, innovative system) and the traditional system

Indicators	Means		Mann–Whitney *U*	Z	*p*-value	Cohen's *d*
	Innovative system (DE) (*N* = 153)	Traditional system (*N* = 192)				
Intrinsic learning M	2.81	2.82	14642.5	–0.05	n.s.	0
Intrinsic self-development M	3.16	3.28	13271	–1.56	n.s.	0.16
Identified M	3.12	3.25	13372.5	–1.44	n.s.	0.17
Introjected Positive M	2.77	2.90	13737.5	–1.04	n.s.	0.15
Introjected Negative M	2.77	2.99	13120	–1.72	n.s.	0.24
External M (Parents)	2.64	3.09	9891.5	–5.24	< 0.001	0.60
External M (Teachers)	2.73	2.94	12695	–2.17	< 0.05	0.26
External M (General)	2.66	2.87	12641.5	–2.26	< 0.05	0.23
Relative Autonomy Index	1.06	0.36	12580.5	2.29	< 0.05	0.22
Perception of school	3.45	3.24	12995.5	1.84	0.06	0.20
Relationships with teachers	3.69	3.83	13788.5	–0.98	n.s.	0.13
Attitude towards school	5.45	5.18	12685	2.23	< 0.05	0.19
Academic achievement	4.21	4.23	6154.5	–0.34	n.s.	0.02

Note. M — motivation.

The difference between motivational profiles of students in two educational systems is illustrated in Figure 2.

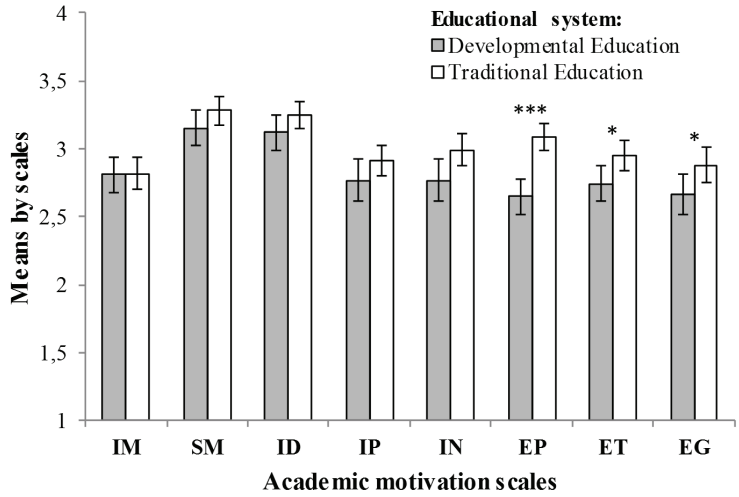


Figure 2. Profiles of academic motivation in the traditional system and the system of developmental education (vertical bars indicate limits of 95% CI, * p<.05, *** p<.001). IM — intrinsic learning motivation, SM — intrinsic self-development motivation, ID — identified motivation, IP — positive introjected motivation, IN — negative introjected motivation, EP — external motivation (Parents), ET — external motivation (Teachers), EG — external motivation (General).

Table 3

Correlations of academic motivation scales with school climate indicators and academic achievement

Indicators	School climate			Academic achievement
	Perception of school	Relationships with teachers	Attitude towards school	
Intrinsic learning M	0.70***	0.60***	0.67***	0.14*
Intrinsic self-development M	0.57***	0.56***	0.58***	0.13*
Identified M	0.57***	0.55***	0.55***	0.19**
Introjected Positive M	0.18***	0.26***	0.21***	–0.01
Introjected Negative M	0.20***	0.30***	0.21***	0.04
External M (Parents)	–0.08	0.15**	–0.05	–0.15*
External M (Teachers)	–0.21***	–0.06	–0.14**	–0.16**
External M (General)	–0.16**	–0.10	–0.11*	–0.13*
Relative Autonomy Index	0.56***	0.41***	0.52***	0.21**
N	345	345	345	259
α	0.87	0.89	—	0.87

*Note. * p < .05, ** p < .01, *** p < .001, two-tailed, α — Cronbach's α.*

Correlations between academic motivation scales, indicators of school climate, and academic achievement are presented in Table 3. At the zero-order level, school climate variables showed strong positive correlations ($r > 0.5$) with intrinsic and identified motivation; moderate positive correlations with both types of introjected motivations; and weak to moderate negative correlations with extrinsic types of motivation. Academic achievement showed positive significant correlations with both types of intrinsic motivation, identified motivation and relative autonomy index. Also weak but significant negative correlations with academic achievement were shown by all subscales of external motivation.

For deeper investigation of the relations among academic achievement, motivation, educational system, and school climate, we conducted structural equation modeling analyses using MPLUS 7.4 in the subsample of students that included data on academic achievement ($N = 259$). This sample included all students (N = 192) educated in the traditional system and 67 (44%) students in a developmental one. We did not succeed to get the student's grades from one DE school participated in the study. However, comparison of this subsample of DE students with the other DE students showed the absence of significant differences among the variables included in the model.

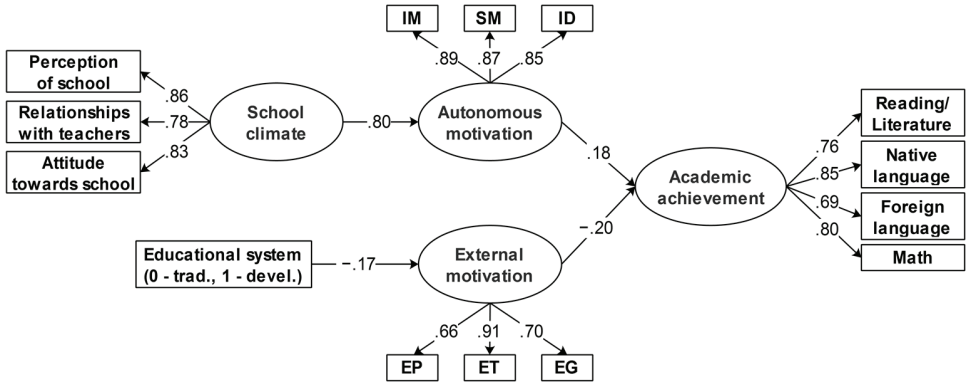


Figure 3. The structural model of relations between educational system, school climate, academic motivation, and academic achievement (all coefficients are standardized and significant at $p < 0.05$, $N = 259$). IM — intrinsic motivation (motivation to learn), SM — self-development intrinsic motivation, ID — identified motivation, EP — external motivation (parents), ET — external motivation (teachers), EG — external motivation (general).

Building the model, we specified two motivational factors of autonomous and external motivation that predicted the latent variable of academic achievement. Introjected motivations were excluded from this analysis to show two extreme types of motivational functioning. External motivation was predicted by the educational system, while autonomous motivation was predicted by school climate. To analyze this model, we used the robust maximum likelihood (MLR) estimator and the full-information maximum likelihood algorithm (Enders & Bandalos, 2001), which estimated parameters on the basis of the available complete data, as well as the implied values of the missing data given the observed data. Figure 2 presents the standardized coefficients that resulted in the model. The path coefficients from

autonomous motivation to academic achievement and from educational system to external motivation were significant at $p < .05$; all the other coefficients were significant at $p < .01$. This model yielded a good fit to the data ($\chi^2 = 137{,}657$; $df = 74$; $p < 0{,}001$; CFI = 0,961; NNFI = 0,952; RMSEA = 0,058; 90% CI for RMSEA: 0,043-0,073; PCLOSE = 0,190; $N = 259$).

The model demonstrates that autonomous motivation is positively associated with academic achievement, while external motivation shows a negative impact on academic achievement, as SDT predicts. The model also shows that educational system and school climate are independent factors, and each has its special motivational consequences. As the model suggests, autonomous motivation shows a strong association with school climate, but is not related to educational system. At the same time, studying in the innovative educational system makes a moderately negative impact on students' external (controlled) motivation (that is, for those in the DE system, external motivation tended to be lower than for those in the traditional system), which in turn is associated with lower academic achievement.

Discussion

The El'konin–Davydov system of developmental education is the most famous system of instruction in Russia. It is grounded in psychological theory, the ideas of scholars such as Vygotsky, Bruner, Davydov, El'konin, Repkin, and Zuckerman. Our results show that DE does have a certain potential for the development of children. In particular, the results indicate that students in DE classes had a more favorable attitude towards school and study and a more positive pattern of motivation, expressed in lower external motivation. Although previous research was focused on showing the role of intrinsic and autonomous motivation in the effective learning process (Garon-Carrier et al., 2016), we were able to show that an educational system might be effective by lowering the level of controlled (i.e. external) motivation in elementary school students. We found that all three types of external motivation were lower in DE classes than in traditional classes. We suggest that this might be due to the methods of collective discussion and mutually active interaction, which downplay the teacher's individual pressure and control imposed on, or at least perceived by, students. The low level of external regulation could also be due to the fact that in DE the teacher aims not to favor following rules and instructions, as well as favor obedience and submission, but gravitates to focusing on supporting the students' initiative and independent thinking (Zuckerman, 2014).

Perhaps even more importantly, the lack of traditional evaluation practices in DE classes may positively affect student motivation. The ways in which students are evaluated is one of the most salient classroom factors that can affect student motivation. Depending on how evaluation is structured and presented, students may experience different patterns of motivation and different learning goals (Ames, 1992). For example, SDT research argues for the negative impact of high-stakes testing (Ryan & Weinstein, 2009). Due to the innovative grading practices used in developmental education, students in DE classes may feel more secure about their own competence. As mentioned earlier, in DE classes the students are not given

traditional grades; instead, they learn to evaluate themselves first, using different criteria, then compare these evaluations with the teacher's assessments, which are usually private, so comparison with others is not emphasized. This pedagogical practice is also accompanied by lack of opportunity for parents to control the children by interrogating them about grades, which is a common practice in modern families. As research shows, Russian parents compared to British and American parents are especially prone to be critical regarding children's academic performance (Elliott et al., 2005).

We also found that DE students demonstrated a more positive attitude toward school and learning, i.e. school climate was more positive in DE classes than in traditional classes. This could be explained from an SDT perspective. Teachers can foster students' well-being and motivation to learn by supporting their psychological needs for autonomy, competence, and relatedness (Ryan & Deci, 2000, 2017). Collective problem-focused discussions practiced both in small groups and with the whole class may support children's competence and autonomy needs by having the teacher listen more, offer encouragement (making fewer "should/must" statements), and provide the children with the opportunity to think and act independently and perceive themselves as creators of knowledge, subjects, not objects of the learning process. The competence need could also be better satisfied in DE classes due to the grading practices considered above. Greater relatedness may be fostered by the team-based projects and collaborative learning in discussion sessions practiced in DE. A recent study (Trenshaw et al., 2016) showed that students' feelings of connectedness could be improved by transforming instructors from graders into mentors, and peers from competitors into valued partners in learning. These relationships motivate students to challenge themselves as it becomes safe to fail and try again.

Why were intrinsic and identified motivations not higher in DE classes than in classes with traditional system of education? It was an unexpected result which basically contradicts the core principles of DE (see Selevko, 2006), as well as the results obtained by Voronkova (2003) on adolescents from DE classes. Perhaps this is due to the age of our sample, which has not yet fully felt the results of developmental education. The main reasons we hypothesized that children in DE would feel greater intrinsic motivation are the innovative problem-based methods of teaching, which try to capture students' interest, and the theoretical content of education in DE. However, despite we did not find any differences in intrinsic motivation between DE and traditional classes, although it is worth mentioning that the relative autonomy index (RAI), which shows the dominance of autonomous motivation over controlled motivation, was higher in DE classes. We suggest that a decrease in the level of external (controlled) motivation may be the first step towards raising the level of intrinsic motivation in the future. The more positive school climate found in DE classes also supports this claim. On the other hand, recent studies in math achievements of elementary school children show that the role of intrinsic motivation might be overestimated (Garon-Carrier, Guay, Dionne et al., 2016) which means that our expectations regarding higher intrinsic motivation in DE classes were not justified.

In the present study, the correlations among school climate and autonomous motivation confirm previous findings (Guay et al., 2008). Even more importantly,

our results about school climate, intrinsic motivation, and academic achievement are in good agreement with the results of the study by Fan and Williams (2017) on high school students, which showed the role of intrinsic motivation as a mediator of school climate's influence on academic achievement. Using a sample of elementary students, we found that autonomous motivation (a general desire to know, to develop one's potential, and to value learning) is associated with both positive relationships with teachers and a positive attitude towards school, as well as with academic achievement. This finding is also in line with humanistic theories of learning (Maslow, 1954; Rogers, 1961) and the SDT perspective, which both assert that acceptance, relatedness, and respect from teachers is critical to students' desire to learn and develop their potential.

Limitations

The limitation of this study is the broad nature of the DE concept, which was not measured in detail. Future researchers might evaluate DE teachers' goals of education, to explore their understanding of the content, methods, forms of education and grading system they actually use while teaching. Also with this type of design we cannot be sure that the distribution of children in TE and DE classes was random, which could also affect out results (although we know from school administrations that no special selection was undertaken in either DE or TE classes). Clearly, to better understand how the DE approach affects student motivations to learn and make educational progress, more research is needed.

Another limitation of the study is its cross-sectional design; it is especially important for the claim that school climate makes an impact on academic motivation. In future studies a longitudinal design is needed, so that climate variables and motivational variables as dependent variables can be measured at least twice to test the causal nature of these relationships.

Conclusion

The results highlight that the children from developmental education classes compared to children from traditional system of education classes demonstrate more positive attitude towards school and study as well as more beneficial motivational profile including lower levels of external academic motivations, which means that they feel less controlled by parents, teachers and grades. No differences in intrinsic and identified motivations as well as the level of academic achievement were found. However, the pedagogical mechanisms that underlie these phenomena are not completely clear, and further research should show what needs to be changed in the system of traditional education in Russia, so that it becomes more consistent with modern psychological theories of learning and learning motivation.

Acknowledgements

The study received the financial support of the Russian Foundation for Basic Research, project No. 18-013-00386. We thank all the children and school administrators who allowed us to conduct this study.

References

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Andrews, F.M., & Withey, S.B. (1976). *Social indicators of well-being: Americans' perceptions of life quality*. New York: Plenum Press. https://doi.org/10.1007/978-1-4684-2253-5

Assor, A., Vansteenkiste, M., & Kaplan, A. (2009). Identified versus introjected approach and introjected avoidance motivations in school and in sports: The limited benefits of self-worth strivings. *Journal of Educational Psychology*, 101, 482-497. https://doi.org/10.1037/a0014236

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal*, *32*(3), 627–658. https://doi.org/10.3102/00028312032003627

Bruner, J. S. (1960). *The process of education*. Cambridge, MA: Harvard University Press.

Cohen, J., McCabe, E.M, Michelli, N.M & Pickeral, T. (2009). School Climate: Research, Policy, Teacher Education and Practice. *Teachers College Record*, *111* (1), 180-213. Retrieved from http://www.tcrecord.org/Content.asp?ContentId = 15220

Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J-P., Séguin, J., Vitaro, F., & Tremblay, R.E. (2016). Intrinsic motivation and achievement in mathematics in elementary school: A longitudinal investigation of their association. *Child Development, 87*, 165–175. https://doi.org/10.1111/cdev.12458

Gordeeva, T.O. (2014). Bazovye tipy motivatsii deyatel'nosti: potrebnostnaya model' [Basic types of motivation of human activity: The need model]. *Vestnik Moskovskogo Universiteta. Seria 14. Psychologiya [Moscow University Psychology Bulletin],* 3, 63–78. Retrieved from http://msupsyj.ru/articles/detail.php?article = 5179

Gordeeva, T.O. (2015). *Psikhologiya motivatsii dostizheniya* [Psychology of Achievement Motivation]. Moscow: Smysl.

Davydov, V.V. (1996). *Teoriya razvivayushchego obucheniya* [The theory of developmental education]. Moscow: Pedagogika.

Davydov, V. (2008). *Problems of developmental instruction: A theoretical and experimental psychological study*. New York: Nova Science Publishers, Inc.

Enders, C.K., & Bandalos, D.L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Fan, W., & Williams, C. (2018). The mediating role of student motivation in the linking of perceived school climate and academic achievement in reading and mathematics. *Frontiers in Education*, *3*, 50. https://doi.org/10.3389/feduc.2018.00050

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 233–240. https://doi.org/10.1037/a0012758

Guiffrida, D., Lynch, M. F., Wall, A., & Abel, D. (2013). Do reasons for attending college affect academic outcomes? A test of a motivational model from a self-determination theory perspective. *Journal of College Student Development*, *54*, 121–139. https://doi.org/10.1353/csd.2013.0019

Gustafsson, J.-E., & Nilsen, T. (2016). The impact of school climate and teacher quality on mathematics achievement: A difference-in-differences approach. In *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2, pp. 81–95). https://doi.org/10.1007/978-3-319-41252-8_4

Hardre, P.L., & Reeve, J. (2003). A motivational model of rural students' intentions to persist in, versus drop out of, high school. *Journal of Educational Psychology*, *95*(2), 347–356. https://doi.org/10.1037/0022-0663.95.2.347

Harter, S. (2003). The development of self-representation during childhood and adolescence. In M. R. Leary, & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 610-642). New York: Guilford.

Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment*, *6*(2), 149–158. https://doi.org/10.1037/1040-3590.6.2.149

Jia, Y., Way, N., Ling, G., Yoshikawa, H., Chen, X., Hughes, D., … Lu, Z. (2009). The influence of student perceptions of school climate on socioemotional and academic adjustment: A comparison of Chinese and American adolescents. *Child Development*, *80*(5), 1514–1530. https://doi.org/10.1111/j.1467-8624.2009.01348.x

Maslow, A. (1964). *Motivation and personality.* New York: Harper.

Niehaus, K., Rudasill, K. M., & Rakes, C. R. (2012). A longitudinal study of school connectedness and academic outcomes across sixth grade. *Journal of School Psychology*, *50*(4), 443–460. https://doi.org/10.1016/j.jsp.2012.03.002

Nisskaya, A.K. (2018). School readiness outcomes of different preschool educational approaches. *Psychology in Russia: State of the Art, 11*(1), 43-60. https://doi.org/10.11621/pir.2018.0104

Reyes, M.R., Brackett, M.A., Rivers, S.E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, *104*(3), 700–712. https://doi.org/10.1037/a0027268

Roeser, R.W., & Eccles, J.S. (1998). Adolescents' perceptions of middle school: Relation to longitudinal changes in academic and psychological adjustment. *Journal of Research on Adolescence*, *8*(1), 123–158. https://doi.org/10.3102/0002831209361209

Rogers, C. (1961). *On becoming a person: A therapist's view of psychotherapy.* London: Constable.

Ryan, R.M., & Connell, J.P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*(5), 749–761. https://doi.org/10.1037//0022-3514.57.5.749

Ryan, R.M., & Deci, E.L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R.M., & Deci, E.L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol. 55,* 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R.M., & Deci, E.L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* New York: Guilford Press.

Ryan, R.M., & Weinstein, N. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education, 7*(2), 224–233. https://doi.org/10.1177/1477878509104327

Scrimin, S., Moscardino, U., Altoé, G., & Mason, L. (2018). Attentional bias for academic stressors and classroom climate predict adolescents' grades and socioemotional functioning. *Journal of Research on Adolescence: The Official Journal of the Society for Research on Adolescence*, *28*(1), 245–258. https://doi.org/10.1111/jora.12329

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J. Gustafsson (Eds.), *Teacher Quality, Instructional Quality and Student Outcomes* (pp. 51–80). Springer. https://doi.org/10.1007/978-3-319-41252-8_3

Selevko G.K. (2005). Entsiklopediya obrazovatel'nykh tekhnologii. [Encyclopedia of educational technologies]. Vol. 1. Moscow: Narodnoe obrazovanie.

Sheldon, K.M., Gordeeva, T.O. *The self-determination continuum. Encyclopedia of Personality and Individual Differences*. Springer, 2017.

Sheldon, K.M., Osin, E.N., Gordeeva, T.O., Suchkov, D.D., & Sychev, O.A. (2017). Evaluating the dimensionality of self-determination theory's relative autonomy continuum. *Personality and Social Psychology Bulletin, 43*(9), 1215–1238. https://doi.org/10.1177/0146167217711915

Sychev, O.A, Gordeeva, T.O., Lunkina M.V., Osin, E.N., Sidneva A.N. (in press). Mnogomernaya shkala udovletvorennosti zhizn'yu shkol'nikov [Questionnaire for diagnostics of life satisfaction of primary schoolchildren]. *Psikhologicheskaya nauka i obrazovanie [Psychological science and education]*.

Trenshaw, K.F., Revelo, R.A., Earl, K.A., & Herman, G.L. (2016). Using self-determination theory principles to promote engineering students' intrinsic motivation to learn. *International Journal of Engineering Education, 32*(3), 1194-1207.

Vallerand, R.J., Pelletier, L.G., Blais, M.R., Briere, N.M., Senecal, C., & Vallieres, E.F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement*, *53*(1), 159–172. https://doi.org/10.1177/0013164493053001018

Voronkova, I.V. (2003). Osobennosti uchebnoi motivatzii i intellektualnogo razvitiya starschich podrostkov, obuchayuschichsya v raznuh didactivheskich sistemah [Educational motivation and intellectual development of older adolescents studying in different educational systems]. *Psichologicheskaya nauka i obrazovaniye [Psychological Science and Education],* 4, 80-88.

Vorontsov, A., Zaslavsky, V., L'vovsky, V., Chudinova, E., & El'konin, B. (2013). *Razvivayushchee obuchenie. Model osnovnoy obrazovatelnoy programmy obrazovatelnogo uchrezhdeniya* [Developmental education. The model of basic educational program in educational institutions]. Moscow: Prosveschenie.

Vygotsky, L.S. (2017). The problem of teaching and mental development at school age [Problema obucheniya i umstvennogo razvitiya v shkol'nom vozraste]. *Changing English: Studies in Culture and Education. 24*(4), 359–371. https://doi.org/10.1080/1358684X.2017.1395165

Yang, C., Bear, G.G., Chen, F.F., Zhang, W., Blank, J.C., & Huang, X. (2013). Students' perceptions of school climate in the US and China. *School Psychology Quarterly*, *28*(1), 7–24. https://doi.org/10.1037/spq0000002

Zuckerman, G.A. (2005). Sistema Ėl'konina–Davydova kak resurs povyshenia kompetentnosti rossiiskikh shkol'nikov [The El'konin–Davydov system as a resource for enhancing the competence of Russian schoolchildren]. *Voprosy Psikhologii [Issues of Psychology]*, *4*, 84–95.

Zuckerman, G.A. (2011). Developmental education: A genetic modeling experiment. *Journal of Russian & East European Psychology*, *49*(6), 45–63. https://doi.org/10.2753/RPO1061-0405490603

Zuckerman, G.A. (2014). Developmental education. In A. Yasnitsky, R. van der Veer, & M. Ferrari (Eds.), *The Cambridge handbook of cultural-historical psychology* (pp. 177–202). Cambridge, UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139028097.011

Original manuscript received August 30, 2018
Revised manuscript accepted October 11, 2018
First published online December 20, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

Comparing Spatial Ability of Male and Female Students Completing Humanities vs. Technical Degrees

Elena A. Esipenko[a*], Ekaterina P. Maslennikova[b], Anna V. Budakova[a], Kseniya R. Sharafieva[a], Victoria I. Ismatullina[c], Inna V. Feklicheva[b], Nadezhda A. Chipeeva[b], Elena L. Soldatova[b], Zhanna E. Borodaeva[a], Kaili Rimfeld[d], Nikolas G. Shakeshaft[d], Margherita Malanchini[e], Sergey B. Malykh[c]

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

[b] *South Ural State University, Chelyabinsk, Russia*

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

[d] *King's College, University of London, UK*

[e] *Department of Psychology, University of Texas at Austin, USA*

[*] Corresponding author. E-mail: esipenkoea@gmail.com

Background. Spatial ability (SA) has long been the focus of research in psychology, because it is associated with performance in science, technologies, engineering, and mathematics (STEM). Research has shown that males consistently outperform females in most aspects of SA, which may partially explain the observed overrepresentation of male students seeking STEM degrees.

Objective. This study examines sex and field of study (degree) differences in different aspects of spatial ability and its structure.

Design. We assessed SA by using an on-line gamified battery, which included 10 spatial tests capturing 10 dimensions of spatial ability, among which were mental rotation, spatial visualization, spatial scanning, spatial reasoning, perspective-taking, and mechanical reasoning. The sample consisted of 882 STEM (55% males) and Humanities (20% males) university students in Russia.

Results. Males outperformed females on all assessed components of SA with a small effect size (1–11%). We also found that students from STEM fields outperformed Humanities students on all SA subtests (effect size ranged from 0.2 to 7%). These differences by study choice were not fully explained by the observed over-representation of males in the STEM group. The results of the study suggested no interaction between sex and degree. In other words, on average, males outperformed females, irrespective of whether they were STEM or humanities students; and the STEM advantage was observed for both

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

doi: 10.11621/pir.2018.0403
http://psychologyinrussia.com

males and females. The same unifactorial structure of SA was observed in the STEM and Humanities groups.

Conclusion. Our results are consistent with previous research, suggesting sex and study field differences in SA. Longitudinal research is needed to explore the causal mechanisms underscoring these differences.

Keywords: Spatial ability (SA), degree (field of study), gender differences, individual differences, STEM

Introduction

Spatial ability (SA) is the capacity to manipulate mental representations of objects and to understand the relationships between them (Wai, Lubinski & Benbow, 2009, Rimfeld, et al., 2017). It also represents a set of cognitive skills that help people in everyday life. SAs have been linked to a number of life outcomes. For example, better-developed SAs are associated with higher achievements in science, technology, engineering, and mathematics (hereafter identified as STEM) (Uttal et al., 2013, Shakeshaft, 2016).

Research has suggested that spatial ability is a factor separable from general intelligence, but no consensus exists on the exact nature and structure of SA (Harle & Towns, 2011; Shakeshaft, 2016). Several recent studies suggest that SA has mostly a unifactorial structure in adults, i.e., different aspects of SA show high correlation with each other (Rimfeld et al., 2017). However, it remains unclear, whether the observed differentiation among different SA tests results from measurement error or reflects partial construct independence. It is also not clear whether the structure of SA differs as a function of people's study or occupation. For example, for people whose university degree involves only some spatial tasks, it could be expected that these abilities will form a factor which is separate from other (untrained) aspects of SA.

In line with this hypothesis, research has suggested that spatial abilities may be associated with training in technical specialties, as students with technical degrees have been shown to have better spatial abilities (Rodán, et al., 2016). Initial strengths in SA may influence the choice of the study/professional direction (Kruglik, 2013, Kell, Lubinski, et al., 2013). In turn, STEM education may further develop SA. This is because students in STEM courses deal with spatially relevant material, such as geometry, more often than humanities students (Baenninger & Newcombe, 1989).

For example, chemistry students have been shown to outperform art students in mental rotation (Peters et al., 1995). Part of the explanation for this advantage, may be that chemistry students study molecular geometry, which requires 3D visualization for the construction of molecular structures (Harle & Towns, 2011). The results of a longitudinal study of the spatial abilities of engineering students also support the hypothesis that spatial elements of training can improve spatial abilities. The study showed that the average score of female engineering students on the Purdue Spatial Visualization Test: Rotations (PSVT: R; Guay & McDaniel, 1977) increased over the course of studying for their degree (Sorby & Veurink, 2010).

Another robust group difference in SA is the male advantage: males of all ages consistently outperform females on spatial ability tasks (Pannini et al., 2016; Vuok-

simaa et al., 2010; Hirnstein et al., 2014). The effect sizes of the differences vary from modest to moderate across studies, with the largest effect sizes of nearly 1.0 SD shown for 3-D mental rotation tasks (Harle & Towns, 2011, Halpern & Collaer, 2005). For example, in a study of mental rotation, although both sexes improved in spatial ability with age, males outperformed females ($\eta_p^2 = .16$) at all ages (9-23 years old) (Geiser et al., 2008). This result was replicated in another study of spatial abilities in adults (20-70 years; $\eta_p^2 = .13$) (Titze et al., 2010).

The observed sex differences in SA may partially explain the disproportion of girls and boys in STEM fields, i.e. the STEM gender gap observed in many countries. Indeed, the percentage of females among STEM graduates ranges from 12.4% in Macao to 40.7% in Algeria; the median is 25.4% (Stoet & Geary, 2018). This underrepresentation is surprising, considering the absence of sex differences in mathematical performance. For example, a recent study has found that 15-year old girls performed similarly or better than boys in science in 67% out of 67 countries participating in the 2015 Program of International Students Assessment (PISA). Another study suggests that in nearly all countries more girls are capable of pursuing college-level STEM study than currently enroll (Stoet & Geary, 2018).

One limitation of most previous studies is that they measured only one or a few aspects of SA. Therefore, there is a lack of understanding as to whether sex and study/occupational differences are present for some or all aspects of spatial ability (Harle & Towns, 2011). More insight is also needed into why differences in spatial ability exist between representatives of different occupations. Existing research suggests that students in STEM fields on average show greater spatial ability than students in Humanities for two reasons: 1) students with higher SA choose to study STEM; and 2) studying STEM improves SA. Exploring the structure of SA in STEM vs. Humanities students can provide new information on this issue.

Objectives

The current study investigates sex and study field differences in SA, using a gamified online test battery, in a sample of STEM and Humanities students. The battery consists of 10 different domains for SA, including 2D and 3D visualization, mental rotation, spatial relations, spatial planning, mechanical reasoning, spatial orientation, and spatial decision making. Based on previous research, we formulated three main hypotheses:

1. Men will outperform women on all 10 components of SA, with a small average effect.
2. STEM degree students will outperform Humanities degree students, with small average effect. This difference will not be explained solely by sex.
3. Both the STEM and Humanities groups will show a unifactorial structure of spatial abilities.

Design

Participants

The data were collected from 882 participants (38% males; mean age = 19.69; standard deviation = 2.09) attending leading universities in different cities in Russia. The

participants were studying for different degrees/occupations: Science, Technology, Engineering, and Mathematics (STEM) disciplines (N = 446, 55% males), and Humanities (Hum) (N = 406, 20% males). The STEM group included students studying for Technical and Natural science degrees. Additional analyses were performed on the three groups separately (Technical, Natural Sciences, and Humanities). These analyses (available from the authors) revealed that the Natural Sciences and Technical sciences groups produced a very similar pattern of results. Therefore, for the present report, we combined these groups in one STEM group.

Procedure

All participants gave consent for their participation. No reward was given. The data collection was anonymous, with only ID numbers entered into the database. First, participants completed the online spatial test battery called the King's Challenge (KC). Each item had a time limit. After the time was up, the test automatically switched to the next subtest. The average time for all subtests of the KC battery was approximately one hour. In addition to the KC battery, participants filled in a demographic questionnaire and took a verbal abilities test. This was not time-limited, and took no longer than 20 minutes. The study was approved by the Ethical Committee for Interdisciplinary Investigations, Tomsk State University.

Measures

Demographic questions. Participants completed a questionnaire that included demographic questions about their age, sex, university, and city of study, as well as information on the professional degree they were seeking.

Verbal questionnaire "My Vocab". The "My Vocab" test (Golovin, 2015) was used to assess verbal ability. The baseline design of the test is similar to the Mill Hill Vocabulary Scale (Raven, 1998). Russian vocabulary size was measured using a short version of the test (see Maslennikova et al., 2017). The test estimates the vocabulary of respondents aged 15 to 21 years old. The test consists of 99 test words, comprised of 95 real words and 4 fake words. The latter are included in order to assess whether participants complete the task honestly. This small number of test words allows the measurement of passive vocabulary volume.

We used the total score (the number of words that participants knew the meaning of). Data from participants who marked at least one non-existent word were excluded from the analysis. The resulting distribution showed a negative skew. Therefore, we excluded data from participants whose scores were less than 19. This resulted in a close-to-normal distribution. The verbal test was included in this study in order to assess whether the samples from different universities/cities were comparable.

Spatial ability battery "King's Challenge". Spatial ability was tested using an online gamified battery called "King's Challenge" (Rimfeld et al., 2017). The battery includes 10 spatial tests capturing the major putative dimensions of spatial ability. The 10 subtests are described in *Table 1*. At the end of the battery, participants received feedback on their performance across the 10 subtests. For additional information about the King's Challenge battery, see supplementary online information in Rimfeld et al. (2017).

Table 1

Description of the subtests of King's Challenge battery

Subtest name	N of items	Time limit per item (sec)	Description
Cross-sections	15	20	visualizing cross-sections of objects
2D drawing	5	45	sketching a 2D layout of a 3D object from a specified viewpoint
Pattern assembly	15	20	visually combining pieces of objects together to make a whole
Elithorn mazes	10	7	joining together as many dots as possible from an array
Mechanical reasoning	16	25	multiple-choice naive physics questions
Paper folding	15	20	visualizing where the holes are situated after a piece of paper is folded and a hole is punched through it
3D drawing	7	70	sketching a 3D drawing from a 2D diagram
Shape rotation	15	20	mentally rotating objects
Perspective-taking	15	20	visualizing objects from a different perspective
Mazes	10	25	searching for a way through a 2D maze in time limited task

Statistical analysis. ANOVA, exploratory factor analysis and confirmatory factor analysis of the data were conducted using software SPSS 22.0 and JASP 0.8.6.0.

Results

The participants' vocabulary corresponded to the reported average values of vocabulary of men and women of their age and level of education (Golovin, 2015). We further analyzed the vocabulary to assess the homogeneity of the sample. There were no significant differences in the volume of vocabulary between men and women ($F = .79$, $p = .37$, $\eta^2 = .002$), or field of study ($F = 4.11$, $p = .04$, $\eta^2 = .006$). The volume of vocabulary differed significantly among respondents from different cities ($F = 9.08$, $p<.001$); however, these differences were small ($\eta^2 = .066$). Details of these analyses are available from the authors.

Tables 2 and 3 present descriptive statistics for each subtest of the "King's Challenge" (KC) battery, by sex and degree. The variables were standardized over the whole sample. The tables also present the levels of difficulty of the tasks, expressed as a proportion of correct responses.

Sex differences

ANOVA results showed that males significantly outperformed females on all subtests. Effect sizes ranged from .01 (for pattern assembly, paper folding and shape rotation) to .11 (for mechanical reasoning). All comparisons by sex are presented in *Table 2*. The size of the effect on the total KC score was .05.

Table 2

Results of the analysis of variance by sex for the King's Challenge (KC) subtests

Group		N	% of correct responses	M (SD)	Levene's test	ANOVA	η^2
					p	F	
Cross-sections	male female	336 564	48.5 (25.15) 40.8 (25.37)	7.29 (3.81) 6.35 (3.55)	.08	13.93***	.02
2D drawing	male female	325 545	76.8 (21.38) 70.8 (21.32)	3.84 (1.07) 3.51 (1.09)	.09	19.07***	.02
Pattern assembly	male female	331 557	45.5 (20.82) 41.0 (19.80)	6.71 (3.16) 6.15 (2.97)	.16	7.01**	.01
Elithorn	male female	298 460	87.8 (9.43) 82.4 (10.02)	8.77 (0.95) 8.22 (1.02)	.07	54.08***	.07
Mechanical reasoning	male female	330 551	68.4 (15.92) 56.6 (15.93)	10.88 (2.58) 9.03 (2.54)	.78	108.86***	.11
Paper folding	male female	329 547	57.5 (28.14) 53.0 (28.80)	8.56 (4.24) 7.93 (4.27)	.46	4.41*	.01
3D drawing	male female	327 542	41.2 (27.70) 33.2 (24.40)	2.86 (1.94) 2.32 (1.71)	.00	17.69***	.02
Shape rotation	male female	327 540	50.4 (28.62) 48.5 (26.62)	8.21 (4.03) 7.48 (4.01)	.76	6.87**	.01
Perspective-taking	male female	327 539	39.6 (30.82) 23.8 (23.55)	5.84 (4.63) 3.57 (3.54)	.00	66.50***	.07
Mazes	male female	327 537	57.7 (19.93) 51.9 (20.10)	5.74 (2.02) 5.13 (2.01)	.63	18.48***	.02
KC Total score	male female	327 537	54.8 (17.03) 47.1 (15.96)	67.97 (20.02) 58.90 (18.76)	.22	45.06***	.05

*Note. * p < .05. ** p < .01. *** p < .001. KC = King's Challenge*

Levene's test shows that for most subtests and the total KC score, the variances were homogeneous. The two exceptions were: 3D drawing (rotation and visualization) and perspective-taking (spatial orientation and visualization).

Differences by degree (field of study)

Table 3 presents the comparisons according to the student's degree of study. Significant differences (p<.001) were found between the STEM and Humanities groups for all subtests and the total score of KC. The size of effect varied from .02 (for pattern assembly) to .10 (total KC score).

Levene's test showed homogeneous variances for all tests except 2D, 3D, and perspective-taking.

After controlling for sex, the effect sizes for some of the tests were significantly reduced, such as for mechanical reasoning (from .07 to .02) and perspective-taking (from .06 to .02). However, for other tests (2D, 3D, and total score), this change was less pronounced.

Table 3

Analysis of variance by field of study

Group		N	% of correct responses	M (SD)	Levene's test	ANOVA	η^2	η^2 after control sex
					p	F		
Cross-sections	STEM Hum	453 418	49.9 (24.49) 36.7 (25.00)	7.49 (3.65) 5.89 (3.48)	.50	43.18***	.05	.03
2D drawing	STEM Hum	445 400	78.2 (19.47) 66.9 (22.29)	3.91 (.97) 3.33 (1.11)	.00	64.44***	.08	.05
Pattern assembly	STEM Hum	447 413	45.2 (20.73) 39.7 (19.36)	6.77 (3.11) 6.00 (2.89)	.17	13.88***	.02	.01
Elithorn	STEM Hum	401 335	85.8 (10.09) 83.1 (10.01)	8.58 (1.00) 8.31 (.10)	.65	13.39***	.02	.002
Mechanical reasoning	STEM Hum	442 411	65.0 (16.66) 56.6 (16.08)	10.41 (2.66) 9.03(2.56)	.16	59.01***	.07	.02
Paper folding	STEM Hum	439 409	60.7 (28.08) 48.0 (27.73)	9.08 (4.22) 7.21 (4.11)	.54	42.53***	.05	.04
3D drawing	STEM Hum	435 407	43.9 (26.37) 27.5 (22.62)	3.07 (1.85) 1.94 (1.60)	.00	90.05***	.08	.07
Shape rotation	STEM Hum	435 405	53.8 (29.15) 46.9 (26.12)	8.40 (4.03) 7.03 (3.93)	.63	24.80***	.03	.002
Perspective-taking	STEM Hum	435 404	36.1 (29.39) 22.2 (23,69)	5.44 (4.40) 3.43 (3.61)	.00	51.84***	.06	.02
Mazes	STEM Hum	434 403	57.3 (19.00) 50.6 (21.00)	5.71 (1.90) 5.02 (2.10)	.06	24.34***	.03	.01
KC Total score	STEM Hum	434 403	55.1 (15.96) 45.6 (14.96)	68.33 (19.03) 56.22 (18.38)	.45	87.40***	.10	.06

*Note: * p < .05. ** p < .01. *** p < .001. STEM = STEM degree; Hum = Humanities degree; KC = King's Challenge*

Sex by Study field (degree) Interaction

We performed a 2 (STEM versus Humanities) x 2 (male versus female) ANOVA (See *Table 4*). The results revealed a major effect of sex ($F = 14.57$, $p < .00$, $\eta^2 = .016$); and a main effect of degree ($F = 49.02$, $p<.001$, $\eta^2 = .06$). The contribution of the degree was greater than the contribution of sex. No interaction was found between sex and degree ($F = .01$, $p<.94$, $\eta^2 = .00$).

Table 4

2×2 ANOVA assessing the effect of sex and degree with interaction between the 2 factors

Cases	Sum of Squares	Degrees of freedom	Mean squares	F statistic	*p* value	η^2
Degree	16914.17	1	16914.17	49.02	<.001	.06
Sex	5028.66	1	5028.66	14.57	<.001	.02
Degree*Sex	2.696	1	2.70	.01	.94	.00
Residual	287422.80	833.00	345.05			

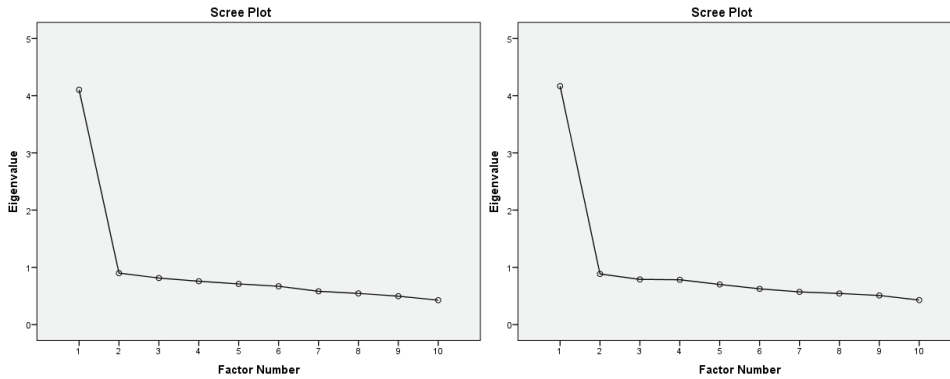
Number of observations = 837; R^2 = .11; adjusted R^2 = .108.

Factor analysis

First, we tested the 10 variables using both the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (MSA) and Bartlett's test of sphericity. For the STEM group, the KMO value was .90 and the chi-square value was 1007.79 ($p < .001$). For the Humanities group, the KMO value was .91 and chi-square value was 848.86 ($p < .001$). These indicated that the data could be used with factor analysis (Hair et al.,1998).

We determined the number of factors using an eigenvalue of more than one. For the STEM group, one factor (eigenvalue 4.10) explained 41.01% of the variance. For the Humanities group, one factor (eigenvalue 4.16) explained 41.65% of the variance (see *Table 5*).

Examination of the scree plots (*Fig. 1*) revealed an obvious "elbow" after the first factor. The results of the factor analysis show that the STEM and Humanities groups have similar unifactorial structures of spatial abilities.



STEM Humanities

Figure 1. Scree Plots illustrating the factor structure of spatial abilities
(Left panel: STEM sample; Right panel: Humanities sample).

Table 5

Factor analysis

	STEM	Humanities
	Factor 1	Factor 1
3D drawing	.73	.73
Paper folding	.67	.69
Cross-sections	.65	.67
2D drawing	.64	.61
Shape rotation	.63	.60
Mechanical reasoning	.58	.60
Perspective-taking	.55	.57
Pattern assembly	.51	.52
Mazes	.47	.48
Elithorn	.40	.43

Discussion

The results show that our first hypothesis was supported. Males on average outperformed females in all spatial tasks, with effect sizes ranging from .02 to .11. These results are consistent with previous studies, demonstrating male advantage in SA in adults, with similar effect sizes (Titze et al., 2010; Linn & Petersen, 1985). As in previous studies, the largest differences were observed in tasks that required understanding of everyday physics, and the ability to make a fast decision about trajectory (mechanical reasoning and Elithorn mazes subtests). The better performance by males on these tasks is in line with the hypothesis that boys have greater experience with certain tasks that promote greater development of spatial cognition, such as activities that involve eye-hand coordination, estimating trajectories of moving objects, or moving about within a complex spatial configuration (Cherney, 2010).

Surprisingly, mental rotation (shape rotation subtest) showed small sex differences in this study (.2). Previous research showed a greater effect size (.5-.9) for this ability (Linn & Petersen, 1985; Parsons, 2004). This difference may be due to the differences in the tasks presented. The 2D figure with four different pieces inside which we used in our study may require somewhat different processing than figures of the Vandenberg Mental Rotations task, often used in previous studies. In the present study, the details which could be markers for rotation were obvious, and rotation required only left/right axis; this is different from when stimuli require rotation along both the top/bottom axis and the left/right axis (Peters et al., 1995). However, the small effect size is unlikely to be explained by the reduced difficulty of the task. As shown by the proportion of correct responses, the task was difficult

(approximately 50% correct responses). In addition, there was no correspondence between the difficulty of the tasks and the effect size.

Our second hypothesis was also confirmed. The STEM group outperformed the Humanities group in all SA subtests and the total score (effect size ranged from .002 to .7). The biggest advantage of STEM degree students was observed for 2D and 3D drawing, as well as on the KC total score (5-7%). The observed differences between degrees (fields of study), such as in 3D drawing, may be partly explained by the STEM students' extensive use of computer technologies, such as software for design and simulation, which often use 3D manipulation of objects (Metz, et al., 2012; Sorby & Veurink, 2010).

The results of the study suggested no interaction between sex and degree. In other words, males on average outperformed females, irrespective of whether they were STEM or Humanities students; and the STEM advantage was observed for both males and females. Moreover, the field of study differences were not fully explained by the over-representation of males in the STEM group. The differences between the two degrees were reduced but did not disappear after correcting for sex.

The reasons for over-representation of men in STEM and of women in humanities which was observed in this study, remain poorly understood, and may include gender stereotypes regarding educational choices (Ramaci, et al., 2017). The STEM gender gap may also emerge partly because students select their field of study based on their evaluation of their intra-individual strengths and weaknesses, irrespective of their objective ability (e.g., Stoet & Geary, 2018). According to previous research on common educational practice, students tend to select their area of study based on an evaluation of their intra-individual strengths and weaknesses (e.g., Gardner, 2016). Research has shown that in a majority of the countries studied, boys, on average, performed better in math (100% of countries) and science (97% of countries) than in reading. By contrast, girls, on average, showed higher performance in reading than in math and science in all countries (Stoet & Geary, 2015).

Finally, our study's third hypothesis was also confirmed. Despite average differences in performance, the structure of spatial abilities for STEM and Humanities groups was very similar. For both groups, one factor explained a very similar proportion of the overall variance (41 and 42%) and showed a unifactorial structure of SA. This result suggests that, if pursuing STEM education has an influence on SA, this influence is global, affecting all aspects of SA rather than a sub-set of practiced skills. An alternative explanation is that the SA advantage shown by STEM students, partly accounts for why they are pursuing STEM, rather than resulting from their practice while studying for their degree. Longitudinal research is needed to further investigate these alternative explanations.

Conclusion

This investigation explored the spatial ability of male and female students completing Humanities and STEM university degrees. Males on average performed better on spatial tasks than females. STEM students outperformed Humanity students in all SA tasks, and these differences weren't completely explained by the over-representation of males in the STEM group. There was no statistical interaction between

sex and field of study. Both groups had a unifactorial structure of spatial abilities. Longitudinal research is needed for a better understanding of the processes underlying the group differences we observed.

Limitations

This study had a number of limitations. First, the homogeneity of the sample was measured by a short verbal ability test, rather than a measure of general cognitive ability. Second, the results showed heterogeneity of variance on a number of subtests, which may indicate task-specific technical issues. Third, the study was not longitudinal and did not allow for evaluating whether studying for a STEM degree can enhance SA.

Acknowledgements

This work was supported by the Russian Ministry for Education and Science project № 25.8905.2017/9.10.

References

Baenninger, M. & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex roles*, *20*(5–6), 327–344. https://doi.org/10.1007/BF00287729

Cherney, I.D. & Voyer, D. (2010). Development of a Spatial Activity Questionnaire I: Items Identification. *Sex Roles*, *62*(1–2), 89–99. https://doi.org/10.1007/s11199-009-9710-9

Gardner, A. (2016). How important are GCSE choices when it comes to university? London, UK: Which? Retrieved from http://university.which.co.uk/advice/gcse-choices-university/how-important-are-gcse-choices-when-it-comes-to-university.

Geiser, C., Lehmann, W., & Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence*, *36*(6), 556–563. https://doi.org/10.1016/j.intell.2007.12.003

Golovin, G.V. (2015). Izmerenie passivnogo slovarnogo zapasa russkogo yazyka [Receptive vocabulary size measurement for Russian language]. *Socio- i psiholingvisticheskie issledovaniya* [Socio-Psycholinguistic Research], *3*, 148–159.

Guay, R.B. & McDaniel, E.D. (1977). The relationship between mathematics achievement and spatial abilities among elementary school children. *Journal for Research in Mathematics Education*, *8*(3), 211–215. https://doi.org/10.2307/748522

Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1998). *Multivariate data analysis*. Upper Saddle River.

Halpern, D.F., & Collaer, M.L. (2005). Sex differences in visuospatial abilities: More than meets the eye. In P. Shah & A. Miake (Eds.), *The Cambridge Handbook of Visuospatial Thinking* (pp. 170–212). New York: Cambridge University Press.

Harle, M. & Towns, M. (2011). A Review of Spatial Ability Literature, Its Connection to Chemistry, and Implications for Instruction. *Journal of Chemical Education*, *88*(3), 351–360. https://doi.org/10.1021/ed900003n.

Hirnstein, M., Andrews, L.C., & Hausmann, M. (2014). Gender-Stereotyping and Cognitive Sex Differences in Mixed- and Same-Sex Groups. *Archives of Sexual Behavior*, *43*(8), 1663–1673. https://doi.org/10.1007/s10508-014-0311-5

Kell, H.J., Lubinski, D., Benbow, C.P., & Steiger, J.H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological science*, *24*(9), 1831–1836. https://doi.org/10.1177/0956797613478615

Kruglik, A.V. (2013). Prostranstvennyy faktor v strukture intellektual'nykh sposobnostey uchashchikhsya [Spatial factor in students' intellectual abilities structure]. Psikhologicheskii zhurnal [*Psychological journal*], *1–2*, 112–116. Retrieved from http://elib.bsu.by/handle/123456789/106262

Linn, M.C. & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, *56*(6), 1479–1498. https://doi.org/10.2307/1130467

Maslennikova, E.P., et al. (2017). Slovarnyi zapas, kak pokazatel' verbal'nogo intellekta: primenenie ekspress-metodiki otsenki slovarnogo zapasa [Vocabulary Size as a Verbal Intelligence Index: Application of the Express Methods of Estimating Vocabulary Size]. Vestnik yuzhnoural'skogo gosudarstvennogo universiteta. Seriya Psikhologiya [*Bulletin of the South Ural State University. Ser. Psychology*], *10*(3), 63–69. https://doi.org/10.14529/psy170306

Metz, S.S., Donohue, S., & Moore, C. (2012). Spatial skills: A focus on gender and engineering. *International Journal of Science Education*, *31*, 1–8.

Pannini, G. et al. (2016). Testing the Testosterone Transmission Hypothesis: An Investigation of Twin Performance On Spatial Ability Tasks. *Intelligence Research (ISIR). Abstracts*, 87.

Parsons, T.D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J.G., & Rizzo, A.A. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia*, *42(4),* 555–562. https://doi.org/10.1016/j.neuropsychologia.2003.08.014

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain and cognition*, *28*(1), 39–58. https://doi.org/10.1006/brcg.1995.1032

Ramaci, T., Pellerone, M., Ledda, C., Presti, G., Squatrito, V., & Rapisarda, V. (2017). Gender stereotypes in occupational choice: a cross-sectional study on a group of Italian adolescents. *Psychology Research and Behavior Management*, *10*, 109–117. https://doi.org/10.2147/PRBM.S134132

Rimfeld, K., Shakeshaft, N.G., Malanchini, M., Rodic, M., Selzam, S., Schofield, K., & Plomin, R. (2017). Phenotypic and genetic evidence for a unifactorial structure of spatial abilities. *The Proceedings of the National Academy of Sciences in the United States of America, 114*(10), 2777–2782. https://doi.org/10.1073/pnas.1607883114

Rodán, A., Contreras, M.J., Elosúa, M.R., & Gimeno, P. (2016). Experimental but Not Sex Differences of a Mental Rotation Training Program on Adolescents. *Fronters in Psychology*, *7,* 1050. https://doi.org/10.3389/fpsyg.2016.01050.

Sakamoto, M. & Spiers, M.V. (2014). Sex and cultural differences in spatial performance between Japanese and North Americans. *Archives of Sexual Behavior*, *43*(3), 483–491. https://doi.org/10.1007/s10508-013-0232-8

Shakeshaft, N.G., Rimfeld, K., Schofield, K.L., Selzam, S., Malanchini, M., Rodic, M., & Plomin, R. (2016). Rotation is visualisation, 3D is 2D: using a novel measure to investigate the genetics of spatial ability. *Scientific reports*, *6*, 30545. https://doi.org/10.1038/srep30545

Sorby, S. & Veurink, N. (2010). Are The Visualization Skills Of First Year Engineering Students Changing? Paper presented at 2010 Annual Conference & Exposition, Louisville, Kentucky. Retrieved from https://peer.asee.org/16026

Stoet, G. & Geary, D.C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, *48*, 137–151. https://doi.org/10.1016/j.intell.2014.11.006.

Stoet, G. & Geary, D.C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological science*, *29*(4), 581–593. https://doi.org/10.1177/0956797617741719

Titze, C., Jansen, P., & Heil, M. (2010). Mental rotation performance and the effect of gender in fourth graders and adults. *European Journal of Developmental Psychology*, *7*(4), 432–444. https://doi.org/10.1080/17405620802548214

Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological bulletin*, *139,* 352–402. https://doi.org/10.1037/a0028446

Vuoksimaa, E., Kaprio, J., Kremen, W.S., Hokkanen, L., Viken, R.J., Tuulio-Henriksson, A., & Rose, R.J. (2010). Having a Male Co-Twin Masculinizes Mental Rotation Performance in Females. *Psychological Science*, *21*(8)*,* 1069–1071.https://doi.org/10.1177/0956797610376075

Wai, J., Lubinski, D., & Benbow, C.P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4)*,* 817–835. https://doi.org/10.1037/a0016127

Original manuscript received August 25, 2018
Revised manuscript accepted October 20, 2018
First published online December 20, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

Spouses' Psychological States and Family Relations in Families with Natural and Induced Pregnancies

Tatiana G. Bokhan[a]*, Olga V. Terekhina[a], Marina V. Shabalovskaya[a,b], Svetlana B. Leshchinskaia[a], Anna V. Silaeva[a,b], Elena A. Naku[a], Fatos Selita[a,c], Lyubov A. Agarkova[d]

[a] *National Research Tomsk State University, Tomsk, Russia,*

[b] *Siberian State Medical University, Tomsk, Russia*

[c] *Department of Psychology, Goldsmiths, University of London, United Kingdom*

[d] *Research Institute of Obstetrics, Gynecology and Perinatology, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russia*

* Corresponding author. E-mail: btg960@mail.ru

Background. Psychological tension in the family, along with stress and mental and physical illness, are linked to the reproductive health of parents, as well as to the outcomes of infertility treatments and pregnancy overall.

Objective. To compare stress and negative affect (depression, irritability, and anxiety) in families with induced pregnancies (in-vitro fertilization, IFV) vs. natural pregnancies. The relationship between negative affect and stress in pregnant women was explored in both groups. Finally, the study investigated links between negative affect and partner relationships.

Design. The sample included 308 women and 278 men from couples with natural conception, and 131 women and 102 men from couples with an IVF pregnancy.

Results. Relatively low levels of negative affective states and stress were found in families with both natural and induced pregnancies. Moderate correlations were found between women's negative affect and their stress level in both groups. Significant correlations were found in both groups between negative psychological states of the spouses, as well as between negative psychological states and warmth/hostility in marital relations.

Conclusion. The results suggest that psychological states, stress levels, and links between psychological states and quality of family relations are similar in families with IVF and those with natural pregnancies. Further longitudinal research is needed to explore the direction of causal links between the psychological states of the spouses, and between their psychological states and the quality of family relations.

***Keywords*:** pregnancy, IVF, infertility, family relationships, stress, psychological states

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

doi: 10.11621/pir.2018.0404
http://psychologyinrussia.com

Introduction

Psychological states of spouses may influence communication and the quality of the relationship between them. In turn, the quality of spousal relations may influence their psychological states (Kouros, Papp, & Cummings, 2008; Vujeva & Furman, 2011). Psychological tension in the family, along with stress and mental and physical illness, have been shown to affect the reproductive health of the parents, as well as the outcomes of infertility treatments and pregnancy overall (Ebbesen et al., 2009; Fadeeva, Vostrikov, & Garganeeva, 2011; Frederiksen, Farver-Vestergaard, Skovgård, Ingerslev, & Zachariae, 2015; Galhardo, Cunha, & Pinto-Gouveia, 2011).

Psychological States of Spouses during IVF Treatment

A number of studies have investigated the psychological states of spouses while undergoing in-vitro fertilization (IVF) treatment, as well as during the pregnancy (Feklicheva et al., 2017; Greil, Shreffler, Schmidt, & McQuillan, 2011; Klemetti, Raitanen, Sihvo, Saarni, & Koponen, 2010; Naku et al. 2016; Petrova, Podolhov, Gzgzyan, & Ngauri, 2013; Purewal, Chapman, & Van Den Akker, 2017). Research suggests that diagnosed infertility can be almost as stressful as the loss of a relative, incurable disease, divorce, and other extremely traumatic events (Naku, Kovas, Bohan, Terehina, &Vidyakina, 2017). In some people, in-vitro fertilization can lead to an inferiority complex, psychological suffering, and anxiety (Wichman, Ehlers, Wichman, Weaver, & Coddington, 2011; Lin & Chueh, 2016). Such psychological problems may occur when it is impossible to conceive a child naturally and after unsuccessful IVF attempts (Filippova, 2009; Hynes, Callan, Terry, & Gallois, 2011). One or both partners (spouses) may have a depressed *mood*, decreased energy, and low general well-being (Haimovici et al., 2018; Maroufizadeh, Karimi, Vesali, & Omani Samani, 2015; Pasch et al., 2016; Williams, Marsh, & Rasgon, 2007). Negative emotions of people suffering from infertility can also be accompanied by *negative self-perception*, a negative and inconsistent self-image, and *self-blame*, which can negatively affect family life satisfaction (Filippova, 2009; Greil, Slauson-Blevins, & McQuillan, 2010).

Studies of anxiety among pregnant women and their partners participating in an IVF treatment have produced inconsistent results (Gourounti et al., 2012; Hjelmsted, Widström, Wramsby, & Collins, 2003; Klock & Greenfeld, 2000). One study showed that women and men undergoing IVF had a higher level of anxiety about losing the pregnancy than people with natural conception (Hjelmsted, Widström, Wramsby, Matthiesen, & Collins, 2003). However, another prospective longitudinal study found that the women conceiving through IVF (but not the naturally conceiving women) had on average decreased anxiety and increased self-esteem during pregnancy (Klock & Greenfeld, 2000). Another study found that fertility-related stress and state anxiety positively correlated with avoidance coping and low perception of personal control in women undergoing IVF (Gourounti et al., 2012).

Apart from anxiety, other negative affective states have also been found in women undergoing IVF (Petrova et al., 2013; Seok Kee, Jung, & Lee, 2000; Sbaragli et al., 2008; Zaharova & Yakupova, 2015). For example, one study showed that

more women in this group tended to ignore actual and potential problems, to idealize pregnancy and motherhood, and to have euphoric and unrealistic ideas about their future child and about themselves (Zaharova & Yakupova, 2015).

Family Relations

Much research has found a link between marital conflicts and psychological disorders (Choi & Marks, 2008; Du Rocher Schudlich, Papp, & Cummings, 2011; Kouros & Cummings, 2011; Pellerone & Micciché, 2014). Lack of mutual understanding in the family may lead to depression, *estrangement*, a decline in psychological and physical health, and a decrease in the partners' ability to work. Specifically, for families engaged in the IVF procedure, spousal relationships may face serious challenges at all stages: decision making, participation in the IVF program, period of pregnancy, childbirth, and child development (Greil et al., 2010). Special circumstances associated with IVF, such as diagnosed infertility (both male and female), unsuccessful conception attempts, and a complicated pregnancy, can have a negative impact on the psychological state of the partners and their family relationships, and may undermine the development of parental identity (Faria, Grieco, & Barros, 2012).

To sum up, research suggests that induced pregnancy, such as through IVF, is accompanied by higher levels of anxiety and other negative affective states, strained family relations, and increased risk of miscarriage (Massey et al., 2016). In addition, it is often preceded by unsuccessful IVF attempts, which may also lead to elevated stress. Therefore, IVF pregnancy is considered to be more stressful, on average, than natural pregnancy. However, some studies did not find differences between families with naturally conceived and induced pregnancies. Moreover, to date limited research is available on the association between family relationships and psychological states of partners in families who undergo IVF, compared to families with natural pregnancy.

Methods

Hypotheses

The present study compared affective states in families with induced and natural pregnancies, and explored the association between negative affect and marital relations. Based on previous research, the following five hypotheses were formulated:

1. Families with induced pregnancy will on average experience greater stress and negative affective states than families with natural pregnancies.
2. Negative affect will be correlated with women's stress during the pregnancy.
3. Negative affect will be associated with more problematic marital relations in both types of families.
4. Psychological states of spouses will be modestly correlated—i.e., the partners, on average, will show some similarity in their psychological states.
5. There will be a modest to moderate correlation between warmth or hostility that partners report towards their partners and what their partners perceive about them.

Participants

Participants were part of the Prospective Longitudinal Interdisciplinary Study (PLIS) conducted in Russia (see Voronina, Bohan, Terehina, Malykh, & Kovas, 2016, for details). The study was approved by the Ethics Committee for Interdisciplinary Investigations, Tomsk State University. All participants provided written informed consent. The sample included 439 women and 380 men: 308 women and 278 men were from couples with natural conception, and 131 women and 102 men with an IVF pregnancy. All participants were recruited from four family-planning clinics in three cities of Russia's Siberian Federal Districts.

Measures

Men completed a *Questionnaire for the Father* and women completed a *Questionnaire for the Mother* during the first and third trimesters of pregnancy.

1. *Irritability, Depression and Anxiety* scale (IDA; Snaith, Constantopoulos, Jardine, & McGuffin, 1978). The IDA, translated and adapted for use in Russian, contains 18 statements, such as "I feel cheerful", "I feel I might lose control and hit or hurt someone", and "I get angry with myself or call myself names". Responses to the statements are measured on a 4-point Likert scale, ranging from 0 (No, not at all) to 3 (Yes, definitely). The questionnaire measures the intensity of depression, outward irritability, inward irritability, and anxiety. For each scale there is a cut-off point, with exceeding values indicating the presence of depression (4–6 points), anxiety (6–8 points), outward irritability (5–7 points), and inward irritability (4–6 points).

2. *Emotional State during Pregnancy Scale* (Rice et al. 2010). This single-item measure, translated and adapted for use in Russian, was completed only by women. The woman indicates on an 11-point Likert scale (from 0 to 10) how stressed and worried she feels (10 means "calm and relaxed", 0 means "stressed and worried"). The data were collected twice, evaluating three different periods of the pregnancy: in the first trimester, women reported about the first 12 weeks of pregnancy; in the third trimester, they responded about the period from 13 to 25 weeks and the period from the 26th week of pregnancy. Values from 0 to 3 indicate high stress, from 4 to 7 — medium (optimum), and from 8 to 10 — low.

3. *Iowa Family Interaction Rating Scales* (Melby et al., 1998). The measure includes nine items, each assessed on a 7-point Likert scale (1–never, 7–always). Each spouse assessed their own emotional warmth (five questions) and hostility (five questions) towards the partner, as well as the perceived warmth and hostility of the partner towards them. The warmth of the relationship was calculated as the mean score of: (a) the woman's warmth towards her partner, (b) the woman's perceived warmth of her partner towards her, (c) the man's warmth towards his partner, and (d)) the man's perceived warmth of his partner towards him. Hostility was calculated as the mean score of (a) the woman's hostility towards her partner, (b) the woman's perceived hostility of her partner towards her, (c) the man's hostility towards his partner, and (d) the man's perceived hostility of his partner towards him.

The overall index of the relationships was also estimated. Higher values on the warmth scale indicated warmer relationships; higher values on the hostility scale indicated greater hostility. The hostility scale was reversed, so that lower values indicated greater hostility. The overall index was calculated as the sum of the two

scales, so that low values indicated problems in family interactions and high values indicated good relationships.

The statistical analysis was performed with the IBM SPSS Statistics Version 23.0 software package. The following analyses were performed to test the hypotheses: descriptive statistics, *Spearman's* rank *correlation,* the *Mann-Whitney* U test, and correlation comparisons using *Fisher*-Z-Transformation.

Results

Negative Affective States of Pregnant Women

The Irritability, Depression and Anxiety scale was used to measure negative affective states. The mean scores of all parameters in both groups lie within the normal range (see Table 1).

Table 1

Descriptive statistics, Irritability, Depression and Anxiety scale (pregnant women)

	Depression (0–15 points)		**Outward Irritability (0–12 points)**		**Inward Irritability (0–12 points)**		**Anxiety (0–15 points)**	
	Natural	**IVF**	**Natural**	**IVF**	**Natural**	**IVF**	**Natural**	**IVF**
N	278	102	278	102	278	102	278	102
Mean	3.14	3.26	3.43	2.93	1.18	1.00	5.51	5.50
Median	3.00	3.00	3.00	2.50	1.00	1.00	5.00	5.00
Mode	2.00	3.00	1.00	2.00	0.00	0.00	6.00	4.00
Std. Deviation	2.15	2.14	2.19	2.03	1.26	1.16	2.88	2.88
Dispersion	4.63	4.59	4.84	4.12	1.60	1.34	8.29	8.27
% of women with elevated levels	6.4	3.9	4.7	2.3	0.3	0	14.9	12.2
Mann-Whitney U test	13,481.5		12,448		13,100		14,205	
p	0.55		0.05		0.21		0.98	

An increased level of negative psychological states was observed in some women in both groups. High levels of depression (> 6 points) were observed in 6.4% of women with natural conception and 3.9% of women with IVF. High levels of outward irritability (> 7 points) were observed in 4.7% of women with natural conception and 2.3% of women with IVF. High levels of inward irritability (> 6 points) were observed in 0.3% of women with natural conception and none of the women with IVF. High levels of anxiety (> 8 points) were observed in 14.9% of women with natural conception and 12.2% of women with IVF.

The *Mann-Whitney* U test revealed significant differences in outward irritability between the two groups, with significantly higher levels in women with natural pregnancy than women with induced pregnancy.

Stress of Women during Pregnancy

As shown in Table 2, the mean scores of women's stress level in both groups (induced vs. natural pregnancy) during all three trimesters were in the normal range (4–7 points). Mean scores for stress in both groups during the third trimester were lower than in the first and second trimesters, but the difference was not significant.

Table 2

Descriptive statistics for stress during pregnancy

	Stress					
	0–12 weeks		**13–25 weeks**		**from 26 weeks**	
	Natural	**IVF**	**Natural**	**IVF**	**Natural**	**IVF**
N	295	122	280	105	272	98
Mean	6.74	6.47	6.90	6.70	7.16	7.19
Median	7.00	6.00	7.00	7.00	8.00	7.00
Mode	8	6	8	8	8	6
Std. Deviation	2.27	2.09	2.16	2.59	1.94	1.94
Dispersion	5.16	4.38	4.66	6.71	3.76	3.75
% of women with 'at risk' stress level (0–3)	7.5	7.4	6.1	12.4	4	3.1
Mann-Whitney U test	16,472.5		14,409		13,250	
p	0.17		0.76		0.93	

In each trimester, women scoring below 3 points were identified as an at-risk group for developing mental health problems, potentially having an *unfavorable* influence on the *pregnancy and the postpartum period*. In the IVF group, increased stress was observed in 7.4% of women in the first trimester, 12.4% in the second trimester, and 3.1% in the third trimester. In the natural conception group, increased stress was observed in 7.5% of women in the first trimester, 6.1% in the second trimester, and 4% in the third trimester. The Mann-Whitney U test did not show significant differences between women with natural and induced pregnancy.

Association Between Negative Affect and Stress in Pregnant Women

Spearman's rank *correlation* was used to assess the relationship between negative affect and the stress of pregnant women (see Table 3). In the natural conception group, significant modest to moderate negative correlations were observed between all psychological states and stress level in all trimesters. Negative correlations indicate that higher stress is associated with worse psychological states (the stress level scale is reversed, with low scores corresponding to high stress). The highest correlation was found between stress in the third trimester and depression ($r = -0.51$; $p = 0.00$) and anxiety ($r = -0.58$; $p = 0.00$), suggesting that women who experienced greater stress also experienced significantly more negative affect.

In the IVF group, significant correlations were observed between depression and stress in all trimesters of pregnancy, between outward irritability and stress in

the first and second trimesters, and between anxiety and stress in the second and third trimesters. The highest correlation was found between depression and stress in the second trimester (r = -0.56; p = 0.00).

Fisher Z was used to test whether the observed correlations between negative affect and stress differed significantly between the IVF and natural pregnancy groups of women. The results showed that correlation between stress in the first trimester and inward irritability was greater in the natural conception group. Correlation between stress in the second trimester and depression was significantly greater in women with an IVF pregnancy than in the natural conception group. In the third trimester, correlations between stress and anxiety, stress and outward irritability, and stress and inward irritability were greater in the natural conception group than in the IVF group.

Table 3

Relationship between negative psychological states of pregnant women and their stress level

		Stress								
		first trimester			**second trimester**			**third trimester**		
		Natural	**IVF**	**Z (p value) Nat vs IVF**	**Natural**	**IVF**	**Z (p value) Nat vs IVF**	**Natural**	**IVF**	**Z (p value) Nat vs IVF**
Depression	r	**–0.18****	**–0.25***		**–0.31****	**–0.56****		**–0.51****	**–0.42****	
	p	0.00	0.01	0.61 (p=0.27)	0.00	0.00	**2.64 (p=0.00)**	0.00	0.00	–0.94 (p=0.17)
	N	268	95		274	100		268	93	
Outward Irritability	r	**–0.13***	**–0.22***		**–0.23****	**–0.24***		**–0.29****	–0.02	
	p	0.03	0.03	0.77 (p=0.22)	0.00	0.01	0.09 (p=0.46)	0.00	0.86	**–2.30 (p=0.01)**
	N	270	96		277	101		271	94	
Inward Irritability	r	**–0.26****	0.13		**–0.19****	–0.06		**–0.27****	–0.04	
	p	0.00	0.19	**–3.29 (p=0.00)**	0.00	0.54	–1.12 (p=0.13)	0.00	0.71	**–1.95 (p=0.03)**
	N	269	96		275	101		269	94	
Anxiety	r	**–0.27****	–0.17		**–0.34****	**–0.40****		**–0.58****	**–0.35****	
	p	0.00	0.10	–0.87 (p=0.19)	0.00	0.00	0.59 (p=0.28)	0.00	0.00	**–2.45 (p=0.01)**
	N	269	96		276	101		270	94	

*Note. **Correlation significant at p = 0.01. *Correlation significant at p = 0.05. Lower scores for stress indicate greater stress; therefore, the negative sign of the associations indicates positive associations between stress and negative affect (more stress is associated with more negative affect). Z = Fisher Z, comparison between IVF and natural pregnancy groups. The sign of Z can be ignored, with significance level indicating whether the stronger correlation (positive or negative) is significantly stronger.*

Negative Psychological States of Partners of Pregnant Women

The mean scores of women's partners' psychological states were overall similar to the women's scores, and were also in the normal range (see Table 4). The Mann-Whitney U test showed that the differences between the two groups were not significant.

Table 4

Descriptive statistics, Irritability, Depression and Anxiety scale (men)

	Depression		**Outward Irritability**		**Inward Irritability**		**Anxiety**	
	Natural	**IVF**	**Natural**	**IVF**	**Natural**	**IVF**	**Natural**	**IVF**
N	278	102	278	102	278	102	278	102
Mean	3.05	3.08	3.15	3.10	1.47	1.24	4.43	4.36
Median	3.00	3.00	3.00	3.00	1.00	1.00	4.00	4.00
Mode	3.00	4.00	3.00	1.00	0.00	0.00	3.00	4.00
Std. Deviation	2.01	1.78	2.30	2.38	1.53	1.30	2.64	2.40
Dispersion	4.03	3.16	5.30	5.65	2.36	1.70	6.97	5.76
% of men with elevated levels	4.8	2.3	3.5	3.9	0.9	0	5.8	3.8
Mann-Whitney U test	12,166		12,491		11,858.5		12,254.5	
p	0.61		0.78		0.28		0.81	

Relationship Between Negative Psychological States of Spouses and Quality of Marital Relations

Descriptive statistics for men's and women's experienced and perceived warmth and hostility are presented in Tables 5 and 6.

Table 5

Descriptive statistics of warmth (Iowa scale)

	Women's warmth		**Women's perceived warmth**		**Men's warmth**		**Men's perceived warmth**	
	Nat.	**IVF**	**Nat.**	**IVF**	**Nat.**	**IVF**	**Nat.**	**IVF**
N	277	101	274	100	277	92	280	92
Mean	27	28.3	27.2	286	27.4	29	27.7	29.4
Median	29	29	30	30,5	29	30	30	31
Mode	35	35	35	35	35	35	35	35
Std. Deviation	6.8	5.5	7.6	6.4	6.9	5.6	7.5	6.7
Dispersion	46.3	30.8	58.4	41.1	48.3	31	55.9	44.8

Table 6

Descriptive statistics of hostility (Iowa scale)

	Women's hostility		Women's perceived hostility		Men's hostility		Men's perceived hostility	
	Nat.	IVF	Nat.	IVF	Nat.	IVF	Nat.	IVF
N	278	100	275	101	270	91	278	92
Mean	20.9	22.4	22.1	23.4	21	22.5	20	20.6
Median	22	23	24	24	22	23	21	22
Mode	24	26	26	24	22	24	20	22
Std. Deviation	5	4.1	5.7	4.5	5	4	5.6	5.4
Dispersion	25	17.1	32.7	20.1	24.7	16.4	31.5	29.5

Table 7 presents the descriptive statistics for emotional warmth and hostility of the relationship, the composite index of marital relationships, and the results of comparative analysis of the groups. The Mann-Whitney U test demonstrates that all scores of marital relations in our sample are slightly higher in the IVF group than in the natural conception group, which indicates warmer and less hostile relationships.

Table 7

Warmth and hostility of the relationship and composite relationship index (Iowa scale)

	Warmth		Hostility (reversed)		Composite	
	Nat.	IVF	Nat.	IVF	Nat.	IVF
N	265	89	261	88	252	87
Mean	27.29	29	20.9	22.25	24.1	25.65
Median	29	30	22	23	25.5	26.5
Mode	35	34.25	22	25.25	26.9	20.25
Std. Deviation	6.4	5.175	4.75	3.72	5.3	4.21
Dispersion	41.2	26.8	22.6	13.8	28.2	17.7
Mann-Whitney U test	9,990		9,766		9,149	
p	0.03		0.04		0.02	

Note. The composite index was calculated as the sum of the two scales, where the scale of hostility was reversed (high scores indicate low hostility).

Table 8 presents correlations between family relations (warmth, hostility, and overall relations) and psychological states (depression, anxiety, inward and outward irritability), separately for the IVF and natural pregnancy groups, and

Table 8

Relationship between negative psychological states of spouses and marital relations

Pairs of variables (r)	Women Nat. N=247-263	Women IVF N=84-86	Z (p value) Nat vs IVF Women	Men Nat N=246-263	Men IVF N=85-89	Z (p value) Nat vs IVF Men
Warmth–Depression	–0.18**	–0.30**	1.0 (p=0.16)	–0.26**	–0.14	–0.97 (p=0.17)
Warmth–Outward Irritability	–0.20**	–0.30**	0.79 (p=0.21)	–0.24**	–0.21	–0.31 (p=0.38)
Warmth–Inward Irritability	–0.17**	–0.25*	0.65 (p=0.26)	–0.18**	–0.17	–0.10 (p=0.46)
Warmth–Anxiety	–0.22**	–0.08	–1.07 (p=0.14)	–0.19**	–0.17	–0.16 (p=0.44)
Hostility–Depression	–0.24**	–0.32**	0.68 (p=0.25)	–0.22**	–0.27*	0.45 (p=0.33)
Hostility–Outward Irritability	–0.35**	–0.42**	0.57 (p=0.28)	–0.31**	–0.22*	–0.79 (p=0.21)
Hostility–Inward Irritability	–0.30**	–0.20	–0.86 (p=0.19)	–0.25**	–0.16	–0.73 (p=0.23)
Hostility–Anxiety	–0.29**	–0.13	–1.39 (p=0.08)	–0.21**	–0.09	–0.96 (p=0.17)
Relations–Depression	–0.35**	–0.35**	1.23 (p=0.11)	–0.26**	–.22*	–0.36 (p=0.36)
Relations–Outward Irritability	–0.30**	–0.40**	0.93 (p=0.18)	–0.30**	–.23*	–0.56 (p=0.29)
Relations–Inward Irritability	–0.25**	–0.25*	0 (p=0.5)	–0.25**	–0.16	–0.43 (p=0.33)
Relations–Anxiety	–0.27**	–0.12	–1.21 (p=0.11)	–0.27**	–0.15	–0.44 (p=0.33)

*Note. **Correlation significant at p = 0.01. *Correlation significant at p = 0.05. Z = Fisher Z, comparison between IVF and natural pregnancy groups. The sign of Z can be ignored, with significance level indicating whether the stronger correlation (positive or negative) is significantly stronger.*

separately for men and women. The results show that family relations modestly correlate negatively with all psychological states, with lower warmth and greater hostility associated with more negative affect. Several correlations in the IVF groups (both men and women) did not reach significance; however, the correlation coefficients, made separately for men and women, showed that the correlations for IVF vs. natural pregnancy groups did not differ significantly (see Fisher Z in Table 8). Since the Ns in the IVF groups were smaller, it is likely that the sample was underpowered to detect some of the weak associations.

Correlations Between Negative Psychological States of Spouses

Spearman's rank *correlation* was used to assess whether the psychological states of spouses were correlated. In the natural conception group, depression experienced by men positively correlated with women's depression ($r = 0.30$, $p = 0.00$), as well as with women's outward irritability ($r = 0.17$, $p = 0.00$) and anxiety ($r = 0.14$, $p = 0.02$). Men's depression also modestly correlated with their spouse's stress level in the second trimester ($r = -0.16$, $p = 0.01$) and third trimester ($r = -0.15$, $p = 0.01$); the negative sign of the correlation is due to low scores corresponding to high stress. Men's outward irritability positively correlated with all women's negative psychological states ($r = 0.17–0.19$; $p < 0.01$), with the exception of inward irritability. Men's inward irritability positively correlated with their partners' inward irritability ($r = 0.18$; $p = 0.00$), outward irritability ($r = 0.21$; $p = 0.00$), and anxiety ($r = 0.12$; $p = 0.04$). Men's anxiety positively correlated with all negative psychological states of the pregnant women and their stress level in the second and third trimester. The strongest correlations were observed between depression in both partners ($r = 0.30$; $p = 0.00$), between men's anxiety and women's depression ($r = 0.26$; $p = 0.00$), and between men's and women's anxiety ($r = 0.24$; $p = 0.00$).

In the couples with induced pregnancy, fewer significant correlations were observed, likely due to an underpowered sample. Significant positive correlations were observed between both partners' outward irritability ($r = 0.23$; $p = 0.03$), men's inward irritability and women's outward ($r = 0.25$; $p = 0.02$) and inward ($r = 0.25$; $p = 0.02$) irritability.

Correlations Between Partner-Reported and Perceived Partner's Warmth and Hostility

Analysis revealed moderate significant correlations between partner-reported and perceived warmth and hostility; as well as between partners' actual warmth and hostility in both groups (see Tables 9 and 10). Correlations were also observed between men's and women's perceived warmth and actual hostility and between the warmth of one partner and the hostility of the other.

Table 9

Correlations between partner-reported and perceived warmth/hostility in the natural conception group

		Woman's warmth towards man	Woman's hostility towards man	Woman's perceived warmth of her partner	Woman's perceived hostility of her partner
Man's warmth towards woman	R	0.47**	0.36**	0.58**	0.41**
	p	0.00	0.00	0.00	0.00
	N	271	272	268	269
Man's hostility towards woman	R	0.44**	0.58**	0.50**	0.66**
	p	0.00	0.00	0.00	0.00
	N	264	266	261	263
Man's perceived warmth of partner	R	0.58**	0.51**	0.60**	0.49**
	p	0.00	0.00	0.00	0.00
	N	274	275	271	272
Man's perceived hostility of partner	R	0.42**	0.68**	0.45**	0.57**
	p	0.00	0.00	0.00	0.00
	N	272	274	269	271

*Note. **Correlation significant at p = 0.00. Positive correlations between actual/perceived warmth and hostility indicate negative associations (hostility scores were reversed).*

Table 10

Correlations between partner-reported and perceived warmth/hostility in the IVF group

		Woman's warmth towards man	Woman's hostility towards man	Woman's perceived warmth of her partner	Woman's perceived hostility of her partner
Man's warmth towards woman	R	0.45**	0.29**	0.60**	0.48**
	P	0.00	0.01	0.00	0.00
	N	90	89	89	90
Man's hostility towards woman	R	0.40**	0.34**	0.53**	0.49**
	P	0.00	0.00	0.00	0.00
	N	89	88	88	89
Man's perceived warmth of partner	R	0.45**	0.41**	0.59**	0.54**
	P	0.00	0.00	0.00	0.00
	N	90	89	89	90
Man's perceived hostility of partner	R	0.43**	0.54**	0.56**	0.47**
	P	0.00	0.00	0.00	0.00
	N	90	89	89	90

*Note. **Correlation significant at p ≤ 0.01. Positive correlations between actual/perceived warmth and hostility indicate negative associations (hostility scores were reversed).*

Discussion

The first hypothesis of the study was not supported. Families with induced pregnancy did not experience greater stress and negative affect than families with natural pregnancies. The results showed relatively low stress for the majority of women, both with natural and induced pregnancy. Nevertheless, a small proportion of women in both groups experienced elevated levels of stress. There was some indication that stress levels were uneven across the trimesters, with a greater proportion of women in the IVF group showing elevated stress in the second trimester. The differences were not statistically significant, but the pattern of these results was consistent with previous literature on the prevalence of stress in infertile women (Hashemieh, Neisani Samani, & Taghinejad, 2013). Elevated stress in this group can be associated with complications during pregnancy or with women's fears based on information about the possible difficulties and risks of pregnancy (Crespo & Bestard, 2016).

Most women in both groups had scores in the normal range for depression, irritability, and anxiety. Nevertheless, in both groups a small proportion of women had elevated levels of these negative psychological states. Women with a natural pregnancy had greater outward irritability than women with an induced pregnancy. This may be due to the "desired baby" effect, which allows women who conceived with IVF to be more resilient to hormone-related irritability and emotional instability experienced during pregnancy. Previous research suggested that 100% of women undergoing IVF treatment view their pregnancy as desirable (and long awaited), whereas this proportion is smaller in women with a natural pregnancy (Naku et al., 2016).

Most men in both groups also had scores in the normal range for negative psychological states during their partners' pregnancy; however, there was a small proportion of participants with high levels of depression, irritability, and anxiety in both groups, which is in line with previous research (Darwin et al., 2017).

The second hypothesis of the study was supported. Experience of more negative psychological states (depression, anxiety, irritability) was associated with greater stress experienced by women during pregnancy. Modest to moderate correlations were observed between all negative psychological states and stress level in all trimesters of pregnancy in both groups. The results also indicated that the association between stress and negative affect may be particularly strong in the third trimester. The findings of associations between stress and negative affect are in line with research that found associations between distress and depression and other negative states (McLaughlin & Hatzenbuehler, 2009; Schneiderman, Ironson, & Siegel, 2005). Irritability and anxiety can be accompanied by an increase in stress during pregnancy; in turn, stress can contribute to irritability, anxiety, and depression. These findings highlight the importance of providing support and advice to women during the pregnancy on how to deal with negative emotional states.

In the group of women with IVF, outward irritability was accompanied by an increase in stress only during the first and second trimesters; in the second and third trimesters, stress was correlated with anxiety. Such dynamics in the IVF group can be explained by the fact that in the last trimesters of pregnancy, the focus of the spouses in the IVF group shifts to worrying about maintaining a healthy preg-

nancy, in light of the known elevated risk of miscarriage after IVF. Future research is needed to replicate the dynamics observed in this study. Comparison of the correlation coefficients showed that in the first trimester, correlation between stress and inward irritability was greater in the natural pregnancy group than in the IVF group. In the second trimester, correlation between stress and depression was significantly greater in the group of women with an induced pregnancy. In the third trimester, stress was correlated more strongly with anxiety, outward irritability, and inward irritability in the natural pregnancy group than in the IVF group.

The third hypothesis of the study was also supported. Modest to moderate associations between the relationship of spouses and their psychological states were observed. Comparisons of the correlation coefficients between psychological states and marital relations showed similar correlations in the IVF and natural pregnancy groups, with no significant differences. Lower warmth and greater hostility towards the partner was associated with higher levels of negative affect in both the IVF and natural pregnancy groups. Some associations in the IVF groups did not reach significance, which may be explained by the underpowered sample. These results are consistent with previous research that found associations between spousal relations and psychological states (Figueiredo et al., 2008; Tanner Stapleton et al., 2012).

The fourth hypothesis was also supported. The significant positive modest correlations between psychological states of pregnant women and their partners observed in our study indicate an emotional link and interdependency of some spouses during this time (Figueiredo et al., 2008).

Finally, *the fifth hypothesis* was also supported. The correlations between partner-reported and perceived warmth and hostility were moderate, ranging from .41 to .68 in the natural conception group, and from 0.41 to 0.6 in the IVF group. These results suggest that most people are overall accurate in their perceptions of their partner's warmth or hostility towards them. However, some tend to misinterpret each other's attitudes or have problems with evaluating or expressing their own attitudes.

Conclusion

The results suggest that psychological states, stress, and links between psychological states and quality of family relations are similar in families with IVF and natural pregnancies. The levels of negative states were relatively low in both groups. Higher stress was associated with worse emotional states of women during all trimesters of pregnancy. A moderate correlation was observed between spouse-reported warmth/hostility and perceived warmth/hostility. Negative psychological states were modestly related to the quality of family relations.

Limitations

The present study was based on an opportunistic sample, recruiting women through family-planning clinics; therefore, the participants in the two groups were not specifically matched on any socio-demographic parameters. However, all families came from four clinics in the same general area of Russia, and were therefore comparable. The sample is part of an ongoing longitudinal study, which is con-

tinuously growing. At the time of the current data analyses, the groups remained relatively small, and unequal in size, which limits the statistical power to find weak associations. However, the results point to overall similarities between the groups. Further longitudinal research is needed to explore the direction of causal links between psychological states of spouses, and between their psychological states and the quality of family relations.

Acknowledgment

This research was supported by "The Tomsk State University Competitiveness Improvement Program" grant (No. 8.1.11.2018).

References

Choi, H. & Marks, N.F. (2008). Marital conflict, depressive symptoms, and functional impairment. *Journal of Marriage and the Family, 70*(2), 377–390. https://doi.org/10.1111/j.1741-3737.2008.00488.x

Crespo, E., & Bestard, J. (2016). Psychosocial needs of women and their partners after successful assisted reproduction treatment in Barcelona. *Reproductive Biomedicine & Society Online, 3,* 90–99. https://doi.org/10.1016/j.rbms.2017.04.001

Darwin, Z., Galdas, P., Hinchliff, S., Littlewood, E., McMillan, D., & McGowan, L. (2017). Fathers' views and experiences of their own mental health during pregnancy and the first postnatal year: A qualitative interview study of men participating in the UK Born and Bred in Yorkshire (BaBY) cohort. *BMC Pregnancy and Childbirth, 17,* 45. https://doi.org/10.1186/s12884-017-1229-4

Du Rocher Schudlich, T. D., Papp, L. M., & Cummings, E. M. (2011). Relations between spouses' depressive symptoms and marital conflict: A longitudinal investigation of the role of conflict resolution styles. *Journal of Family Psychology, 25*(4), 531–540. https://doi.org/10.1037/a0024216

Ebbesen, S.M.S, Zachariae, R., Mehlsen, M.Y., Thomsen, D., Højgaard, A., Ottosen, L., Petersen, T., & Ingerslev, H.J. (2009). Stressful life events are associated with a poor in vitro fertilization (IVF) outcome: A prospective study. *Hum Reprod., 24,* 2173–2182. *https://doi.org/10.1093/humrep/dep185*

Fadeeva, N.I., Vostrikov, V.V., & Garganeeva, A.E. (2011). Emotsional'noe sostoianie i sotsial'nye osobennosti patsientok v rezul'tativnykh i nerezul'tativnykh tsiklakh ekstrakorporal'nogo oplodotvoreniia [Emotional state and social characteristics of patients with successful and unsuccessful IVF cycles. *Nevrologicheskii vestnik* [Journal of Neurology], *2,* 32–36. Retrieved from http://www.fesmu.ru/elib/Article.aspx?id=241453.

Faria, D. E. P., Grieco, S. C., & Barros, S. M. O. (2012). The effects of infertility on the spouses' relationship. *Revista da Escola de Enfermagem da USP, 46*(4), 794–801. https://doi.org/10.1590/S0080-62342012000400002

Feklicheva, I.V., Chipeeva, N.A., Voronina, I.D., Soldatova, E.L., Maslennikova, E.P., Shabalovskaya, M. ., Agarkova, L.A., Malykh, S.B., & Kovas, Yu.V. (2017). Vzaimosviaz' mezhdu otnosheniem k budushemu rebënku i otnosheniiami mezhdu roditel'iami i sem'iakh so spontannoi i indutsirovannoi bremennost'iu [Relationship between the attitude to an unborn child and the relations between parents in families with spontaneous and induced pregnancy]. *Akusherstvo i ginekologiia* [Obstetrics and Gynecology], *10,* 78–83. https://doi.org/10.18565/aig.2017.10.78-83

Figueiredo, B., Field, T., Diego, M., Hernandez Reif, M., Deeds, O., & Ascencio, A. (2008). Partner relationships during the transition to parenthood. *Journal of Reproductive and Infant Psychology, 26*(2), 99–107. https://doi.org/10.1080/02646830701873057

Filippova, G.G. (2009). Psikhologicheskaia adaptatsiia k beremennosti i roditel'stvu v supruzheskikh parakh s besplodiem i nevynashivaniem v amneze [Psychological adaptation to pregnancy and parenthood in married couples with infertility and miscarriage in anamnesis]. *Perinatal'naia psikhologiia i psikhologiia roditel'stva* [Perinatal Psychology and Psychology of Parenthood], *4,* 102–107.

Frederiksen, Y., Farver-Vestergaard, I., Skovgård, N.G., Ingerslev, H.J., & Zachariae, R. (2015). Efficacy of psychosocial interventions for psychological and pregnancy outcomes in infertile women and men: A systematic review and meta-analysis. *BMJ Open, 5*(1), e006592. https://doi.org/10.1136/bmjopen-2014-006592

Galhardo, A., Cunha, M., & Pinto-Gouveia, J. (2011). Psychological aspects in couples with infertility. *Sexologies, 20*(4), 224–228. https://doi.org/10.1016/j.sexol.2011.08.005

Gourounti, K., Anagnostopoulos, F., Potamianos, G., Lykeridou, K., Schmidt, L., & Vaslamatzis, G. (2012). Perception of control, coping and psychological stress of infertile women undergoing IVF. *Reprod Biomed Online, 24*(6), 670–679. https://doi.org/10.1016/j.rbmo.2012.03.002

Greil, A.L., Shreffler, K.M., Schmidt, L. & McQuillan, J. (2011). Variation in distress among women with infertility: Evidence from a population-based sample. *Human Reproduction, 26*(8), 2101–2112. https://doi.org/10.1093/humrep/der148

Greil, A.L., Slauson-Blevins, K., & McQuillan, J. (2010). The experience of infertility: A review of recent literature. *Sociology of Health & Illness, 32*(1), 140–162. https://doi.org/10.1111/j.1467-9566.2009.01213.x

Haimovici, F., Anderson, J.L., Bates, G.W., Racowsky, C., Ginsburg, E.S., Simovici, D., & Fichorova, R.N. (2018). Stress, anxiety, and depression of both partners in infertile couples are associated with cytokine levels and adverse IVF outcome. *Am J Reprod Immunol, 79* (4), e12832. https://doi.org/10.1111/aji.12832

Hashemieh, C., Neisani Samani, L., & Taghinejad, H. (2013). Assessment of anxiety in pregnancy following Assisted Reproductive Technology (ART) and associated infertility factors in women commencing treatment. *Iranian Red Crescent Medical Journal, 15*(12), e14465. https://doi.org/10.5812/ircmj.14465

Hjelmsted, A., Widström, A.M., Wramsby, H., Matthiesen, A.S., & Collins A. (2003). Personality factors and emotional responses to pregnancy among IVF couples in early pregnancy: A comparative study. *ActaObstetGynecolScand, 82*(2), 152–161.

Hjelmstedt, A., Widström, A.M., Wramsby, H., & Collins, A. (2003). Patterns of emotional responses to pregnancy, experience of pregnancy and attitudes to parenthood among IVF couples: A longitudinal study. *Journal of Psychosomatic Obstetrics & Gynecology, 24*(3), 153–162, https://doi.org/10.3109/01674820309039669

Hynes, G.J., Callan, V.J., Terry, D.J., & Gallois, C. (2011). The psychological well-being of infertile women after a failed IVF attempt: The effects of coping. *Br. J. Med. Psychol, 65*(3), 269–278.

Klemetti, R., Raitanen, J., Sihvo, S., Saarni, S., & Koponen, P. (2010). Infertility, mental disorders and well-being—a nationwide survey. *Acta Obstetricia et Gynecologica Scandinavica, 89,* 677–682. https://doi.org/10.3109/00016341003623746

Klock, S.C. & Greenfeld, D.A. (2000). Psychological status of in vitro fertilization patients during pregnancy: A longitudinal study. *Fertility and Sterility, 73*(6), 1159–1164. https://doi.org/10.1016/S0015-0282(00)00530-6

Kouros, C.D. & Cummings, E.M. (2011). Transactional relations between marital functioning and depressive symptoms. *American Journal of Orthopsychiatry, 81,* 128–138.

Kouros, C.D., Papp, L.M., & Cummings, E.M. (2008). Interrelations and moderators of longitudinal links between marital satisfaction and depressive symptoms among couples in established relationships. *Journal of Family Psychology, 22*(5), 667–677. https://doi.org/10.1037/0893-3200.22.5.667

Lin, Y.H. & Chueh, K.H. (2016). Somatic symptoms, sleep disturbance and psychological distress among women undergoing oocyte pick-up and in vitro fertilisation-embryo transfer. *J ClinNurs., 25*, 11–12, 1748–1756. https://doi.org/10.1111/jocn.13194

Maroufizadeh, S., Karimi, E., Vesali, S., & Omani Samani, R. (2015). Anxiety and depression after failure of assisted reproductive treatment among patients experiencing infertility. *International Journal of Gynecology & Obstetrics, 130,* 253–256. https://doi.org/10.1016/j.ijgo.2015.03.044

Massey, A.J., Campbell, B.K., Raine-Fenning, N., Pincott-Allen, C., Perry, J., & Vedhara, K. (2016). Relationship between hair and salivary cortisol and pregnancy in women undergoing IVF. *Psychoneuroendocrinology, 74,* 397–405. https://doi.org/10.1016/j.psyneuen.2016.08.027

McLaughlin, K.A., & Hatzenbuehler, M.L. (2009). Stressful life events, anxiety sensitivity, and internalizing symptoms in adolescents. *Journal of Abnormal Psychology, 118*(3), 659–669. https://doi.org/10.1037/a0016499

Melby, J., Conger, R., Book, R., Rueter, M., Lucy, L., Repinski, D., & Scaramella, L. (1998). *The Iowa Family Interaction Rating Scales* (5th ed.). Unpublished document, Iowa State University, Institute for Social and Behavioral Research.

Naku, E.A., Bohan, T.G., Ulyanich, A.L., Shabalovskaya, M.V., Tosto, M., Terehina, O.V., & Kovas, Yu.V. (2016). Psikhologicheskie kharakteristii zhenshchin, prokhodiashchikh lechenie po programme EKO [Psychological characteristics of women undergoing IVF treatment]. *Voprosy ginekologii, akusherstva i perinatologii* [Issues of Gynecology, Obstetrics and Perinatology], *15*(6), 23–30. https://doi.org/10.20953/1726-1678-2016-6-23-30

Naku, E.F., Kovas, Yu.V., Bohan, T.G., Terehina, O.V., & Vidyakina, T.A. (2017). Faktory negativnykh psikhoemotsional'nykh sostoianii zhenshchin, prokholdiashchikh lechenie besplodiia po programme EKO [Factors of negative psycho-emotional states among women undergoing IVF treatment]. *Sibirskii psikhologicheskii zhurnal* [Siberian Psychological Journal], *63,* 119–136. https://doi.org/10.17223/17267080/63/9

Pasch, L.A., Holley, S.R., Bleil, M. E., Shehab, D., Katz, P.P., & Adler, N.E. (2016). Addressing the needs of fertility treatment patients and their partners: Are they informed of and do they receive mental health services? *Fertility and Sterility, 106*(1), 209–215.e2. https://doi.org/10.1016/j.fertnstert.2016.03.006

Pellerone, M. & Micciché, S. (2014). Prenatal attachment and anxiety: Women who decide to try in vitro fertilization and women who procreate naturally. *Psychol. Res., 4*(6), 419–427. https://doi.org/10.17265/2159-5542/2014.06.002

Petrova, N.N., Podolhov, E. N., Gzgzyan, A.M., & Ngauri, D.A. (2013). Psikhicheskie rasstroistva i lichnostno-psikhologicheskie osobennosti u zhenshchin s besplodiem pri lechenii EKO [Mental disorders and personality-psychological features in women with infertility in treatment with IVF]. *Obzor psikhiatrii i meditsinskoi psikhologii* [Review of Psychiatry and Medical Psychology], *2,* 42–49.

Purewal, S., Chapman, S., & Van Den Akker, O. (2017). A systematic review and meta-analysis of psychological predictors of successful assisted reproductive technologies. *BMC Research Notes, 10*(711). https://doi.org/10.1186/s13104-017-3049-z

Rice, F., Harold, G.T., Boivin, J., van den Bree, M., Hay, D.F., & Thapar, A. (2010). The links between prenatal stress and offspring development and psychopathology: Disentangling environmental and inherited influences. *Psychological Medicine, 40*(2), 335–345. https://doi.org/10.1017/S0033291709005911

Sbaragli, C., Morgante, G., Goracci, A., Hofkens,T., DeLeo, V., & *Castrogiovanni, P.* (2008). Infertility and *psychiatric morbidity. Fertility and Sterility, 90,* 2107–2111. https://doi.org/10.1016/j.fertnstert.2007.10.045

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology, 1*, 607–628. https://doi.org/10.1146/annurev.clinpsy.1.102803.144141

Seok Kee, B., Jung, B.J., & Lee, S.H.J. (2000). A study on psychological strain in IVF patients. *Assist Reprod Genet, 17*(8), 445–448. https://doi.org/10.1023/A:1009417302758

Snaith, R.P., Constantopoulos, A.A., Jardine, M.Y., & McGuffin, P. (1978). A clinical scale for the self-assessment of irritability. *British Journal of Psychiatry, 132,* 163–171. https://doi.org/10.1192/bjp.132.2.164

Tanner Stapleton, L.R., Schetter, C.D., Westling, E., Rini, C., Glynn, L.M., Hobel, C.J., & Sandman, C.A. (2012). Perceived partner support in pregnancy predicts lower maternal and infant distress. *Journal of Family Psychology, 26*(3), 453–463. https://doi.org/10.1037/a0028332

Voronina, I.D., Bohan, T.G., Terehina, O.V., Malykh, S.B., & Kovas, Y. (2016). Demograficheskie pokazateli, obraz zhiznii i zdorov'e v sem'iakh s estestvennoi i indutsirovannoi bremennost'iu v Rossii i Velikobritanii [Demographics, lifestyle and health in families with natural and induced pregnancy in Russia and the UK]. *Teoreticheskaia i eksperimental'naia psikhologiia* [Theoretical and Experimental Psychology], *9*(4), 63–76.

Vujeva, H.M. & Furman, W. (2011). Depressive symptoms and romantic relationship qualities from adolescence through emerging adulthood: A longitudinal examination of influences. *Journal of Clinical Child and Adolescent Psychology, 40,* 123–135. https://doi.org/10.1080/15374416.2011.533414

Wichman, C.L., Ehlers, S.L., Wichman, S.E., Weaver, A.L., & Coddington, C. (2011). Comparison of multiple psychological distress measures between men and women preparing for in vitro fertilization. *Fertility and Sterility, 95,* 717–721. https://doi.org/10.1016/j.fertnstert.2010.09.043

Williams, K.E., Marsh, W.K., & Rasgon, N.L. (2007). Mood disorders and fertility in women: A critical review of the literature and implications for future research. *Human Reproduction Update, 13*(6), 607–616. https://doi.org/10.1093/humupd/dmm019

Zaharova, E.I. & Yakupova, V.A. (2015). Vnutrenniaia materinskaia pozitsiia zhenshchin, bremennost' kotorykh nastupila s pomoshch'iu EKO [Internal maternal position of women who became pregnant using IVF]. *Natsional'nyi psikhologicheskii zhurnal* [National Psychological Journal], *1*(17), 96–104. https://doi.org/10.11621/npj.2015.0111

Original manuscript received August 20, 2018
Revised manuscript accepted October 18, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

To Tell or Not to Tell: The Ethics and Law of Disclosing Health-Related Genetic Information to Family Members

Robert Chapman[a], Johanna Devereux[b], Victoria Nanau[a], Vanessa Smereczynska[a], Daria Matsepuro[c] & Fatos Selita[c*]

[a] *Department of Psychology, Goldsmiths, University of London, UK*

[b] *Guy's and St Thomas' NHS Foundation Trust, London, UK*

[c] *Tomsk State University, Tomsk, Russia*

* Corresponding author. Email: ftselita@gmail.com

Background. Genetic conditions and susceptibilities differ from other diseases and health-related risks. Genetic information is shared between blood relatives, and therefore a genetic finding can have implications for the wider family.

Objective. The present study investigates people's views on issues related to disclosing genetic information to relatives. Specifically, the study assesses opinions in relation to two issues: 1) whether people have a moral obligation to share their genetic data with family members; and 2) whether healthcare providers should have a legal obligation to share such data when consent is withheld.

Design. A public engagement event was held based on the real-life court case of *ABC vs the UK National Health Service (NHS)*. Participants were provided with information in three phases: first, about the case; then, with progressively more details of the case; and finally, with other relevant information. After being given each portion of information, the participants were asked to disclose their views on the rights and responsibilities related to the sharing of this information.

Results. The results clearly demonstrate that people hold strong and polarized views regarding confidentiality, and the moral and legal duties to disclose genetic information to family members. Even when withholding information could have an adverse impact on the health and life choices of relatives, participants disagreed about the legal obligations for healthcare providers to disclose a person's genetic information to those relatives.

Conclusion. The results suggest that the issues of privacy and disclosure of genetic information are complex and divisive.

***Keywords*:** Genetics, ethics, public engagement, patient confidentiality, duty of care, data access rights.

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

doi: 10.11621/pir.2018.0405
http://psychologyinrussia.com

Introduction

Advances in genetic research and technologies have led to genetic testing being adopted more routinely as part of healthcare treatment in many countries (Botkin et al., 2014). For example, in the UK, the Chief Medical Officer's 2016 report "Generation Genome" focused on the increasing importance of genomics in healthcare treatment and management (Davies, 2017). Most medical conditions have complex genetic etiologies in that they are influenced by many genetic factors (e.g. Balmain, Gray, & Ponder, 2003). However, some conditions follow relatively simple patterns of inheritance and expression.

Many countries have adopted new-born screening procedures to detect some of these conditions for early intervention and treatment. The number of disorders which are included in new-born screening varies in different countries, from two to more than 20 (Bodamer, Hoffmann, & Lindner, 2007; Kelly, Makarem, & Wasserstein, 2016). In the UK, new-born screening is performed for disorders with recognizable latent or early symptomatic stages, and for which prevention programs or treatments are already available.

Based on these criteria, UK screening includes nine congenital disorders: sickle cell disease, cystic fibrosis, congenital hypothyroidism, phenylketonuria (PKU), medium-chain acylCoA dehydrogenase deficiency (MCADD), maple syrup urine disease (MSUD), isovaleric acidaemia (IVA), glutaric aciduria type 1 (GA1), and homocystinuria (pyridoxine unresponsive) (HCU) (NHS, 2018). There are many other single gene disorders for which molecular (genetic) diagnoses are available, but screening for them is not offered population-wide. This is due to the lack of an effective treatment (aside from symptom management) and the rarity of these conditions in the general population without a previous family history of the disease. One such conditions is Huntington's disease (HD; see *Figure 1*).

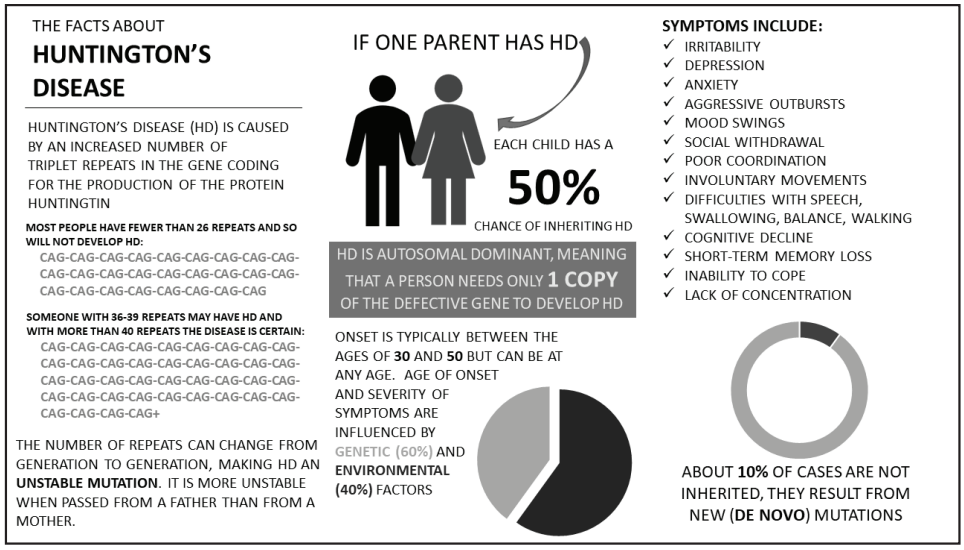


Figure 1. Information about Huntington's Disease

Availability of health-related genetic information to blood relatives

Genetic conditions and susceptibilities differ from other diseases and health-related risks because genetic information is shared among blood relatives. For example, knowing an individual's smoking status is not informative with regards to the health of his or her relatives. By contrast, anything one individual might learn about their personal genetic health risks indicates possible risks for family members, proportionate to the degree of genetic relatedness (e.g. 100% for identical twins; 50% for parents, children, or siblings; 25% for first-degree cousins, aunts and uncles, grand-parents, or half-siblings). This is particularly true for disorders with relatively simple genetic inheritance patterns, like HD (see *Figure 1*).

In the absence of effective treatments, individuals may find such information either useful or harmful. For example, when considering pre-symptomatic testing for HD, some feel that having a result is important for family planning decisions, while others are concerned about the physical and psychological impact of having such test results (Evers-Kiebooms et al., 2002; van der Steenstraten, Tibben, Roos, van de Kamp, & Niermeijer, 1994). This presents an ethical and legal dilemma regarding whether such information should be disclosed to family members and by whom.

For individuals this is not currently a legal dilemma, but an ethical one, since there is no law forbidding or forcing individuals to disclose genetic test results to their relatives. However, for medical organizations, the situation is becoming increasingly complicated in this genomic era. Oaths of conduct and ethics are taken by healthcare professionals all over the world (Crawshaw, 1994; Green, 2017; Hulkower, 2016), many of which are based on the Hippocratic oath of the 3rd century CE (Loudon, 1994). The following statements are included in the modern Hippocratic oath:

"I will respect the privacy of my patients, for their problems are not disclosed to me that the world may know," and "I will prevent disease whenever I can, for prevention is preferable to cure."

These two obligations may be difficult to reconcile in the context of genetic information. Consider this example. A woman undergoes a genetic test, which uncovers a BRCA1 gene mutation, which is known to significantly increase the risk of developing breast and ovarian cancer. She opts to undergo a prophylactic double mastectomy surgery and the removal of her ovaries to reduce her risk of developing cancer (Hartmann et al., 2001). The BRCA1 gene is not routinely screened for during healthcare evaluations, but is tested for by women with a known family risk or specific tumor histologies. Therefore, information about the mutation could be important for the woman's sister, who, as a first degree relative, has a 50% risk of carrying the same mutation.

The woman does not want to disclose the result of her test to her sister. Should the woman's doctor/medical provider respect her confidentiality, in line with the "respect privacy" part of the Hippocratic oath? Or does the provider have an ethical and/or legal obligation to inform the sister, in line with the "prevent disease" aspect of the oath? In other words, should the woman's confidentiality take precedence over the health risk to the sister, or the other way around?

These questions are no longer hypothetical and have already reached the courts. For example, a case of non-disclosure of HD is currently being decided in the UK courts (*ABC v St George's Healthcare NHS Foundation Trust [2017] EWCA Civ 336', 2017)*. In this case, action has been brought against the UK National Health Service (NHS) by a patient on the basis that the NHS owed a duty of care to disclose her father's HD diagnosis to her. The daughter asserts that, had she been informed of the risk, and her own diagnosis was confirmed, she would have terminated her pregnancy. The High Court struck her claim on the grounds that there was no reasonably arguable duty of care owed to the daughter by the NHS. The decision of the High Court was appealed by the daughter, and the Court of Appeal reversed the High Court's decision, remitting the case for trial. The differing decisions of the UK courts reflect the complexity of the moral, ethical, and legal issues in such cases.

The increasing use of genetics in medicine, coupled with the relatedness between family members, means that the concepts of medical ethics, consent, and confidentiality need to be revisited. Any such re-evaluation of established social norms should include broad and thorough public discussion. This is especially important when revisions of the social norms could fundamentally change how we conceptualize the "self."

Many ethical and moral dilemmas are difficult to resolve because people hold strong and polarized views on them. These include the issues of abortion, capital punishment, gun ownership, and the age of consent, for which laws also differ across legal systems. Societies now face an additional dilemma: whether or not to disclosue genetic information to family members. People's views on this issue have not been adequately explored.

The present study investigates public opinions about who should have a right to genetic information, and whether this information can be disclosed to relatives without the consent of the person to whom that information relates. In particular, we explore people's views on two questions:

1. Do people have a moral obligation to share their genetic data with family members?
2. Should healthcare providers have a legal obligation to share such data when consent is withheld?

Objectives

The main aim of our study was to establish whether there is a relative consensus regarding the disclosure of genetic information to family members, or whether opinions relating to this issue are polarized in society. This information is important for updating regulations and developing guidelines for the use of genetic data.

Method

Participants

Data were collected from individuals attending a public science engagement event run by The Accessible Genetics Consortium (www.tagc.world) at a university in London in 2018. The event was part of the Genes & Tonic series, which is dedicated

to promoting genetic knowledge and its application to different areas, such as education, medicine, and justice. The event was advertised to the public through social media, websites of the university and partner organizations, posters, and word of mouth.

During the event, participants heard talks by a geneticist, a lawyer, and a genetic counsellor, and engaged in activities during which data were collected. Refreshments were provided twice during the event. Data were recorded for 35 participants (22 female) of a median age of 31.88 (SD =14.1, range 18 to 80). Fourteen participants had only secondary school education, 10 had completed undergraduate university degrees, and 11 had completed postgraduate studies.

All participants provided written consent for the use of their data in this research. The research was approved by the Goldsmiths University Psychology Department Ethics Committee.

Measures and Procedure

Upon arrival, participants were provided with a welcome packet and a unique identifier number. This packet contained forms for participants to provide responses to specific questions asked throughout the event, as well as questions to capture demographic information: age, sex, occupation, and highest education level achieved.

Throughout the event, participants were provided with details drawn from the above-mentioned court case (*ABC v St George's Healthcare NHS Foundation Trust [2017] EWCA Civ 336', 2017)* concerning HD.

Participants were asked to give their opinions three times during the event, as progressively more details about the case, as well as other information, were released to them in successive waves.

At the outset of the evening, participants were provided with background information about the symptoms, progression, and prognosis for HD, and current treatment options. They were also presented with the following overview of the case notes (people's names were altered for the event):

- 2007: Having shot and killed his wife, a man (Fred) was sentenced to a hospital order and a restiction order (related to mental illness). In this connection, his adult children attended family therapy at the same hospital.
- January 2009: Fred's doctors first suspected that he might have HD. They urged him to tell his family; he informed his brother but refused to tell his three daughters.
- 09/02/2009: a diagnosis of HD was confirmed through genetic testing.
- December 2009–January 2010: Healthcare professionals repeatedly urged Fred to disclose his diagnosis to his daughters. Fred withheld consent.
- 08/23/2010: One of his daughters (Claire) was accidentally told by Fred's doctor that her father was diagnosed with HD.
- Late 2010: Claire began the process of suing the NHS for not providing this information officially at the time of diagnosis.

Participants at the event were also advised of the familial risk to Claire and her sisters, i.e. that they each had a 50% risk of developing HD. The fact that HD can impair cognitive function, and that this might have had an impact on the father's

ability to understand his and his family's situation, was not emphasized to the participants, although some did comment on this in their feedback. Participants were then asked to give their opinions on a 7-point scale (1=not at all to 7=definitely) on the following three statements:

1) The patient (daughter) had a *right* to know about her father's diagnosis.
2) The National Health Service (NHS) should have been *legally* obliged to provide this information to the daughters.
3) The father had a *moral* responsibility to provide this information.

Participants were also invited to provide written feedback and comments during each wave of data collection.

The second wave of data collection was preceded by the additional information that the daughter (Claire) was pregnant at the time her father's diagnosis was confirmed. The participants were also informed that she attested that, if she had known of her father's diagnosis, she would have terminated her pregnancy, given her own risk of developing HD and the risk to her unborn child. The participants' opinions were collected again as described above.

The third wave of data collection was preceded by additional information. This time participants were given a hypothetical scenario that a cure for HD had been discovered, but was only effective if begun before symptoms appeared. In this scenario, genetic testing for HD by the NHS was still only available to patients exhibiting symptoms, or to those with a known family history of the disorder. As such, the daughter would only have been able to access the cure if she knew about her father's diagnosis. The participants' opinions were collected for the third time as described previously. It was made clear to them that the scenario was hypothetical (no cure for HD currently exists), but that this is an active area of research.

Participants were asked to record their unique ID on each form they submitted during the evening. Unfortunately, not all participants provided their unique ID on each form they submitted, which led to missing data on some aspects of the analyses (those involving all three waves).

Participants were also asked whether they had ever had a genetic test, if they knew anyone with a genetic condition, or if they had such a condition themselves. They were also asked how influential religion was in informing their opinions and decisions (not at all, somewhat, or greatly influenced) and how confident they were in their genetic knowledge on a scale of 0 (none) to 100 (entirely confident).

Results

The average participant's confidence in genetic knowledge was relatively low: 36.91 (SD=26.27; range 0 to 85). Regarding the influence of religion on their opinions, 31 participants (88.6%) reported not being influenced at all; three participants (8.6%) reported being influenced to some extent; and one participant (2.9%) reported strong influence.

As can be seen in *Figure 2*, there was a general consensus that the father had a moral obligation to provide information to his daughter about his diagnosis. However, opinions were more divided when it came to the daughter's right to such information. Participants' responses were even more polarized when they were asked

whether the NHS should be legally obliged to disclose genetic information when consent has been withheld. Even following the final wave of information, when participants had been advised that the daughter was pregnant, and that, hypothetically, there was a cure for HD, 22% of participants still felt that there should be no legal obligation placed on the NHS to disclose the father's diagnosis to the daughter.

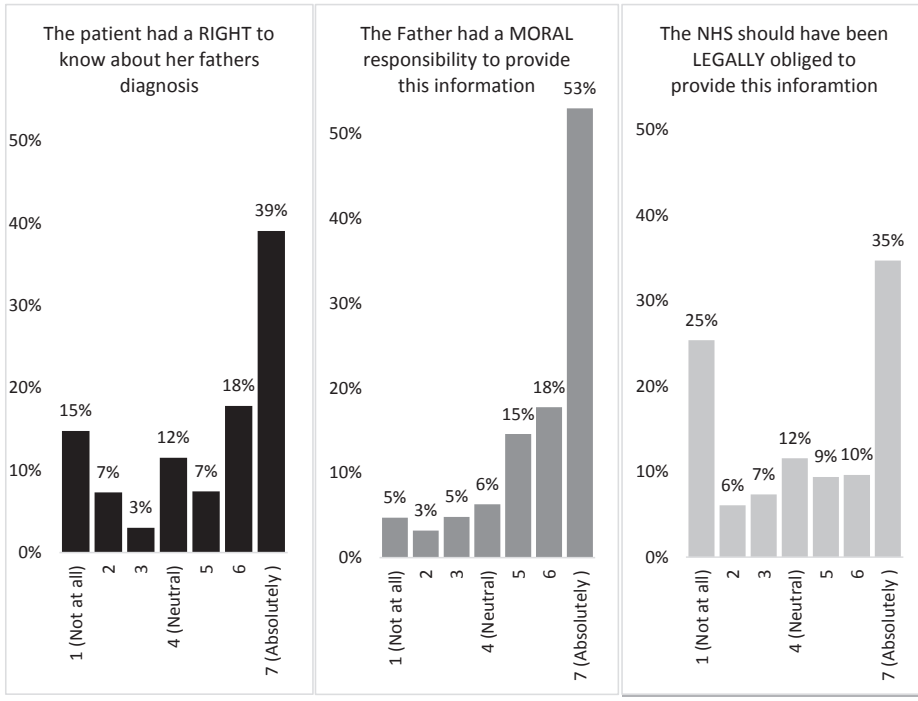


Figure 2. Summed percentage (across the 3 waves) of participants' responses to the 3 statements. Note: Percentage of responses (rather than participant numbers) are reported

As can be seen in *Figure 3*, participants' opinions remained relatively stable throughout the questioning, with only small increases for each statement across the three waves of data collection.

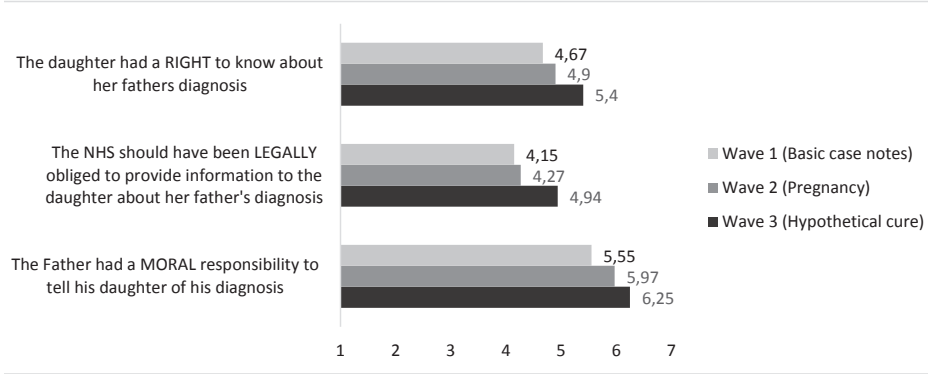


Figure 3. Average score on a scale of 1–7 represented for each wave and each question

Some potential group differences emerged, although the sample was underpowered to test these statistically. Men tended to be more inclined toward mandating disclosure of the genetic information than women, particularly for the first two questions (*Figure 4*). Participants who had a genetic condition, either themselves or in their family, were also more inclined towards disclosure (*Figure 5*). As only four participants (11.5% of the sample) stated that their opinions were influenced by religion to some degree, group analyses of religion are not presented.

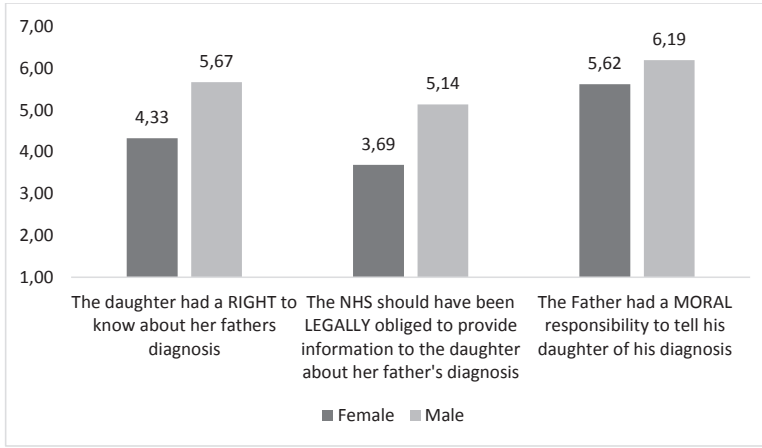


Figure 4. Average score on a scale of 1–7 for men (N = 9) and women (N = 13) for each question.

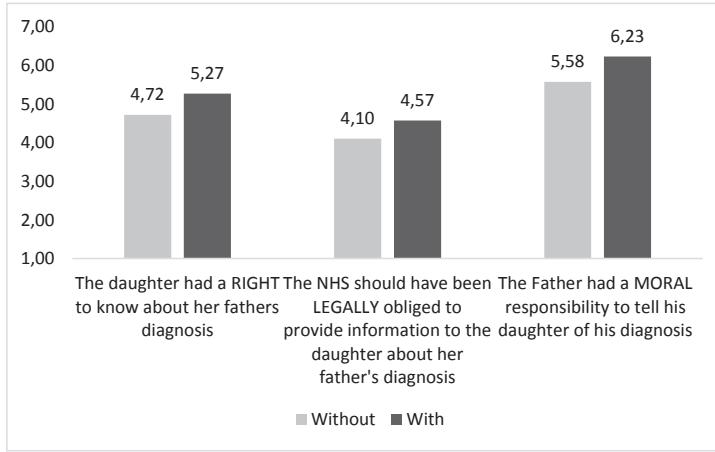


Figure 5. Average score on a scale of 1–7 for participants without (N = 17) and with (N = 5) a genetic condition, either themselves or in the family.

Qualitative analysis

The unscripted responses provided by participants clearly demonstrate their strong and polarized views. For example, when in favor of a patient's privacy, participants made such statements as: "If your DNA isn't your own, what is?"; "It remains Fred's

right alone. Regardless of consequences."; "The NHS were refused permission by the father, a violation of this goes against doctor-patient confidentiality." When in favor of disclosure of the information to the relatives, participants said, for example: "Any information to do with genetics like this must be shared."; "Families need careful genetic counselling to deal with Huntington's. It leads to early death therefore families need to know because of their children." "Fred is now responsible for 2 lives, so is under a lot of moral obligation. It's the woman's choice if she wants to terminate, not Fred. #Prochoice." The full collection of responses is available from the authors upon request.

Conclusion

The results of this study demonstrate that people hold strong and polarized views on the issue of confidentiality, and the moral and legal duty to disclose genetic information to family members. In particular, participants disagreed about the legal obligations on healthcare providers to disclose a person's genetic information to relatives, even when withholding information could have adverse impacts on the health, wellbeing, and life choices of those relatives.

Although the study sample size was small, it captured a wide range of ages, professions, and educational backgrounds. Conversely, the sample was also homogeneous in that all participants were interested in genetics, and had intentionally attended a genetic science engagement event. Over half the participants were students, and 60% had completed degree-level studies, indicating high levels of educational attainment within the sample. The fact that such a diversity of views is present within this sample suggests that the issue of privacy and disclosure of genetic information is complex and divisive.

The results also showed that exposure to the same information, including expert talks on genetics, law, and genetic counselling, did not lead to significantly increased similarity in participants' views. A bigger and more representative study is needed to further explore demographic and other factors that may influence people's views on these matters. For example, the results indicated that having a genetic condition in the family may lead to viewing disclosure of genetic information to family members more favorably.

The case presented in this study is relatively simple, as there is a single known genetic cause for HD. As discussed in the introduction, most diseases have much more complex etiologies, with a mixture of genetic and environmental factors potentially contributing. This etiological complexity makes risk estimates much harder, since genetic information is highly probabilistic. With this increased uncertainty, deciding on ethical and legal responsibility becomes even more complicated. Nevertheless, the need for clarity on these matters is urgent.

However, research shows that societies are unprepared for making informed decisions on updates to laws and policies. For example, a recent study revealed poor genetic literacy among the public when people were presented with multiple-choice questions about simple genetic concepts (Chapman et al., 2018). A lack of knowledge, combined with strong polarized opinions, is a worrying position for society to be in. The present research highlights the need for improved genetic education, but also suggests that simply knowing the facts will not resolve the di-

lemmas in this area. Further research is needed into the reasons for the strong polarized opinions as a step toward finding optimal solutions. Public engagement will be fundamental to these endeavors.

Limitations

The data presented here only represent the views of a small number of participants who were already engaged enough with genetics and genomics to attend the event. The numbers were insufficient to allow for meaningful inferential analyses, and this should be addressed in future studies. However, the findings that opinions are polarized and somewhat stable certainly does warrant further investigation.

Acknowledgements

We are grateful to all the participants for their contribution and to the members of InLab (www.inlab.co.uk) for their support with planning and running the "Genes & Tonic: Know Thy Father" event. This work was supported by the Russian Ministry for Education and Science (project No. 25.8562.2017/9.10).

References

ABC v St George's Healthcare NHS Foundation Trust [2017] EWCA Civ 336: Does a doctor owe a duty of care to disclose a hereditary disease to a patient's child? (2017, May 23). Retrieved from https://clinicalnegligence.blog/2017/05/23/abc-v-st-georges-healthcare-nhs-foundation-trust-2017-ewca-civ-336-does-a-doctor-owe-a-duty-of-care-to-disclose-a-hereditary-disease-to-a-patients-child/

Balmain, A., Gray, J., & Ponder, B. (2003). The genetics and genomics of cancer. *Nature Genetics*, *33*(3s), 238–244. https://doi.org/10.1038/ng1107

Bodamer, O.A., Hoffmann, G.F., & Lindner, M. (2007). Expanded newborn screening in Europe 2007. *Journal of Inherited Metabolic Disease*, *30*(4), 439–444. https://doi.org/10.1007/s10545-007-0666-z

Botkin, J.R., Lewis, M.H., Watson, M.S., Swoboda, K.J., Anderson, R., Berry, S. A., … & Wilfond, B. (2014). Parental Permission for Pilot Newborn Screening Research: Guidelines from the NBSTRN. *Pediatrics*, *133*(2), e410–e417. https://doi.org/10.1542/peds.2013-2271

Chapman, R., Likhanov, M., Selita, F., Zakharov, I., Smith-Woolley, E., & Kovas, Y. (2018). New literacy challenge for the twenty-first century: genetic knowledge is poor even among the well-educated. *Journal of Community Genetics*, 1–12. https://doi.org/10.1007/s12687-018-0363-7

Crawshaw, R. (1994). The Hippocratic oath. Is alive and well in North America. *BMJ : British Medical Journal*, *309*(6959), 952.

Davies, S.C. (2017). Annual report of the Chief Medical Officer 2016: generation genome. UK Government. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/631043/CMO_annual_report_generation_genome.pdf

Evers-Kiebooms, G., Nys, K., Harper, P., Zoeteweij, M., Dürr, A., Jacopini, G., …& Simpson, S. (2002). Predictive DNA-testing for Huntington's disease and reproductive decision making: a European collaborative study. *European Journal of Human Genetics*, *10*(3), 167–176. https://doi.org/10.1038/sj.ejhg.5200781

Green, B. (2017). Use of the Hippocratic or other professional oaths in UK medical schools in 2017: practice, perception of benefit and principlism. *BMC Research Notes*, *10*. https://doi.org/10.1186/s13104-017-3114-7

Hartmann, L.C., Sellers, T.A., Schaid, D.J., Frank, T.S., Soderberg, C.L., Sitta, D.L., …& Jenkins, R B. (2001). Efficacy of Bilateral Prophylactic Mastectomy in BRCA1 and BRCA2 Gene Mutation Carriers. *JNCI: Journal of the National Cancer Institute*, *93*(21), 1633–1637. https://doi.org/10.1093/jnci/93.21.1633

Hulkower, R. (2016). The History of the Hippocratic Oath: Outdated, Inauthentic, and Yet Still Relevant. *Einstein Journal of Biology and Medicine*, *25*(1), 41–44. https://doi.org/10.23861/EJBM20102542

Kelly, N., Makarem, D.C., & Wasserstein, M.P. (2016). Screening of Newborns for Disorders with High Benefit-Risk Ratios Should Be Mandatory. *The Journal of Law, Medicine & Ethics*, *44*(2), 231–240. https://doi.org/10.1177/1073110516654133

Loudon, I. (1994). The Hippocratic oath. *BMJ (Clinical Research Ed.)*, *309*(6951), 414.

NHS. (2018). Newborn blood spot test - NHS.UK. Retrieved 21 May 2018, from https://www.nhs.uk/conditions/pregnancy-and-baby/newborn-blood-spot-test/

van der Steenstraten, I.M., Tibben, A., Roos, R.A., van de Kamp, J.J., & Niermeijer, M.F. (1994). Predictive testing for Huntington disease: nonparticipants compared with participants in the Dutch program. *American Journal of Human Genetics*, *55*(4), 618–625.

Original manuscript received July 12, 2018
Revised manuscript accepted September 30, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

The Role of Mathematical and Trait Anxiety in Mental Fatigue: an EEG Investigation

Ekaterina S. Zhban*[a], Maxim V. Likhanov[b], Ilya M. Zakharov[c], Elizaveta M. Bezrukova[a], Sergey B. Malykh[c]

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

[b] *Educational Fund "Talent and success", Centre "Sirius", Sochi, Russia*

[c] *Psychological Institute, Russian Academy of Education, Russia*

* Corresponding author. E-mail: z.katherina@gmail.com

Background. Mental fatigue is a state of tiredness, decreased motivation, and increased aversion to performing a task. Mental fatigue is associated with the length of engagement in an activity (time-on-task) and the degree of cognitive effort required. In addition, mental fatigue can be affected by personality characteristics, such as trait or domain-specific anxiety. There is a lack of research into associations between mental fatigue and trait anxiety, as well as specific types of anxiety such as math anxiety.

Objective. This study investigates whether the level of mental fatigue manifested in an EEG taken during the performance of a mixed problem-solving task, is associated with math and trait anxiety.

Design. An EEG recording was performed on participants in a resting state with their eyes closed in two runs, both before and after they performed a task. The task consisted of three types of stimuli: arithmetic, algebraic, and lexical.

Results. The results showed that the EEG correlates of fatigue changed between the first and second runs. These changes were not linked with mathematics anxiety. Some significant EEG effects were found for trait anxiety: people with high trait anxiety appeared more aroused and showed less fatigue effects. However, these results did not reach the level of significance after correction for multiple comparisons.

Conclusion. Overall, our results are in line with the motivational control theory, according to which mental fatigue "resets" when a person switches from one task to another. In our study, the experimental paradigm consisted of three types of tasks, a format which might have prevented fatigue. We discuss the implications of the study for further research into the links between anxiety and mental fatigue.

Keywords: Mathematics anxiety (MA), trait anxiety (TA), EEG, mental fatigue

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

doi: 10.11621/pir.2018.0406
http://psychologyinrussia.com

Introduction

Mental fatigue is a state that generally manifests itself as tiredness, decreased motivation, and aversion to perform a task (Boksem & Tops, 2008; Matthews, Desmond, & Hitchcock, 2012). It develops as a function of the time-on-task and depends on the amount of cognitive effort required (Boksem, Meijman, & Lorist, 2005; Lorist et al., 2000). Mental fatigue can be operationalized by various methods, including neuroimaging. Craig and colleagues (2012) reviewed 17 experimental papers where mental fatigue was measured by changes in the spectral power of different EEG bands. According to this review, theta, beta, and alpha band activity and their ratios are likely to increase with the increase of mental fatigue.

The relationships between mental fatigue, time-on-task, and cognitive effort are not always linear. Studies show that people differ in their vulnerability to mental fatigue: Some people can continue their work for a long time without any effects on their performance, while others start showing the signs of fatigue very early (Matthews et al., 2012). Fatigue is frequently connected with sleep loss, circadian rhythms, and health issues. However, it is also linked to personality traits, such as introversion, neuroticism, sensitivity to punishment, behavioral inhibition, and anxiety (Matthews et al., 2012). Here we focus on the association between mental fatigue and two types of anxiety–trait and math anxiety.

A link between mental fatigue and trait anxiety (TA) was previously reported by several studies; it was shown to range from 0.28 to 0.59 (Craig, Tran, Wijesuriya, & Boord, 2006; Jiang et al., 2003; Wijesuriya, Tran, & Craig, 2007). The classical view of mental fatigue explains the phenomenon as "a lack of sufficient steady state energy to power … cognitive work" (Hancock, Desmond, & Matthews, 2012, p. 67). If this is the case, then this link between mental fatigue and trait anxiety can be explained by an increased need for resources by people with high anxiety, which is suggested by the Processing Efficiency Theory (PET; Eysenck & Calvo, 1992) and Attentional Control Theory (ACT; Eysenck, Derakshan, Santos, & Calvo, 2007). These two theories are generally used to interpret the mechanisms of anxiety's effects on performance.

The PET assumes that anxiety reduces the working memory capacity available for task processing, as engagement in an anxiety-related task causes people to experience stress and intrusive thoughts, and to ruminate on these intrusive thoughts. The ACT extends PET, and postulates that anxiety in general affects the ability to allocate attentional and cognitive resources to task performance. Due to an impaired inhibitory function, anxious people become more distracted by task-irrelevant stimuli, regardless of whether these stimuli are external (i.e., conventional distractors) or internal (i.e., worrying thoughts, ruminations, etc.).

In summary, researchers concur that people with high trait anxiety not only need resources for task processing, but also resources to deal with anxiety-induced thoughts and to inhibit distracting stimuli. This assumption is confirmed by multiple studies which showed that people with trait anxiety perform more poorly than non-anxious people under stressful conditions, especially in ego-threatening conditions, such as when they are being observed by other people (Edwards, Edwards, & Lyvers, 2017; Rezaei, Hosseini Ramaghani, & Fazio, 2017). Based on the classic interpretation of mental fatigue, this double requirement for resources may make anxious people more vulnerable to mental fatigue.

The PET and ACT framework has also been applied to maths anxiety (MA) (Carretié, Mercado, Tapia, & Hinojosa, 2001; Núñez-Peña & Suárez-Pellicioni, 2014; Wang, Huang, Ma, & Li, 2012). Maths anxiety is defined as "a feeling of tension and anxiety that interferes with the manipulation of numbers and solving of mathematical problems in … ordinary life and academic situations" (Richardson & Suinn, 1972). MA has been shown to play a role in underperformance in mathematics (Artemenko, Daroczy, & Nuerk, 2015; Suárez-Pellicioni, Núñez-Peña, & Colomé, 2016). Research has shown that MA correlates with TA ($r = 0.35$). However, MA cannot be viewed as an expression of TA, since different aspects of MA correlate more highly with one another (0.5–0.8) than with TA (Dowker, Sarkar, & Looi, 2016).

The present study employs the EEG method to explore links between TA, MA, and mental fatigue. Mental fatigue is measured as an increase in theta, alpha, and beta band activity during rest conditions in the course of a cognitive task (Craig, Tran, Wijesuriya, & Nguyen, 2012). We chose a resting state condition because previous studies have demonstrated that electrical brain activity during a resting state might be related to psychological states and emotional regulation processes (Knyazev, 2007; Putman, 2011). Based on previous research, we hypothesized that mental fatigue would positively correlate with TA. We also hypothesised a positive correlation between mental fatigue and MA, as similar mechanisms may be involved in the association between fatigue and all types of anxiety.

Methods

Objective

The aim of the present study is to investigate whether the level of mental fatigue, manifested in EEG activity, is associated with math and trait anxiety.

Participants

All participants were tested at one of the two research centers in Russia. Participants were recruited via social media, as part of an ongoing twin study (Kovas et al., 2013). Two hundred twenty-five participants, including 61 singletons and 82 pairs of twins, took part in the study. Only one twin from each pair was randomly selected, in order to account for twin non-independence (Rijsdijk & Sham, 2002). After the exclusion of participants with missing EEG recordings, the final sample (singletons and one twin from a pair) included 101 participants with mean age of 21.62 ($SD = 4.28$; 33 males, 68 females). All participants were native Russian speakers, and had normal or corrected-to-normal vision and no history of head injury, neurological or psychiatric disorder. Participants gave their informed written consent prior to the experiment. The study was approved by the Ethical Committee for Interdisciplinary Investigations (Tomsk State University).

Design and experimental procedure

Demographic, psychological, and physiological (EEG) data were collected. Participants were requested to fill in a demographic questionnaire that included information about their age and sex. They then completed psychological inventories and the EEG experiment.

Math anxiety. Math anxiety was measured with the Shortened Math Anxiety Rating Scale (sMARS; Alexander & Martray, 1989). This test measures math anxiety by presenting participants with 25 math-related situations. Participants rated, on a 5-point Likert scale, how much anxiety they experienced in each particular situation. A 21-item Russian adaptation of sMARS was used, which has four items removed from the questionnaire based on their poor psychometric properties (see Likhanov et al., 2017). As in previous research (Ashcraft & Moore, 2009), the total score of all items (21 in this study) was used as a singular measure of math anxiety.

Trait anxiety. The level of trait anxiety was measured with a Russian adaptation of the Spielberger Trait Anxiety Questionnaire (Spielberger, 1989; Hanin, 1983). The questionnaire consists of 20 questions on how anxious participants feel in their ordinary life on a 4-point Likert scale. The trait anxiety score was calculated as a simple sum of scores, without applying any coefficients.

EEG recording. Participants were instructed to sit comfortably and try not to move their eyes. The recording was made in a dimly lit, soundproof, and electromagnetically shielded room. The resting state data were recorded two times: before and after the main EEG data collection session, which on average lasted for 90 minutes. Each resting state recording run included five two-minute series with closed and open eyes (10 minutes), which were recorded in the following order: eyes closed — eyes open — eyes closed — eyes open — eyes closed; four minutes with open and six minutes with closed eyes in total. Only the data in the eyes-closed condition was used for further analysis. The resting state after the main session included four two-minute series: eyes closed — eyes open — eyes closed — eyes open (eight minutes in total).

During the main session, participants were to complete three types of problem-solving tasks: lexical, algebraic, and arithmetic. There were 70 trials of each task type, with 35 correct and 35 incorrect trials in each task, all of which were completely randomized (see Bloniewski, Likhanov et al., 2018).

The data were collected with the BrainVision PyCorder software (Brain Products). The EEG data were recorded from the scalp using 64 channels (63 EEG + VEOG) via Ag/AgCl electrodes placed according to the international 10-20 system. The EEG data were recorded without online filtering, with the sampling rate of 500 Hz. The Cz electrode was used as a reference, and the FPz was used as a ground. The data were then preprocessed offline using BrainVision Analyzer 2. The recorded data were downsampled to 256 Hz and filtered with a 0.1 -40 Hz filter.

The data were then re-referenced to an averaged reference and manually cleaned from artifacts, with noisy channels excluded. To remove blink and vertical eye-movement artifacts, independent component analysis (ICA) was performed on the following electrodes: VEOG — AFz, HEOG — FT9, and FT10. After ICA, the excluded channels were topographically interpolated, and semiautomatic artifact rejection was conducted. Fast Fourier transformation (FFT) was applied to 2s artifact-free epochs at every channel to extract spectral power values of the signal

in selected frequency bands (alpha, theta, and beta). The FFT results were averaged across at least 150 epochs for each channel.

In this study the changes in spectral power were investigated at seven electrodes: Fz, F3, F4, F7, F8, Cz and Pz. The rationale was that the effects of mental fatigue tend to gravitate towards midline and frontal electrodes in theta rhythm (Strijkstra, Beersma, Drayer, Halbesma, & Daan, 2003; Wascher et al., 2014); frontocentral and parietal regions for alpha (Cantero, Atienza, Salas, & Gómez, 1999; Oken, Salinsky, & Elsas, 2006; Santamaria & Chiappa, 1987), and frontal regions for the beta spectral band (Craig et al., 2012).

Mental fatigue measures. The following EEG bands were used: alpha (8-12 Hz), beta (12-30 Hz), and theta (4-7 Hz). We estimated absolute spectral power in selected frequency bands. Relative power was also used in the present study as a more reliable measure of spectral power (Nuwer, 1988) and was computed as follows:

1. relative alpha spectral power = alpha spectral power /(alpha spectral power + theta spectral power + beta spectral power)
2. relative theta spectral power = theta spectral power /(alpha spectral power + theta spectral power + beta spectral power)
3. relative beta spectral power = beta spectral power /(alpha spectral power + theta spectral power + beta spectral power)

An increase in theta, alpha, and beta band activity indicates an increase in mental fatigue (Cheng & Hsu, 2011; Craig et al., 2012; Eoh, Chung, & Kim, 2005).

We also used two composite measures, which have been shown to correlate with fatigue. First, we used the Task Load Index (TLI), which reflects the degree of neural activation during effortful task performance (Gevins & Smith, 2003). TLI is calculated as the ratio of theta activity at frontal midline sites, to alpha at parietal sites (here as theta at Fz to alpha at Pz). It is expected to increase with the development of mental fatigue (Cheng & Hsu, 2011; Kamzanova, Matthews, Kustubayeva, & Jakupov, 2011). Second, we used the Relative Energy Index (REI), which has been shown to correlate positively with an increase in mental fatigue, as fatigue is expected to increase with a decrease of energy (Cheng & Hsu, 2011; Eoh et al., 2005). REI is calculated as ratio of (alpha + theta) to beta activity.

To reduce the number of statistical comparisons, in this study mental fatigue was estimated as the difference in spectral power between run one and run two — the difference score. The difference score was estimated for each frequency band and the two coefficients (TLI and REI). The difference score was also used in a correlational analysis, to explore the link between mental fatigue and anxiety.

Mental fatigue may be related to a participant's state, e.g. how fatigued one already is prior to the experiment. Therefore, we calculated partial correlations between math and trait anxiety and the difference score for all parameters at all electrodes, controlling for EEG spectral power in run one. Partial correlations were computed between the two types of anxieties and both absolute and relative power values. The false-discovery rate (FDR) was applied as a correction for multiple comparisons.

Results

First, we tested whether the changes in mean spectral power for all ranges and spectral ratios, which occurred between run one and run two in the regions of interest, were associated with mental fatigue. We used a non-parametric Wilcoxon signed-rank test as an assumption of the normality of the mean difference between spectral power in alpha, theta, and beta ranges, and spectral ratios in run one and run two were violated for most of the electrodes. The analysis showed that absolute band power increased significantly at all electrodes in all bands (see *Table 1*). The relative band power before and after performing the task was significantly different for all electrodes for alpha (decreased) and theta (increased) bands, but not for the beta band (see *Table 2*). The spectral ratios (TLI and REI) did not change significantly (see *Table 3*).

Table 1

Descriptive statistics and Wilcoxon Signed Rank Tests for absolute activity

Power band	Electrode	Mean (SD) for run 1	Mean (SD) for run 2	Mean (SD)for difference score	*Z*	*p*–value	Effect size
Alpha	Fz	0.57 (0.26)	0.62 (0.30)	0.05 (0.30)	–2.27	.01	0.23
	F3	0.55 (0.24)	0.60 (0.26)	0.06 (0.24)	–3.14	.00	0.32
	F4	0.57 (0.28)	0.58 (0.25)	0.18 (0.75)	–1.75	.04	0.18
	F7	0.67 (0.42)	0.84 (0.67)	0.02 (0.30)	–4.20	.00	0.42
	F8	0.64 (0.38)	0.79 (0.76)	0.14 (0.83)	–3.36	.00	0.34
	Cz	0.48 (0.23)	0.54 (0.38)	0.07 (0.41)	–2.54	.01	0.25
	Pz	0.67 (0.37)	0.73 (0.48)	0.08 (0.39)	–2.28	.01	0.23
Beta	Fz	0.26 (0.13)	0.31 (0.20)	0.09 (0.24)	–4.89	.00	0.49
	F3	0.28 (0.13)	0.34 (0.16)	0.10 (0.16)	–5.44	.00	0.55
	F4	0.29 (0.20)	0.32 (0.12)	0.19 (0.81)	–4.96	.00	0.50
	F7	0.42 (0.35)	0.57 (0.64)	0.06 (0.25)	–5.30	.00	0.53
	F8	0.38 (0.34)	0.53 (0.73)	0.18 (0.85)	–5.69	.00	0.57
	Cz	0.24 (0.16)	0.30 (0.40)	0.1 (0.48)	–4.55	.00	0.46
	Pz	0.28 (0.19)	0.34 (0.25)	0.10 (0.34)	–4.40	.00	0.44
Theta	Fz	0.46 (0.46)	0.56 (0.27)	0.11 (0.64)	–4.67	.00	0.47
	F3	0.43 (0.21)	0.53 (0.22)	0.16 (1.02)	–5.84	.00	0.59
	F4	0.45 (0.24)	0.51 (0.19)	0.20 (0.89)	–4.55	.00	0.46
	F7	0.60 (0.46)	0.80 (0.69)	0.16 (1.37)	–5.55	.00	0.56
	F8	0.54 (0.38)	0.74 (0.77)	0.21 (1.11)	–5.71	.00	0.57
	Cz	0.40 (0.19)	0.50 (0.45)	0.15 (0.94)	–3.93	.00	0.40
	Pz	0.47 (0.31)	0.57 (0.48)	0.17 (1.07)	–4.00	.00	0.40

Note. p<.05 is counted as significant.

Table 2

Descriptive statistics and Wilcoxon Signed Rank Tests for relative activity

Power band	Electrode	Mean (SD) for run 1	Mean (SD) for run 2	Mean (SD) for difference score	z	*p*-value	Effect size
Alpha	Fz	0.43 (0.08)	0.41 (0.08)	–0.02 (0.06)	–2.45	.01	0.25
	F3	0.43 (0.08)	0.4 (0.07)	–0.02 (0.06)	–3.04	.00	0.31
	F4	0.43 (0.07)	0.4 (0.07)	–0.01 (0.06)	–3.13	.00	0.31
	F7	0.40 (0.06)	0.39 (0.06)	–0.02 (0.06)	–2.3	.01	0.23
	F8	0.41 (0.07)	0.39 (0.07)	–0.02 (0.06)	–3.26	.00	0.33
	Cz	0.43 (0.07)	0.41 (0.07)	–0.02 (0.06)	–2.19	.01	0.22
	Pz	0.47 (0.09)	0.44 (0.08)	–0.02 (0.07)	–2.57	.01	0.26
Beta	Fz	0.20 (0.04)	0.21 (0.04)	0.01 (0.04)	–1.3	.1	0.13
	F3	0.23 (0.05)	0.23 (0.04)	0.00 (0.04)	0.52	.7	0.05
	F4	0.22 (0.05)	0.23 (0.04)	0.00 (0.04)	–0.66	.26	0.07
	F7	0.25 (0.05)	0.25 (0.05)	0.00 (0.04)	0.48	.68	0.05
	F8	0.24 (0.05)	0.25 (0.04)	0.01 (0.04)	–0.83	.2	0.08
	Cz	0.21 (0.04)	0.21 (0.04)	0.00 (0.04)	–0.41	.34	0.04
	Pz	0.20 (0.05)	0.21 (0.04)	0.01 (0.04)	–0.84	.2	0.08
Theta	Fz	0.36 (0.07)	0.38 (0.07)	0.01 (0.04)	–2.37	.01	0.24
	F3	0.35 (0.07)	0.37 (0.07)	0.02 (0.04)	–3.76	.00	0.38
	F4	0.35 (0.07)	0.37 (0.07)	0.01 (0.06)	–3.37	.00	0.34
	F7	0.35 (0.06)	0.36 (0.07)	0.02 (0.04)	–2.29	.01	0.23
	F8	0.35 (0.06)	0.36 (0.06)	0.02 (0.05)	–3.3	.00	0.33
	Cz	0.37 (0.06)	0.38 (0.07)	0.01 (0.04)	–1.6	.05	0.16
	Pz	0.33 (0.08)	0.35 (0.07)	0.01 (0.05)	–2.61	.00	0.26

Note. p<.05 is counted as significant.

Significant differences between the two runs were considered to manifest mental fatigue. To reduce the number of statistical comparisons, we calculated the difference scores by subtracting the spectral power at run two from that at run one separately for each frequency band (See *Figure 1*). For the bands that were significantly different, we were interested whether the difference scores for absolute and relative spectral power characteristics correlated with math and trait anxiety.

Table 3

Descriptive statistics and Wilcoxon Signed Rank Tests for spectral ratios difference score

Power band	Electrode	Mean (SD) for run 1	Mean (SD) for run 2	Mean (SD) for difference score	*z*	*p*–value	Effect size
REI	Fz	4.17 (1.15)	3.96 (1.03)	–0.18 (0.94)	–1.25	.11	0.13
	F3	3.65 (1.07)	3.54 (0.92)	–0.09 (0.83)	0.04	.52	0
	F4	3.22 (0.86)	3.17 (0.82)	–0.14 (0.82)	0.09	.54	0.01
	F7	3.73 (1.07)	3.57 (0.92)	–0.03 (0.77)	–0.8	.21	0.08
	F8	3.34 (0.90)	3.18 (0.75)	–0.16 (0.81)	–1.03	.15	0.1
	Cz	4.01 (1.01)	3.89 (0.91)	–0.13 (0.85)	–0.62	.27	0.06
	Pz	4.23 (1.22)	3.99 (1.02)	–0.20 (1.04)	–1.09	.14	0.11
TLI		0.88 (0.66)	0.95 (0.53)	0.04 (0.54)	–0.42	.34	0.04

Note. p<.05 is counted as significant.

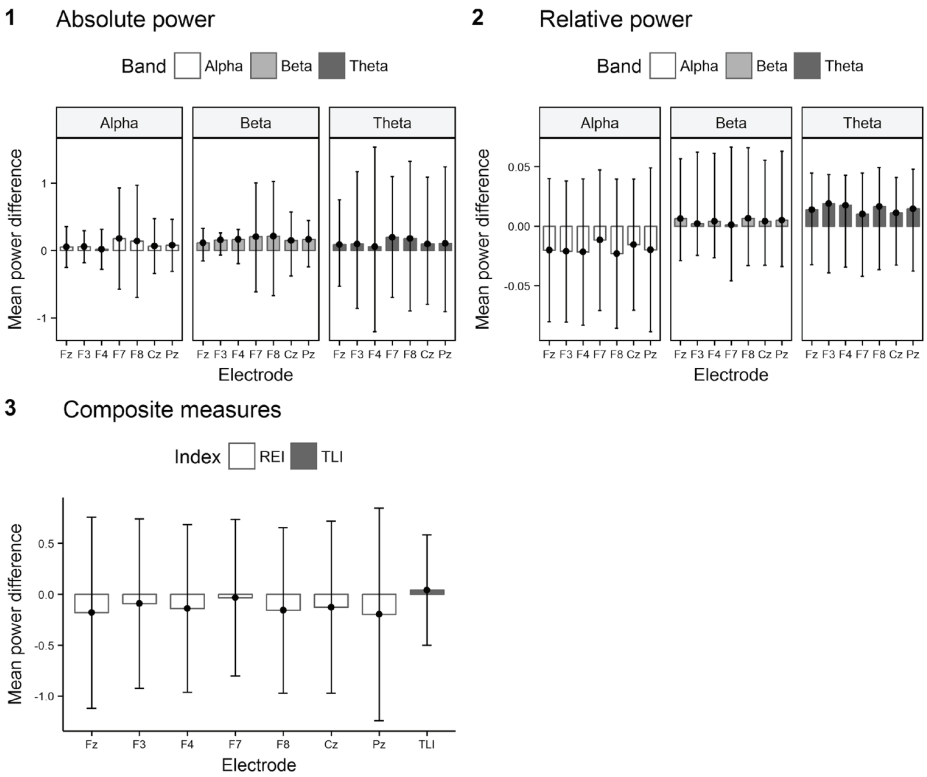


Figure 1. Deviation bar graphs for EEG power difference scores from run one to run two. 1. Absolute power difference; 2. Relative power difference; 3. Power difference for the two composite measures (REI and TLI).

Table 4

Correlations between EEG power difference score and math and trait anxiety

			Math anxiety			Trait anxiety		
Power band		**Electrode**	**Rho**	***p*–value**	***p* FDR adjusted**	**Rho**	***p*–value**	***p* FDR adjusted**
Absolute	Alpha	Fz	0.06	.55	.98	0.07	.54	.7
		F3	0.11	.29	.98	0.08	.49	.7
		F4	0.02	.81	.98	0.03	.76	.86
		F7	0.0	.98	.98	0.1	.36	.7
		F8	–0.15	.16	.98	0.17	.11	.28
		Cz	–0.02	.87	.98	0.1	.36	.7
		Pz	0.14	.17	.98	0.0	.98	.99
	Beta	Fz	0.03	.75	.98	–0.17	.11	.28
		F3	0.02	.88	.98	–0.07	.53	.7
		F4	–0.04	.68	.98	–0.08	.44	.7
		F7	–0.08	.46	.98	–0.05	.67	.8
		F8	–0.08	.45	.98	–0.08	.45	.7
		Cz	–0.02	.87	.98	0.08	.45	.7
		Pz	0.1	.33	.98	–0.03	.81	.89
	Theta	Fz	0.03	.8	.98	–0.06	.59	.74
		F3	–0.01	.9	.98	0.0	.99	.99
		F4	–0.09	.41	.98	–0.07	.53	.7
		F7	–0.06	.57	.98	0.0	.99	.99
		F8	–0.16	.12	.98	0.04	.69	.8
		Cz	–0.06	.55	.98	0.11	.31	.7
		Pz	0.13	.23	.98	–0.07	.53	.7
Relative	Alpha	Fz	0.05	.66	.98	0.26	.02	.15
		F3	0.06	.54	.98	0.2	.06	.19
		F4	0.09	.4	.98	0.22	.04	.15
		F7	0.03	.77	.98	0.24	.02	.15
		F8	–0.07	.54	.98	0.29	.01	.15
		Cz	–0.01	.96	.98	0.09	.38	.7
		Pz	0.01	.89	.98	0.11	.33	.7
	Theta	Fz	–0.05	.63	.98	–0.22	.04	.15
		F3	–0.04	.72	.98	–0.23	.03	.15
		F4	–0.09	.39	.98	–0.22	.04	.15
		F7	–0.04	.68	.98	–0.22	.04	.15
		F8	–0.06	.55	.98	–0.23	.03	.15
		Cz	0.0	.96	.98	–0.18	.09	.26
		Pz	0.04	.69	.98	–0.22	.04	.15

Note. p< .05 is counted as significant.

Table 5
Partial correlations between EEG power difference score and mats and trait anxiety

			Math anxiety			Trait anxiety		
Power band		**Electrode**	**Rho**	***p*–value**	***p* FDR adjusted**	**Rho**	***p*–value**	***p* FDR adjusted**
Absolute	Alpha	Fz	0.07	.48	.97	0.05	.64	.97
		F3	0.13	.21	.97	0.05	.67	.97
		F4	0.06	.56	.97	–0.00	.99	.97
		F7	0.04	.67	.97	0.07	.51	.97
		F8	–0.11	.28	.97	0.16	.13	.97
		Cz	0.02	.88	.97	0.08	.48	.97
		Pz	0.16	.13	.97	–0.02	.82	.97
	Beta	Fz	0.03	.74	.97	–0.15	.16	.97
		F3	0.02	.87	.97	–0.07	.53	.97
		F4	–0.03	.76	.97	–0.08	.46	.97
		F7	–0.06	.54	.97	–0.04	.67	.97
		F8	–0.06	.54	.97	–0.07	.5	.97
		Cz	–0.01	.94	.97	0.09	.42	.97
		Pz	0.10	.33	.97	–0.03	.79	.97
	Theta	Fz	0.02	.87	.97	–0.04	.72	.97
		F3	–0.01	.95	.97	0.01	.94	.97
		F4	–0.05	.64	.97	–0.03	.75	.97
		F7	–0.02	.85	.97	0.00	.99	.97
		F8	–0.14	.17	.97	0.05	.62	.97
		Cz	–0.04	.72	.97	0.13	.22	.97
		Pz	0.14	.18	.97	–0.07	.51	.97
Relative	Alpha	Fz	0.08	.47	.97	0.20	.06	.29
		F3	0.07	.49	.97	0.12	.25	.59
		F4	0.10	.35	.97	0.14	.2	.53
		F7	0.03	.79	.97	0.19	.08	.32
		F8	–0.08	.45	.97	0.22	.04	.29
		Cz	0.00	.97	.97	0.06	.6	.84
		Pz	0.02	.85	.97	0.05	.61	.84
	Theta	Fz	–0.08	.44	.97	–0.22	.04	.29
		F3	–0.05	.6	.97	–0.20	.06	.29
		F4	–0.10	.35	.97	–0.20	.06	.29
		F7	–0.02	.85	.97	–0.21	.05	.29
		F8	–0.06	.56	.97	–0.22	.04	.29
		Cz	–0.01	.91	.97	–0.17	.1	.36
		Pz	0.04	.71	.97	–0.20	.06	.29

Note. $p < .05$ is counted as significant.

Table 4 presents correlations between the difference scores in band powers and anxiety (math and trait). There were no significant correlations between absolute band power changes and anxiety (math and trait). For math anxiety, no correlations were observed with the relative band power difference score. However, trait anxiety was significantly correlated with difference scores for relative alpha and theta activity for most of the electrodes. To correct for multiple comparisons, we used the FDR-correction (Benjamini & Hochberg, 1995). After this correction, none of the comparisons reached the level of significance (see *Table 4*).

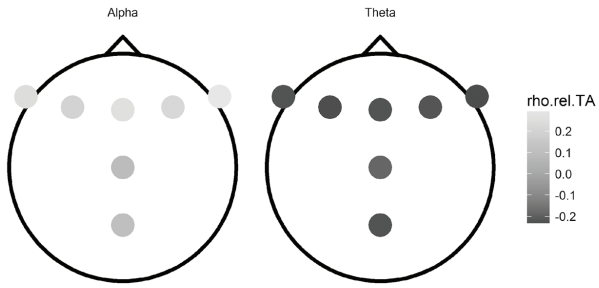


Figure 2. Correlations of power difference score and trait anxiety for alpha and theta relative power

Partial correlations that account for the level of fatigue of a participant at the start of the experiment were not significant for absolute power values. For relative power values, difference scores marginally correlated with trait anxiety at most frontal electrodes in theta band, as well as at the F8 electrode in alpha band (see *Figure 2*). However, no results reached significance after the FDR-correction (see *Table 5*).

Discussion

The aim of this study was to investigate whether the level of mental fatigue during a cognitive task is associated with math and trait anxiety. Mental fatigue was not measured directly. The approach we used in the study was based on the findings which showed that mental fatigue can manifest itself in EEG spectral characteristics at different frequency bands.

We started by exploring differences in EEG spectral power, measured before and after performing math- and language-related cognitive tasks. We analyzed EEG activity in narrow frequency bands together with composite EEG measures (Task Load Index, TLI; and Relative Energy Index, REI). In previous studies, REI was shown to be reversed proportional to energy and related to a state of drowsiness, which is a facet of the state of mental fatigue (Eoh et al., 2005; Lal & Craig, 2001). Our data did not yield significant differences in REI between the two runs. This result might be explained by a relatively short duration of the experiment (approximately an hour and a half). This relatively short duration, combined with a 5-minute break during the experimental task, may have led to an absence of drowsiness. No significant changes were found in the TLI index. Previous research found

changes in TLI with an increase of difficulty (Gevins, Smith, McEvoy, & Yu, 1997). In the present study, the experimental task was not difficult, and the level of difficulty stayed constant throughout the task.

Absolute power in the eyes-closed resting state was higher at run two than at run one in all bands. This is in line with findings by Craig and colleagues (2012). However, the relative power results were not consistent with previous findings (Craig et al., 2012): only theta power increased from run one to run two, while alpha decreased and beta did not show any change. Contrary to our prediction of alpha band power growth with increase of fatigue, as expected from previous research (Craig et al., 2012), relative alpha rhythm decreased in all electrodes except Pz. The existing literature is somewhat inconsistent in establishing the link between alpha activity and fatigue: it was reported to be both decreasing and increasing in different studies (see the review in Craig et al., 2012). In our study, the decrease in relative alpha band activity might be interpreted as a marker of task engagement (Cheng & Hsu, 2011). Such alertness might be a result of the need for participants to switch between several tasks (arithmetic, lexical, and algebraic) during the session, so that they have to stay vigilant. Overall, our data showed contradictory results for alpha band changes in absolute and relative power approaches.

The data on the link between beta and fatigue are also quite unclear. Craig and colleagues (2012) reported beta increases with fatigue (three studies); decreases with fatigue (one study); and no change (two studies). Our data also show no significant EEG differences before and after performing cognitive tasks. There are different ways to calculate EEG spectral power. Here we used the relative power approach which accounts for all of the spectrum. It is possible that the discrepancies between our study and the aforementioned ones in beta change might be caused by the relative power approach we used. In this approach, more data on EEG activity is used for the calculation of the estimation, so that it might be more accurate and reliable in estimation of power change.

Theta band spectral power differences were consistent for both absolute and relative power approaches. Theta power is considered a robust indicator of fatigue (Craig et al., 2012). In our study theta increased from run one to run two, suggesting mental fatigue in our participants. However, this finding is not consistent with the findings for the other indices. Further research is needed to find a plausible explanation for this contradictory result.

Next, we performed analyses to explore the link between EEG spectral power changes and math and trait anxiety. Given that there were no differences in composite measures across the two runs, and mixed evidence for the beta band (no differences in relative power), we focused on alpha and theta bands in further analysis. Alpha and theta bands showed prominent changes after task performance for both absolute and relative spectral power. However, we found no correlations of absolute EEG activity changes with either anxiety. There were also no correlations between relative activity and math anxiety. Trait anxiety was positively correlated with a decrease in relative alpha and negatively with an increase in relative theta. These correlations failed to reach significance after correction for multiple comparisons. However, the indicated associations may be real, as they were consistent in spatial distribution of the effect (there was similar activity on the adjacent electrodes, see

Figures 2). The current study may be underpowered to demonstrate these associations robustly.

The positive correlation between the alpha difference score and TA, shown in our study, might be evidence of people with high anxiety having bigger differences between run one and run two; i.e., people with higher TA have lower alpha after the task completion. This might reflect the fact that these people were more engaged with the task, and thus were more vigilant. This result is in line with previous research which showed that lower alpha might be associated with increased focus (Oken, Salinsky, & Elsas, 2006; Palva & Palva, 2007). Another interpretation which might be inferred from the partial correlation analysis, is that the alpha band spectral power change is affected by the participant's previous state rather than by his or her level of anxiety during the task. In this case a low level of fatigue before the task might have led to lower fatigue by the end of the task, as compared to people who were fatigued before the recording. In future studies, an independent measurement of initial fatigue level should be taken and controlled for during the analysis.

Our data also showed that there was negative correlation between the theta difference score and TA. A possible explanation for that might be related to the arousal effect of anxiety. For people with high anxiety, participating in the experiment could have stressed them that led to mobilizing their resources, and having less effects of tiredness. Both these findings might be discussed in the framework of the optimal performance theory of Yerkes and Dodson (1908), according to which participants in our study were more vigilant because of their increased anxiety, induced by participation in the experiment. These results are also in line with the results of previous studies, which identified the nonlinear effects of anxiety. For example, one study showed that motivation moderated the relationship between math performance and math anxiety (Wang et al., 2015). The study found that math anxiety might have served as a positive arousal factor for people with high motivation.

Contrary to our initial hypothesis, we failed to find any association between math anxiety and fatigue. This may be interpreted in the context of the motivational control theory of cognitive fatigue (Hockey, 2011). According to that theory, fatigue arises when there exists a conflict between a current and an alternative goal (including rest and biological needs). A person's goals are constantly being evaluated, and when demands for task performance increase, he or she experiences a subjective feeling of discomfort. This leads either to increasing effort, or to leaving the effort unchanged by either sticking to a previous goal despite a deficit in performance, or by switching to another goal. Fatigue arises when a person feels a demand for greater effort, and is expected to dissipate if a person chooses to switch to another task or not to increase his or her effort. One of the predictions of the model discussed by Hockey, is that changing the goal resets the need for effort to normal, thus removing fatigue.

Since MA is considered to be a domain-specific anxiety linked to math performance, in our study there may have been different processes underway during the math-related and lexical tasks of people with high MA. For example, MA may have led to greater arousal during math tasks, preventing a buildup of fatigue. It is not clear whether such a buildup would lead to cancelling out the effects of fatigue experienced during tasks; or whether the tasks on the whole were not long enough

by themselves to induce fatigue. TA is not specific to any task, and thus there was no goal change for people with high TA. This may explain why the difference scores were linked to the TA level: people with TA might have experienced the whole experiment as one stressful task that continuously sustained their arousal.

Conclusion

No significant links were found between math and trait anxiety, and mental fatigue. However, some interesting associations were indicated, and need to be explored by a study with better statistical power. Moreover, to better understand the link between math anxiety and mental fatigue, an experiment with only math-related tasks should be designed. Other measures of mental fatigue (e.g. yawning, nodding off, self-reports) should be used to understand the consistency of EEG and other measures of mental fatigue, and to clarify the differential pattern of results for different spectral power bands.

Limitations

First, the goal switching within the experimental task might not have allowed the tracing of the link between MA and fatigue. Second, the length of the experiment might have affected our findings: although there were some changes in EEG associated with mental fatigue, the length of time of the task might have been insufficient for robustly inducing fatigue. Third, our sample might not have been large enough to detect weak correlations. Previous research reported weak to moderate correlations between trait anxiety and mental fatigue. The link between math anxiety and mental fatigue may be even weaker, and therefore undetectable in our sample. Fourth, mental fatigue was not measured directly.

Acknowledgement

This work was supported by the Russian Ministry for Education and Science within the project No.25.8663.2017/9.10

References

Alexander, L. & Martray, C.R. (1989). The development of an abbreviated version of the Mathematics Anxiety Rating Scale. *Measurement and Evaluation in Counseling and Development*, *22*(3), 143–150. https://doi.org/10.1080/07481756.1989.12022923

Artemenko, C., Daroczy, G., & Nuerk, H.-C. (2015). Neural correlates of math anxiety–an overview and implications. *Frontiers in Psychology*, *6*, 1333. https://doi.org/10.3389/fpsyg.2015.01333

Ashcraft, M.H. & Moore, A.M. (2009). Mathematics Anxiety and the Affective Drop in Performance. *Journal of Psychoeducational Assessment*, *27*(3), 197–205. https://doi.org/10.1177/0734282908330580

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. WileyRoyal Statistical Society. https://doi.org/10.2307/2346101

Bloniewski, T., Likhanov, M., Khuc, J., Budakov, A., Zakharov, I., & Kovas, Y. (2018). Test-retest reliability of behavioral and ERP measures in math-related processing. *The European Proceedings of Social & Behavioural Sciences*. https://dx.doi.org/10.15405/epsbs.2018.11.02.13

Boksem, M.A.S., Meijman, T.F., & Lorist, M.M. (2005). *Mental fatigue, motivation and action monitoring*. https://doi.org/10.1016/j.biopsycho.2005.08.007

Boksem, M.A.S. & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*(1), 125–139. https://doi.org/10.1016/j.brainresrev.2008.07.001

Cantero, J.L., Atienza, M., Salas, R.M., & Gómez, C.M. (1999). Alpha EEG coherence in different brain states: an electrophysiological index of the arousal level in human subjects. *Neuroscience Letters*, *271*(3), 167–170. https://doi.org/10.1016/S0304-3940(99)00565-0

Carretié, L., Mercado, F., Tapia, M., & Hinojosa, J.A. (2001). Emotion, attention, and the "negativity bias", studied through event-related potentials. *International Journal of Psychophysiology*, *41*(1), 75–85. https://doi.org/10.1016/S0167-8760(00)00195-1

Cheng, S.-Y. & Hsu, H.-T. (2011). Mental Fatigue Measurement Using EEG. *Risk Management Trends*. https://doi.org/10.5772/16376

Craig, A., Tran, Y., Wijesuriya, N., & Boord, P. (2006). A controlled investigation into the psychological determinants of fatigue. *Biological Psychology*, *72*(1), 78–87. https://doi.org/10.1016/J.BIOPSYCHO.2005.07.005

Craig, A., Tran, Y., Wijesuriya, N., & Nguyen, H. (2012). Regional brain wave activity changes associated with fatigue. *Psychophysiology*, *49*(4), 574–582. https://doi.org/10.1111/j.1469-8986.2011.01329.x

Dowker, A., Sarkar, A., & Looi, C.Y. (2016). Mathematics Anxiety: What Have We Learned in 60 Years? *Frontiers in Psychology*, *7*, 508. https://doi.org/10.3389/fpsyg.2016.00508

Edwards, M.S., Edwards, E.J., & Lyvers, M. (2017). Cognitive trait anxiety, stress and effort interact to predict inhibitory control. *Cognition and Emotion*, *31*(4), 671–686. https://doi.org/10.1080/02699931.2016.1152232

Eoh, H.J., Chung, M.K., & Kim, S.-H. (2005). Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. *International Journal of Industrial Ergonomics*, *35*(4), 307–320. https://doi.org/10.1016/J.ERGON.2004.09.006

Eysenck, M.W. & Calvo, M.G. (1992). Anxiety and Performance: The Processing Efficiency Theory. *Cognition & Emotion*, *6*(6), 409–434. https://doi.org/10.1080/02699939208409696

Eysenck, M.W., Derakshan, N., Santos, R., & Calvo, M.G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Gevins, A. & Smith, M.E. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, *4*(1–2), 113–131. https://doi.org/10.1080/14639220210159717

Gevins, A., Smith, M.E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*(4), 374–385. https://doi.org/10.1093/cercor/7.4.374

Hancock, P.A., Desmond, P.A., & Matthews, G. (2012). Conceptualizing and defining fatigue. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The Handbook of Operator Fatigue* (pp. 63–73). Burlington: Ashgate Publishing Company.

Hanin, Y.L. & Spielberger, C.D. (1983). The development and validation of the Russian Form of the State-Trait Anxiety Inventory. *Series in Clinical & Community Psychology: Stress & Anxiety*, *2*, 15–26.

Hockey, G.R. J. (2011). A motivational control theory of cognitive fatigue. In P.L. Ackerman (Ed.), *Cognitive fatigue: multidisciplinary perspectives on current research and future appli-*

cations (pp. 167–188). Washington, DC: American Psychological Association. https://doi.org/10.1037/12343-008

Jiang, N., Sato, T., Hara, T., Takedomi, Y., Ozaki, I., & Yamada, S. (2003). Correlations between trait anxiety, personality and fatigue. *Journal of Psychosomatic Research*, *55*(6), 493–500. https://doi.org/10.1016/S0022-3999(03)00021-7

Kamzanova, A.T., Matthews, G., Kustubayeva, A.M., & Jakupov, S.M. (2011). EEG Indices to Time-On-Task Effects and to a Workload Manipulation (Cueing). *International Journal of Psychological and Behavioral Sciences*, *5*(8), 19–22.

Knyazev, G. (2007). Motivation, emotion, and their inhibitory control mirrored in brain oscillations. *Neuroscience & Biobehavioral Reviews, 31*(3), 377–395. https://doi.org/10.1016/j.neubiorev.2006.10.004

Kovas, Y., Galajinsky, E.V., Boivin, M., Harold, G.T., Jones, A., Lemelin, J.-P., …& Malykh, S. (2013). The Russian School Twin Registry (RSTR): Project PROGRESS. *Twin Research and Human Genetics*, *16*(1), 126–133. https://doi.org/10.1017/thg.2012.133

Lal, S.K.L. & Craig, A. (2001). A critical review of the psychophysiology of driver fatigue. *Biological Psychology*, *55*(3), 173–194. https://doi.org/10.1016/S0301-0511(00)00085-5

Likhanov, M., Zakharov, I., Kuzmina, Y., Budakova, A., Vasin, G., Malykh, S., & Kovas, Y. (2017). Math Anxiety Does Not Moderate the Link between Spatial and Math Ability. *The European Proceedings of Social & Behavioural Science*, https://doi.org/10.15405/epsbs.2017.12.21

Lorist, M.M., Klein, M., Nieuwenhuis, S., Jong, R., Mulder, G., & Meijman, T.F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology*, *37*(5), 614–625. https://doi.org/10.1111/1469-8986.3750614

Matthews, G., Desmond, P.A., & Hitchcock, E.M. (2012). Dimensional Models of Fatigue. In G. Matthews, P. A. Desmond, C. Neubauer, & P. A. Hancock (Eds.), *The Handbook of Operator Fatigue* (pp. 139–154). Burlington: Ashgate Publishing Company.

Núñez-Peña, M.I. & Suárez-Pellicioni, M. (2014). Less precise representation of numerical magnitude in high math-anxious individuals: An ERP study of the size and distance effects. *Biological Psychology*, *103*, 176–183. https://doi.org/10.1016/j.biopsycho.2014.09.004

Nuwer, M. (1988). Quantitative EEG: I. Techniques and problems of frequency analysis and topographic mapping. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, *5*(1), 1–43. https://doi.org/10.1097/00004691-198801000-00001

Oken, B.S., Salinsky, M.C., & Elsas, S.M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, *117*(9), 1885–1901. https://doi.org/10.1016/J.CLINPH.2006.01.017

Palva, S., & Palva, J.M. (2007). New vistas for α-frequency band oscillations. *Trends in Neurosciences*, *30*(4), 150–158. https://doi.org/10.1016/J.TINS.2007.02.001

Putman, P. (2011). Resting state EEG delta–beta coherence in relation to anxiety, behavioral inhibition, and selective attentional processing of threatening stimuli. *International Journal of Psychophysiology, 80*(1), 63–68. https://doi.org/10.1016/j.ijpsycho.2011.01.011

Rezaei, F., Hosseini Ramaghani, N.A., & Fazio, R.L. (2017). The effect of a third-party observer and trait anxiety on neuropsychological performance: the Attentional Control Theory (ACT) perspective. *The Clinical Neuropsychologist*, *31*(3), 632–643. https://doi.org/10.1080/13854046.2016.1266031

Richardson, F.C. & Suinn, R.M. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551–554. https://doi.org/10.1037/h0033456

Rijsdijk, F.V. & Sham, P.C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in Bioinformatics*, *3*(2), 119–33. https://doi.org/10.1093/bib/3.2.119

Santamaria, J. & Chiappa, K.H. (1987). The EEG of drowsiness in normal adults. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, *4*(4), 327–382. https://doi.org/10.1097/00004691-198710000-00002

Spielberger, C.D. (1989). *State-Trait Anxiety Inventory: Bibliography* (2nd ed.). Palo Alto: CA: Consulting Psychologists Press.

Strijkstra, A.M., Beersma, DG. M., Drayer, B., Halbesma, N., & Daan, S. (2003). Subjective sleepiness correlates negatively with global alpha (8–12 Hz) and positively with central frontal theta (4-8 Hz) frequencies in the human resting awake electroencephalogram. *Neuroscience Letters*, *340*(1), 17–20. https://doi.org/10.1016/S0304-3940(03)00033-8

Suárez-Pellicioni, M., Núñez-Peña, M.I., & Colomé, À. (2016). Math anxiety: A review of its cognitive consequences, psychophysiological correlates, and brain bases. *Cognitive, Affective, & Behavioral Neuroscience*, *16*(1), 3–22. https://doi.org/10.3758/s13415-015-0370-7

Wang, X., Huang, Y., Ma, Q., & Li, N. (2012). Event-related potential P2 correlates of implicit aesthetic experience. *Neuroreport*, *23*(14), 862–866. https://doi.org/10.1097/WNR.0b013e3283587161

Wang, Z., Lukowski, S.L., Hart, S. A., Lyons, I.M., Thompson, L.A., Kovas, Y., …& Petrill, S.A. (2015). Is Math Anxiety Always Bad for Math Learning? The Role of Math Motivation. *Psychological Science*. *26*(12), 1863–1876. https://doi.org/10.1177/0956797615602471

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., …& Gutberlet, I. (2014). Frontal theta activity reflects distinct aspects of mental fatigue. *Biological Psychology*, *96*, 57–65. https://doi.org/10.1016/j.biopsycho.2013.11.010

Wijesuriya, N., Tran, Y., & Craig, A. (2007). The psychophysiological determinants of fatigue. *International Journal of Psychophysiology*, *63*(1), 77–86. https://doi.org/10.1016/j.ijpsycho.2006.08.005

Yerkes, R.M. & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/cne.920180503

Original manuscript received August 09, 2018
Revised manuscript accepted October 30, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

The Factorial Structure of Spatial Abilities in Russian and Chinese Students

Maxim V. Likhanov[1,3], Victoria I. Ismatullina[2], Alexander Y. Fenin[2], Wei Wei[4], Kaili Rimfeld[5], Ekaterina P. Maslennikova[6], Elena A. Esipenko[3], Ksenia R. Sharafieva[3], Inna V. Feklicheva[6], Nadezhda A. Chipeeva[6], Anna V. Budakova[3], Elena L. Soldatova[6], Xinlin Zhou[7] & Yulia V. Kovas*[3,8]

[1] *Educational Fund "Talent and Success", Educational center "Sirius", Russia*

[2] *Psychological Institute of Russian Academy of Education, Russia*

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

[4] *Zhejiang University, China*

[5] *King's College London, United Kingdom*

[6] *South Ural State University, Russia*

[7] *Beijing Normal University, China*

[8] *Goldsmiths, University of London, United Kingdom*

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

Background. Recent research has suggested a unifactorial structure of spatial ability (SA). However, further studies are needed to replicate this finding in different populations.

Objective. This study aims to explore the factorial structure of SA in samples of 921 Russian and 229 Chinese university students.

Design. A gamified spatial abilities battery was administered to all participants. The battery consists of 10 different domains of SA, including 2D and 3D visualization, mental rotation, spatial pattern assembly, spatial relations, spatial planning, mechanical reasoning, spatial orientation, and spatial decision-making speed and flexibility.

Results. The results of the factor analysis showed a somewhat different pattern for different samples. In the Russian sample, the unifactorial structure, shown previously in a large UK sample (Rimfeld et al., 2017), was replicated. A single factor explained 40% of the variance. In the Chinese sample two factors emerged: the first factor explained 26% of the variance and the second factor, including only mechanical reasoning and cross-sections tests, explained 14%. The results also showed that the Chinese sample significantly outperformed the Russian sample in five out of the 10 tests. Russian students showed better performance in only two of the tests. The effects of all group comparisons were small.

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

doi: 10.11621/pir.2018.0407
http://psychologyinrussia.com

Conclusion. Overall, a similar amount of variance in the 10 tests was explained in the two samples, replicating results from the UK sample. Future research is needed to explain the observed differences in the structure of SA.

Keywords: spatial ability (SA), factorial structure, Russian and Chinese students

Introduction

Spatial ability (SA), "the ability to generate, retain, retrieve, and transform well-structured visual images" (Lohman, 1996), has been extensively studied for the last 60 years (Kell & Lubinski, 2013). However, it is still not clear whether spatial ability is a unitary construct or whether it has a more complex structure (Aristova et al., 2018). The results of psychometric studies suggest several components of spatial abilities: spatial visualization (Lohman, 1979), spatial orientation (Hegarty Montello, Richardson, Ishikawa, & Lovelace 2006), spatial imagination (Jansen, 2009), mental rotation (Shepard & Metzler, 1971) and spatial relations (Lohman et al., 1987). Researchers have also suggested different classifications for the numerous components of spatial ability. For example, Uttal and colleagues (2013) suggested a classification based on the combination of intrinsic/extrinsic and static/dynamic categories of spatial skills.

Nevertheless, most studies of SA components have suggested the existence of two groups of components: small- and large-scale spatial abilities (Jansen, 2009). The small-scale group includes spatial visualization, spatial imagination, and mental rotation (Jansen, 2009). These abilities involve mental operations with objects: transformation (Zacks, Mires, Tversky, & Hazeltine, 2000), mental rotation (Blajenkova & Kozhevnikov, 2005), and manipulation (Kozhevnikov & Hegarty, 2001). The large-scale group includes spatial orientation; object location from the spectator's point of view and an ability to assess the direction and distance (Jansen, 2009); navigation ability (Kozhevnikov, Motes, Rasch, & Blajenkova 2006); sense of direction (DeBeni & Pazzaglia, 2006), spatial orientation (Kozhevnikov & Hegarty, 2001); and other abilities. This group of abilities is connected with changes in the spectator's visual perspective, while other objects' positions remain the same (Hegarty & Waller, 2004).

It is unclear whether small- and large-scale skills are independent or a part of a unitary meta-construct. Hegarty and colleagues (2006) discussed four models which purport to explain the relationships between these two types of skills:

1) the "unitary model" assumes that both skills overlap completely;
2) the "total dissociation model" assumes that the skills are distinct;
3) the "partial dissociation model" assumes that the abilities have similarities and differences;
4) and the "mediation model" assumes that small- and large-scale skill are linked, but this link is mediated by a third variable (i.e., intelligence).

The results of meta-analysis of psychometric studies from 1985 to 2014 (Wang et al., 2014) confirmed the dissociation between small- and large-scale abilities, with only a moderate correlation ($r = .27$) between them. Neuropsychological studies also suggest a partial dissociation between small- and large-scale spatial abilities

(Hegarty et al., 2006; Morris & Parslow, 2004). Small-scale spatial ability tasks are associated with the parietal lobes (Kosslyn & Thompson, 2003), whereas large-scale spatial abilities have been linked to the hippocampus (Gogos et al., 2010; Hughdahl et al., 2006) and medial and temporal lobes (Parslow et al., 2004). There are also activation patterns triggered by both groups of tests such as, for example, vision, muscle sense, etc. (Wang, 2014).

Experimental studies further support the partial dissociation model. For example, one study (Jansen, Wiedenbauer, & Hahn, 2010), conducted in adults, showed that large-scale ability can be improved slightly by training which uses small-scale ability tasks ($\acute{\eta}^2 = .12$). However, there was no such effect in another study with the same design, but which involved children (Jansen, 2009). Given these results, a partial dissociation model seems rather plausible, suggesting that small- and large-scale skills can be studied separately. In the present paper, we focus on the structure of SA related to small-scale skills.

Research that attempts to clarify the structure of SA requires administration of a large number of tests, which makes it difficult in practice. One way to reduce the strain on participants during such studies is to use gamified online testing. A recent study by Rimfeld and colleagues (2017) introduced a gamified online spatial abilities battery (King's Challenge), which assesses 10 different small-scale spatial ability domains: 1) mazes task (searching for a way through a 2D maze in a speeded task); 2) 2D drawing (sketching a 2D layout of a 3D object from a specified viewpoint); 3) Elithorn maze (joining together as many dots as possible from an array); 4) pattern assembly (visually combining pieces of objects together to make a whole); 5) mechanical reasoning (multiple-choice naïve physics questions); 6) paper folding (visualizing where the holes are situated after a piece of paper is folded and a hole is punched through it); 7) 3D drawing (sketching a 3D drawing from a 2D diagram); 8) shape rotation (mentally rotating objects); 9) perspective-taking (visualizing objects from a different perspective); and 10) cross-sections (visualizing cross-sections of objects).

The study was conducted on 1367 twin pairs (ages 19–21) from the UK-representative Twins Early Development Study (TEDS). The results of the study suggested a unifactorial structure of small-scale SA. Factor loadings ranged from .44 to .71 for different tests, with perspective-taking and mazes having the lowest loadings, and 3D drawing having the highest (see *Fig. 2* in Rimfeld et al., 2017). The study also showed that individual differences in this construct were largely explained by genetic factors (69%).

More research is needed in order to test whether this unitary structure of SA can be replicated in different cultures. Cross-cultural comparison of the structure of SA may be useful for understanding the sources of individual differences within any population. A study by Sakamoto and colleagues (2014) showed that Chinese and Japanese adults on average demonstrate greater visual spatial ability than U.S. adults (Cohen's $d = 0.5–0.7$). This advantage may be explained by the impact of the Chinese writing system on the development of spatial ability (Flaherty & Connolly, 1995; Gitterman & Sies, 1992; Rodic et al., 2015), or by other cultural differences.

In the present study we aim to replicate the unitary structure which was suggested by Rimfeld and colleagues (2017), using an adaptation of the same King's Challenge battery, in Russian and Chinese samples. The main aims of the study

were: 1) to explore the factorial structure of SA in the two student samples: Russian and Chinese; 2) to compare the factorial structures in both samples with the previously reported findings from the UK population; and 3) to evaluate the relative strengths and weaknesses in performance on the 10 tests, across the two samples.

Methods

Participants

The sample consisted of students from Russia and China.

The Russian sample consisted of 921 participants (16 to 37 years old, M = 19.59, SD = 1.85; 348 males and 573 females). Participants were recruited from five leading universities in Russia: three universities from the top 10, and two from the top 100, out of the more than 500 universities in Russia.

The Chinese sample consisted of 229 participants (17 to 30 years old, M = 19.92, SD = 1.73; 88 males and 114 females). Participants were recruited from top-tier universities: one university from top 15–20, and one university from the top three to five, out of the approximately 300 universities in China. The initial sample included 245 students, but data from 16 participants were excluded because of reported technical problems.

Reference Sample

The current study aims to replicate the factorial structure of SA, previously shown in the UK sample using the same 10-measure instrument. Therefore, we compare our results to the reference sample, the Twin Early Development Study (TEDS; Rimfeld et al., 2017). The TEDS is an unselected sample, representative of the UK population. The results reported in Rimfeld et al. are not directly comparable to the student samples collected in the present study. We conducted additional analyses on the TEDS data, selecting only participants who reported being involved in higher education. Unlike the students in the present study, who came from top universities, the TEDS students were undergraduate and postgraduate students from a large number of unselected universities in the UK and abroad. For this reason, we do not directly compare the results with the UK student sample. In terms of the structure of SA, the results (available from the authors) showed a very similar structure for the TEDS student subsample as for the whole UK sample. Therefore, we compare our results with the results published in Rimfeld et al. (2017).

Measures

Spatial Ability. All participants completed interactive tests from the King's Challenge (KC) gamified interactive battery, translated into Russian and Chinese. As described in the introduction, KC is an online battery that tests 10 major dimensions of spatial ability: mental rotation, spatial visualization, spatial reasoning, perspective-taking, and mechanical reasoning (see Esipenko et al., this issue for more detail on each test). The battery is gamified, so that participants engage in building and protecting the King's Castle, with all tests linked by the same storyline. The total score was computed by summing the results from all 10 tests. Each of the 10

tests included a different number of items, and presented different levels of difficulty (see Esipenko et al., this issue).

Demographic information. Demographic information, including age, sex, and level of education, was collected in both samples, using an online questionnaire.

Procedure

The data for the Russian sample were collected in one session at the participants' universities.

The Chinese participants were recruited by emailing participants and inviting them to complete the on-line test battery. Data from respondents who experienced technical issues in any test were excluded from the analyses.

Results

Descriptive statistics

Table 1 presents the means, standard deviations, standard error, and split-half reliability for the 10 spatial measures for Russian and Chinese samples. Differences in the number of participants across the tests is explained by some respondents not completing the whole battery. According to the criteria suggested by Comrey and Lee (1992), the sample size was adequate (Chinese sample) and excellent (Russian sample) for the performed analyses. See density plots for all variables for both Russian and Chinese samples in *Appendix, Figures 1, 2.*

Table 1

Descriptive statistics for KC tests

	Russia			China		
	n	**Split-half**	**mean (SD)**	**mean (SD)**	**Split-half**	**n**
Cross-sections	921	.74	6.71 (3.65)	8.06 (3.74)	.82	229
2D drawing	918	.80	3.53 (1.23)	3.63 (1.19)	.80	229
Pattern assembly	909	.69	6.37 (3.04)	7.37 (3.10)	.74	229
Elithorn maze	904	.88	7.74 (1.97)	6.99 (2.15)	.90	227
Mechanical reasoning	902	.56	9.76 (2.70)	10.75 (2.43)	.54	225
Paper folding	897	.85	8.17 (4.25)	10.55 (3.54)	.83	224
3D drawing	889	.78	2.52 (1.82)	2.87 (1.80)	.93	224
Shape rotation	886	.82	7.76 (4.05)	8.82 (3.71)	.79	223
Perspective taking	885	.86	4.45 (4.14)	3.29 (3.07)	.80	223
Mazes	883	.60	5.35 (2.03)	5.47 (2.06)	.63	222
Total score	883	NA	62.40 (19.66)	68.24 (17.94)	NA	222

Note. The results of the different subtests cannot be directly compared, as they each included a different number of items.

Given the consistent sex differences reported for spatial ability (see, for example, Tosto et al., 2014; Toivainen et al., 2018), sex differences in the Chinese and Russian samples were examined before factor analysis was done. In the Russian sample, males outperformed females on all tests, with $\acute{\eta}^2$ ranging from 1% to 11% (see Esipenko et al., in this issue for more information). In the Chinese sample, significant sex differences were present for only three out of 10 tests: cross-sections (7%), mechanical reasoning (5%), and mazes (10%). In these tests, males outperformed females: cross-sections (9.69 vs. 7.95); mechanical reasoning (11.45 vs. 10.48); and mazes (6.30 vs 5.16), with Cohen's d ranging from 0.47 to 0.67.

Therefore, for further analysis, sex differences were regressed out from all tests for the Russian sample, and from the three tests for the Chinese sample.

Correlational analysis

First, we performed a Pearson's correlation analysis for the 10 King's Challenge tests, in both the Russian and Chinese samples. All correlations were positive, ranging from r = .15 to r = .80, and were significant after Bonferroni correction for multiple comparisons. The Elithorn maze and mazes tests were the least associated with the other tests (as well as to the total score). The exact correlation coefficients can be found in *Appendix, Table 1* (for the Russian sample) and *Appendix Table 2* (for the Chinese sample).

Factor analysis

Our main aim was to replicate the unifactorial structure reported by Rimfeld et al. (2017) in the unselected UK sample. We first examined the factorability of the 10 tests in the Russian sample. Several well-recognized criteria for the factorability of a factor analysis were used. First, it was observed that 10 out of 10 tests correlated at least .3 with at least one other test, suggesting reasonable factorability (see *Appendix, Table 1*). Secondly, the Kaiser-Meyer-Olkin measure of sampling adequacy turned out to be .93 (above the commonly recommended value of .6; Cerny & Kaiser, 1977); and Bartlett's test of sphericity was significant ($\chi 2\ (45) = 2881.98$, $p<.0001$). The diagonals of the anti-image correlation matrix were also all over .9. Finally, with the exception of the Elithorn maze, the communalities were all above .3 (*see Appendix, Table 3),* further confirming that each test shared some common variance with other tests. Given these overall indicators, factor analysis was deemed to be suitable for all 10 tests.

The PCA scree plot (see *Appendix, Figure 3*) and eigenvalues suggested single factor extraction (eigenvalue equals 4.59). Initial eigenvalues suggested that this factor explained 45.94 percent of the variance. Finally, the Root Mean Square Error of Approximation (RMSEA) of .03 (min = .017, max = .041, CI = .9), the Tucker Lewis Index (TLI) of .983, and the Comparative Fit Index (CFI) of .988 were also consistent with the single factor model. Confirmatory factor analysis showed that factor loadings ranged between .44 (for Elithorn's maze) to .76 (3D drawing), totally accounting for 40% of the overall variation (see *Table 2* for details on factor loadings). These results are consistent with the results from the unselected UK sample, where 42% of the overall variance was explained.

Table 2

Factor loadings for the King's Challenge subtests for both samples.

	Russian sample	Chinese sample	
Subtest name	Factor 1	Factor 1	Factor 2
Cross-sections	.65	.03	.68
2D drawing	.73	.57	.14
Pattern assembly	.56	.35	.23
Elithorn maze	.44	.64	–.05
Mechanical reasoning	.65	.02	.64
Paper folding	.71	.54	.29
3D drawing	.76	.53	.26
Shape rotation	.66	.74	–.09
Perspective taking	.55	.25	.15
Mazes	.52	.60	–.14
Variance explained by a factor	40%	26%	14%
Cumulative variance	40%	40%	

In the Chinese sample, eight of the 10 tests correlated at least .3 with at least one other test, suggesting reasonable factorability (see *Appendix, Table 2*). The Kaiser-Meyer-Olkin measure of sampling adequacy was .82, and Bartlett's test of sphericity was significant (χ^2 (45) = 313.71, $p < .001$). The diagonals of the anti-image correlation matrix were also all over .75. Finally, with the exception of pattern assembly and perspective-taking, the communalities were all above .3, further confirming that the tests shared some common variance with other tests (see *Appendix Table 3*). Given these overall indicators, factor analysis was deemed to be suitable for all 10 tests. Pattern assembly and perspective-taking were also added to the analysis to enable factor structure comparisons across the samples.

In the Chinese sample, scree plot and eigenvalues (3.35 and 1.19 for the first and second factor, respectively) suggested a two-factor structure of the data (see *Appendix, Figure 4*). The first two factors explained 33.5 and 11.9 percent of variance, respectively. Only two tests (cross-sections and mechanical reasoning) mostly loaded on the second factor. Oblimin rotation was used in principal component analysis, because the correlation between factors was .3. Moreover, without rotation, there were several tests which demonstrated cross-loadings of more than .4 to both factors. After rotation was applied, cross-loadings decreased to the level of .3, which clarified the relationships between the clusters of tests.

Confirmatory factor analysis also supported the two-factor model of SA in the Chinese sample. The RMSEA of a single-factor extraction was too high (.064, with min = .039, max = .086, CI = 0.9); while the two-factor model (the one with cross-sections and mechanical reasoning as a second factor) score for RMSEA was .037 (min = 0, max = .067, CI = 0.9). Also, ANOVA showed significant differences be-

tween the fit for the unifactorial and two-factorial models (χ^2 diff. = 10.85, p<.001), with the two-factor model showing a better fit ($\chi^2 = 39.99$) than the unifactorial model ($\chi^2 = 50.85$). Confirmatory factor analysis showed that the two-factor solution explained 40% of the overall variance, factor loadings are presented in *Table 2.*

Means comparison

The performance levels for the two samples were compared using t-tests, implementing the FDR correction for multiple comparisons (See *Appendix, Table 4*). As can be seen from *Figure 1*, the Russian and Chinese samples significantly differed in the total score and individual tests, with the exception of 2D drawing, 3D drawing, and mazes. The Chinese students on average outperformed Russian students on five tests (with Cohen's d ranging from .27 to .58, p < .001). The Russian students outperformed the Chinese sample on the Elithorn maze and perspective-taking (Cohen's d was .30 and .37 for perspective-taking and Elithorn maze, respectively; p<.001) (see *Figure 1*).

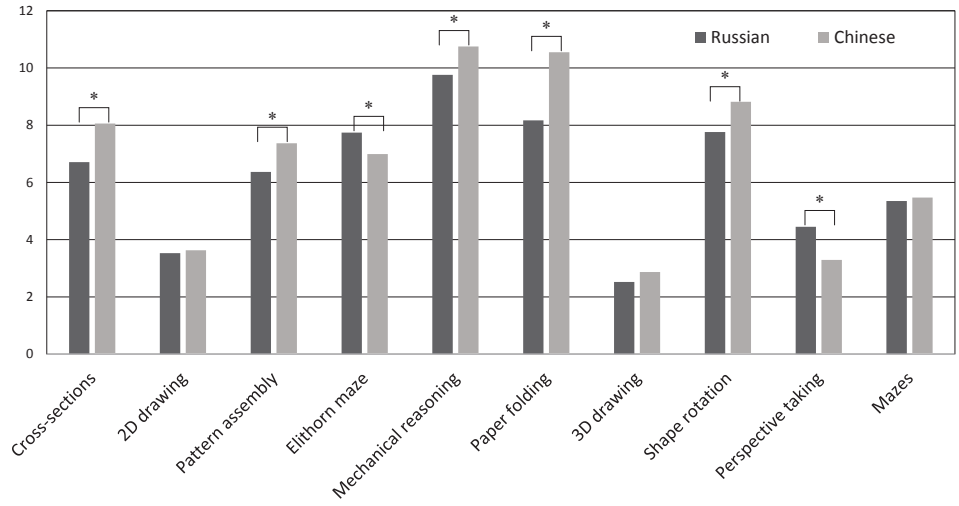


Figure 1. Mean comparisons for Chinese and Russian samples
*Note. * = significant difference at .01 level.*

Discussion

Our aims were: 1) to investigate the factorial structure of spatial abilities in Russian and Chinese samples; 2) to check whether our findings replicate the findings by Rimfeld and colleagues (2017); and 3) to explore the relative strengths and weaknesses in performance on the King's challenge battery across the two samples.

Overall, our data showed somewhat different factor structure for the Russian and Chinese samples. In the Russian sample, the unifactorial structure for both selected (unpublished data) and unselected UK samples was replicated. In the Chinese sample, a second factor emerged. The Russian, Chinese, and unselected UK samples (see *Fig. 2* in Rimfeld et al., 2017) showed very similar factor loadings, although the cross-sections and mechanical reasoning tests formed a separate fac-

tor in the Chinese sample. Several explanations may be proposed for the observed differences in the factorial structure.

The differences may be due to cultural and educational differences between the countries. These cultural differences may also explain the differences in average performance across the countries for these tests. Chinese students demonstrated higher scores for cross-sections and mechanical reasoning (8.06 and 10.75) than the Russian students (6.71 and 9.76).

The advantage in these skills by Chinese students might be partly explained by the features of the written Chinese language. It has previously been shown that both visual-orthographic processing and spatial analysis are essential for learning to read a Chinese character (Tan et al., 2005). Also, spatial ability was linked with writing in Chinese (Flaherty & Connolly, 1995). If this is the case, the cross-sections task may rely on an over-trained ability in the Chinese participants, as the task of retrieving part of the information from a unity (such as getting the shape of a section from a full figure) might be related to the need to compose a Chinese character out of several elements. It is possible that continuous engagement in such language processing leads to superior development of the relevant brain networks, which in turn leads to advantages in spatial ability and mathematics.

The mechanical reasoning test requires reading instructions for each item of the test, since the instructions differ from item to item. This might have led to the formation of separate "verbal" factor composed of these two tests. For other spatial ability tests in the battery, the instructions remained the same for all items. In addition, other tests, such as paper folding, do not rely on character-composing logic, as they do not involve operations with parts to compose a unity. Another test in the battery that may also rely on the ability to construct unity from several parts is pattern assembly. Therefore, this test can also be expected to rely on an underlying verbal component and to load on the verbal factor. In our study we did not observe this, which may be a result of the relatively small sample size.

Interestingly, pattern assembly (along with paper folding and 3D drawing) show relatively high loadings on the second factor (although not exceeding the level of .3 after rotation). This might suggest that these tests are not robustly loading on either of the factors. Furthermore, the loadings for perspective-taking were almost equal for the two factors, and did not exceed .3. Perspective taking is sometimes considered to be a large-scale test (Hegarty & Waller, 2004), and might have less links with other small-scale tests. This is also evident from its correlation coefficients with other tests in the battery in the Chinese sample (all correlations are less than .3). This pattern was not observed in the Russian sample.

The emergence of the second factor in the Chinese sample can be partly explained by the low split-half reliability for the mechanical reasoning test (.56 for the Russian and .54 for the Chinese samples). This may have led to error-related factor differences. Additional analyses on this test showed somewhat strange patterns of correlations between items within the test: some items correlated negatively with each other. This suggests low reliability for this test. Given that there was a time limit of 25 seconds for each item (and a need to read new instructions each time), participants with high spatial ability might have failed to give a correct answer in time, which might have led to the test having a reduced reliability.

To sum up, there could be indeed a second factor in spatial ability for the Chinese sample. The second factor might have emerged because the two tests (cross-sections and mechanical reasoning) are linked to verbal reasoning more than other tests. Further studies with more participants and other measures, e.g. verbal ability, are needed in order to improve our understanding of the SA factor structure in the Chinese population.

Our data also showed significant average differences in performance between the two samples. The Chinese students outperformed the Russian student on five out of 10 spatial tests (cross-sections, pattern assembly, mechanical reasoning, paper folding, and shape rotation). These results are consistent with previous research that demonstrated better average results by Chinese school children in mathematics and science (e.g., Programme for International Student assessment results; PISA; OECD, 2016). The mathematical advantage shown by children of all ages from Asian-Pacific countries has been very well established (Imbo & LeFevre, 2009; Mullis, Martin, & Foy, 2008). This advantage might continue to the university level and be reflected in spatial ability performance.

The observed average differences may also be related to the differences in the level of general intelligence and other relevant abilities between the two samples. The Chinese students were recruited from top-tier universities, where students are highly selected for admission based on their abilities (including mathematical reasoning). Although the Russian students were also recruited from top universities, the Russian selection criteria do not include such tests, but rather depend on the Unified State Exam performance, which varies for different majors/degrees.

Other reasons for the observed average differences may include school curricular differences. The lower general performance of the Russian participants can be related to the exclusion (since 2010) of technical drawing from the school National Curriculum. As a result, skills related to performance in some tests, such as cross-sections and mechanical reasoning, might have been under-trained in the students who participated in the current study. Moreover, in China, in contrast to Western countries, the abacus is still used to teach simple mathematical operations such as addition/subtraction. It has been shown that this leads to increased activation of the brain areas linked to spatial processing (Wu et al., 2009).

However, the Russian participants showed better performance in Elithorn maze and perspective-taking tasks than the Chinese participants. A possible explanation might be that some tests require greater flexibility (e.g., one needs to switch between different strategies to solve the task in the Elithorn maze), which was shown to be lower in Chinese population, in comparison with the Belgian and Canadian populations (Imbo & LeFevre, 2009).

Conclusion

Our data only partially replicated the findings of Rimfeld and colleagues (2017): while the unifactoral structure of spatial ability was replicated in the Russian sample, two factors emerged for the Chinese sample. The study also demonstrated some cross-cultural differences in spatial ability. More research is needed to replicate the pattern of relative strengths and weaknesses across the populations and to explore potential sources of these differences.

Limitations

One limitation of the present study was the relatively small size of the Chinese sample. In addition, some internet connection issues were reported by some participants, which led to a reduced dataset. The samples in the two countries were drawn from student populations from top universities in their countries. However, the selection criteria differed across the universities, and that could have led to differences across the samples in average intelligence and other relevant characteristics.

Acknowledgements

This work was supported by the Ministry of Education and Science of the Russian Federation within project No. 25.8566.2017/9.10

References

Aristova, I., Esipenko, E., Sharafieva, K., Maslennikova, E., Chipeeva, N., Feklicheva, I., … & Kovas, Y. (2018). Sovremennye podkhody v izuchenii prostranstvennykh sposobnostei: struktura i etiologia individual'nykh razlichii [Modern approaches in the study of spatial abilities: a structure and etiology of individual differences]. *Voprosy Psychologii* [Issues of psychology], 1, 118–126.

Blajenkova, O., Motes, M.A., & Kozhevnikov, M. (2005). Individual differences in the representations of novel environments. *Journal of Environmental Psychology*, *25*(1), 97–109. https://doi.org/10.1016/j.jenvp.2004.12.003

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge University Press. https://doi.org/10.1080/00140139508925174

Cerny, C.A. & Kaiser, H.F. (1977). A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate Behavioral Research*, *12*(1), 43-47. https://doi.org/10.1207/s15327906mbr1201_3

Comrey, A.L. & Lee, H.B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, N.J., US: Lawrence Erlbaum Associates, Inc.

Coutrot, A., Silva, R., Manley, E., de Cothi, W., Sami, S., Bohbot, V., … & Hornberger, M. (2017). Planetary-wide organization of human cognition. *bioRxiv,* 188870. http://dx.doi.org/10.1101/188870

De Beni, R., Pazzaglia, F., & Gardini, S. (2006). The role of mental rotation and age in spatial perspective-taking tasks: when age does not impair perspective-taking performance. *Applied cognitive psychology*, *20*(6), 807–821. https://doi.org/10.1002/acp.1229

Esipenko, E., Maslennikova, E., Budakova, A., Sharafieva, K., Ismatullina, V., Feklicheva, I., … & Kovas, Y. (2018) Spatial ability of male and female students completing humanities vs. technical degrees. *Psychology in Russia: State of the Art, 11*(4), 40–52. https://doi.org/10.11621/pir.2018.0403

Flaherty, M. & Connolly, M. (1995). Space perception, coordination and a knowledge of Kanji in Japanese and non-Japanese. *Psychologia*, 38, 229–237.

Gitterman, M.R. & Sies, L.F. (1992). Nonbiological determinants of the organization of language in the brain: A comment on Hu, Qiou and Zhong. *Brain and Language*, 43, 162–165. https://doi.org/10.1016/0093-934X(92)90027-C

Gogos, A., Gavrilescu, M., Davison, S., Searle, K., Adams, J., Rossell, S. L., ... & Egan, G.F. (2010). Greater superior than inferior parietal lobule activation with increasing rotation

angle during mental rotation: an fMRI study. *Neuropsychologia*, *48*(2), 529–535. https://doi.org/10.1016/j.neuropsychologia.2009.10.013

Hegarty, M. & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. https://doi.org/10.1016/j.intell.2003.12.001

Hegarty, M., Montello, D.R., Richardson, A.E., Ishikawa, T., & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34(2), 151–176. https://doi.org/10.1016/j.intell.2005.09.005

Hugdahl, K., Thomsen, T., & Ersland, L. (2006). Sex differences in visuo-spatial processing: an fMRI study of mental rotation. *Neuropsychologia*, 44(9), 1575–1583. https://doi.org/10.1016/j.neuropsychologia.2006.01.026

Imbo, I. & LeFevre, J.A. (2010). The role of phonological and visual working memory in complex arithmetic for Chinese-and Canadian-educated adults. *Memory & Cognition*, 38(2), 176–185. https://doi.org/10.3758/MC.38.2.176

Jansen, P (2009). The dissociation of small- and large-scale spatial abilities in school-age children. *Perceptual and Motor Skills*, *109*(2), 357–361. https://doi.org/10.2466/PMS.109.2.357-361

Jansen, P., Schmelter, A., & Heil, M. (2010). Spatial knowledge acquisition in younger and elderly adults: A study in a virtual environment. *Experimental Psychology*, *57*(1), 54. https://doi.org/10.1027/1618-3169/a000007

Jansen, P., Wiedenbauer, G., & Hahn, N. (2010). Manual rotation training improves direction-estimations in a virtual environmental space. *European Journal of Cognitive Psychology*, *22*(1), 6–17. http://dx.doi.org/10.1080/09541440802678487

Kell, H., Lubinski, D., Benbow, C., & Steiger, J. (2013). Creativity and technical innovation: spatial ability's unique role. *Psychol Sci*. *24*(9), 1831–1836. https://doi.org/10.1177/0956797613478615

Kosslyn, S.M. & Thompson, W.L. (2003). When is early visual cortex activated during visual mental imagery? *Psychological bulletin*, *129*(5), 723. https://doi.org/10.1037/0033-2909.129.5.723

Kozhevnikov, M. & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, *29*(5), 745–756. https://doi.org/10.1016/j.intell.2003.12.001

Kozhevnikov, M., Motes, M.A., Rasch, B., & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology*, *20*(3), 397–417. https://doi.org/10.1002/acp.1192

Lohman, D.F. (1979). *Spatial Ability: A Review and Reanalysis of the Correlational Literature (No. TR-8)*. Stanford, California: Stanford University School of Education

Lohman, D.F., Pellegrino. J.W., Alderton, D.L., & Regian, J.W. (1987). Dimensions and components of individual differences in spatial abilities. In S.H. Irvin & S.E. Newstead (Eds.), *Intelligence and Cognition: Contemporary frames of reference*. Boston: Martinus Nijhoff.

Lohman, D.F. (1996). Spatial ability and g. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 97–116).

Lohman, D.F. & Nichols, P.D. (1990). Training spatial abilities: Effects of practice on rotation and synthesis tasks. *Learning and Individual Differences*, *2*(1), 67–93.

McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological bulletin*, *86*(5), 889.

Morris, R.G. & Parslow, D.M. (2004). Neurocognitive components of spatial memory. In G. Allen (Ed.), *Human spatial memory: Remembering where*, (pp. 217–247). Lawrence Erlbaum Associates, Mahwah.

Mullis, I., Martin, M.O., & Foy, P. (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Eighth and Fourth Grades*.

OECD (2016), PISA 2015 Results (Volume I): Excellence and Equity in Education, PISA, OECD Publishing, Paris. Retrieved from http://dx.doi.org/10.1787/9789264266490-en

Parslow, D.M., Rose, D., Brooks, B., Fleminger, S., Gray, J.A., Giampietro, V., … & Morris, R.G. (2004). Allocentric spatial memory activation of the hippocampal formation measured with fMRI. *Neuropsychology, 18*(3), 450–61. https://doi.org/10.1037/0894-4105.18.3.450

Revelle, W. & Rocklin, T. (1979). Very simple structure: An alternative procedure for estimating the optimal number of interpretable factors. *Multivariate Behavioral Research*, *14*(4), 403–414. https://doi.org/10.1207/s15327906mbr1404_2

Rimfeld, K., Shakeshaft, N.G., Malanchini, M., Rodic, M., Selzam, S., Schofield, K., ... & Plomin, R. (2017). Phenotypic and genetic evidence for a unifactorial structure of spatial abilities. *Proceedings of the National Academy of Sciences*, *114*(10), 2777–2782. https://doi.org/10.1073/pnas.1607883114

Rodic, M., Tikhomirova. T., Kolienko, T., Malykh, S., Bogdanova, O., Zueva, D.Y., … & Kovas, Y. (2015) Spatial complexity of character based writing systems and arithmetic in primary school: a longitudinal study. *Front. Psychol., 6*, 333. https://doi.org/10.3389/fpsyg.2015.00333

Sakamoto, M. & Spiers, M.V. (2014). Sex and cultural differences in spatial performance between Japanese and North Americans. *Archives of sexual behavior*, *43*(3), 483–491. https://doi.org/10.1007/s10508-013-0232-8

Shakeshaft, N.G., Rimfeld, K., Schofield, K.L., Selzam, S., Malanchini, M., Rodic, M., … & Plomin, R. (2016). Rotation is visualization, 3D is 2D: using a novel measure to investigate the genetics of spatial ability. *Scientific reports*, 6. https://dx.doi.org/10.1038%2Fsrep30545

Shen, C., Vasilyeva, M., & Laski, E.V. (2016). Here, but not there: Cross-national variability of gender effects in arithmetic. *Journal of experimental child psychology*, 146, 50–65. https://doi.org/10.1016/j.jecp.2016.01.016

Shepard, R.N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703. https://doi.org/10.1126/science.171.3972.701

State Federal Educational standard for school (2010, November 11). Retrieved from https://xn--80abucjiibhv9a.xn--p1ai/%D0%B4%D0%BE%D0%BA%D1%83%D0%BC%D0%B5%D0%BD%D1%82%D1%8B/938

Tan, L.H., Spinks, J.A., Eden, G., Perfetti, C.A., & Siok, W.T. (2005). Reading depends on writing, in Chinese. *Proc. Natl. Acad. Sci. U.S.A.*, *102,* 8781–8785 https://doi.org/10.1073/pnas.0503523102

Toivainen, T., Pannini, G., Papageorgiou, K.A., Malanchini, M., Rimfeld, K., Shakeshaft, N., & Kovas, Y. (2018). Prenatal testosterone does not explain sex differences in spatial ability. *Sci Rep*., *8*(1), 13653. https://doi.org/10.1038/s41598-018-31704-y

Tosto, M.G., Hanscombe, K.B., Haworth, C.M., Davis, O.S., Petrill, S.A., Dale, P.S., & Kovas, Y. (2014). Why do spatial abilities predict mathematical performance? *Developmental science*, *17*(3), 462–470. https://dx.doi.org/10.1111%2Fdesc.12138

Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R.,Warren, C., & Newcombe, N.S. The malleability of spatial skills: a meta-analysis of training studies. *Psychol Bull., 139*(2), 352–402. doi: 10.1037/a0028446. Epub 2012 Jun 4. https://doi.org/10.1037/a0028446

Wang, Y. (2014). Unveiling the Cognitive Mechanisms of Eyes: The Visual Sensor Vs. the Perceptive Browser of the Brain. *International Journal of Cognitive Informatics and Natural Intelligence* (IJCINI), *8*(1), 36-50. https://doi.org/10.4018/ijcini.2014010103

Wang, L., Cohen, A., & Carr, M. (2014). Spatial Ability at Two Scales of Representation: A Meta-Analysis. *Learning and Individual Differences, 36*(1), 140–144 https://doi.org/10.1016/j.lindif.2014.10.006

Wu, T.H., Chen, C.L., Huang, Y.H., Liu, R.S., Hsieh, J.C., & Lee, J.J. (2009). Effects of long-term practice and task complexity on brain activities when performing abacus-based mental calculations: a PET study. *Eur. J. Nucl. Med. Mol. Imaging, 36*, 436–445 https://doi.org/10.1007/s00259-008-0949-0

Zacks, J.M., Mires, J., Tversky, B., & Hazeltine, E. (2000). Mental spatial transformations of objects and perspective. *Spatial Cognition and Computation*, *2*(4), 315–332. https://doi.org/10.1023/A:1015584100204

Original manuscript received August 3, 2018
Revised manuscript accepted October 14, 2018
First published online December 30, 2018

Appendix

Table 1

Correlations for King's Challenge (KC) tests (Russia)

	1	2	3	4	5	6	7	8	9	10
1. Cross-sections	-									
2. 2D drawing	.47	-								
3. Pattern assembly	.38	.41	-							
4. Elithorn maze	.27	.34	.22	-						
5. Mechanical reasoning	.46	.45	.35	.31	-					
6. Paper folding	.49	.49	.39	.30	.46	-				
7. 3D drawing	.47	.60	.39	.32	.48	.55	-			
8. Shape rotation	.41	.47	.39	.25	.42	.49	.49	-		
9. Perspective taking	.34	.40	.29	.25	.34	.38	.46	.39	-	
10. Mazes	.29	.35	.34	.26	.35	.36	.38	.34	.27	-
11. Total score	.70	.69	.62	.48	.68	.77	.74	.73	.65	.55

Note. All correlations were significant after Bonferroni correction ($p<0.001$)

Table 2

Correlations for King's Challenge (KC) tests (China)

	1	2	3	4	5	6	7	8	9	10
1.Cross-sections										
2. 2D drawing	.32**									
3. Pattern assembly	.36**	.32**								
4. Elithorn maze	.28**	.41**	.32**							
5. Mechanical reasoning	.45**	.30**	.29**	.22**						
6. Paper folding	.43**	.59**	.34**	.39**	.42**					
7. 3D drawing	.40**	.52**	.35**	.36**	.40**	.52**				
8. Shape rotation	.27**	.36**	.39**	.47**	.25**	.45**	.46**			
9. Perspective taking	.23**	.25**	.20**	.17*	.18*	.25**	.24**	.27**		
10.Mazes	.16*	.30**	.22**	.31**	.15*	.36**	.36**	.36**	.15*	
11. Total score	.70**	.70**	.60**	.60**	.70**	.80**	.68**	.70**	.51**	.58**

*Note. ** Correlation is significant at the 0.01 level (2-tailed) after Bonferroni correction, * Correlation is significant at the 0.05 level (2-tailed) after Bonferroni correction.*

Table 3

Communalities for Russian and Chinese sample

	Russian	**Chinese**
Cross-sections	.483	.637
2D drawing	.578	.504
Pattern assembly	.372	.218
Elithorn maze	.262	.484
Mechanical reasoning	.502	.584
Paper folding	.547	.580
3D drawing	.621	.436
Shape rotation	.492	.574
Perspective taking	.395	.096
Mazes	.342	.430

Table 4.

Mean comparison (t-stat counted as Russian sample vs. Chinese)

	N (Ch)	**Mean (Ch)**	**SD (Ch)**	**N (Rus)**	**M (Rus)**	**SD (Rus)**	**t-stat**	**d Cohen**	**adjusted p-value**
Cross-sections	229	8.06	3.74	922	6.71	3.65	-4.89	0.37	.00
2D drawing	229	3.63	1.19	918	3.53	1.23	-1.19	0.09	1.00
Pattern assembly	229	7.37	3.10	909	6.37	3.04	-4.40	0.33	.00
Elithorn maze	227	6.99	2.15	904	7.74	1.97	4.79	0.37	.00
Mechanical reasoning	225	10.75	2.43	902	9.76	2.70	-5.37	0.38	.00
Paper folding	224	10.55	3.54	897	8.17	4.25	-8.63	0.58	.00
3D drawing	224	2.87	1.80	889	2.52	1.82	-2.60	0.19	.15
Shape rotation	223	8.82	3.71	886	7.76	4.05	-3.74	0.27	.00
Perspective taking	223	3.29	3.07	885	4.45	4.14	4.70	0.30	.00
Mazes	222	5.47	2.06	883	5.35	2.03	-0.78	0.06	1.00
Total	222	68.24	17.94	883	62.40	19.66	-4.26	0.30	.00

Note: significant differences between samples are in bold

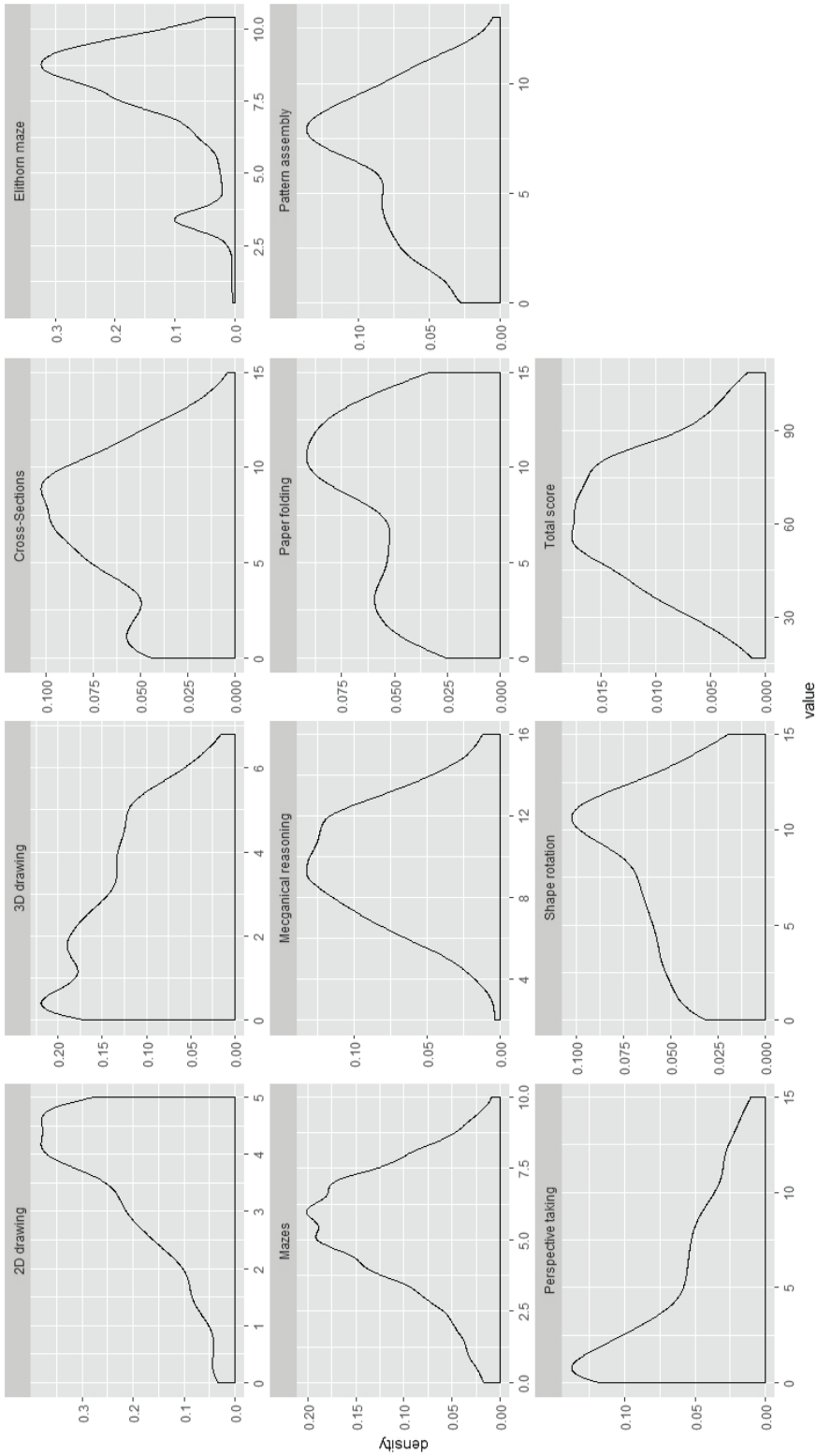


Figure 1. Density plot distributions for all tests (Russian sample)

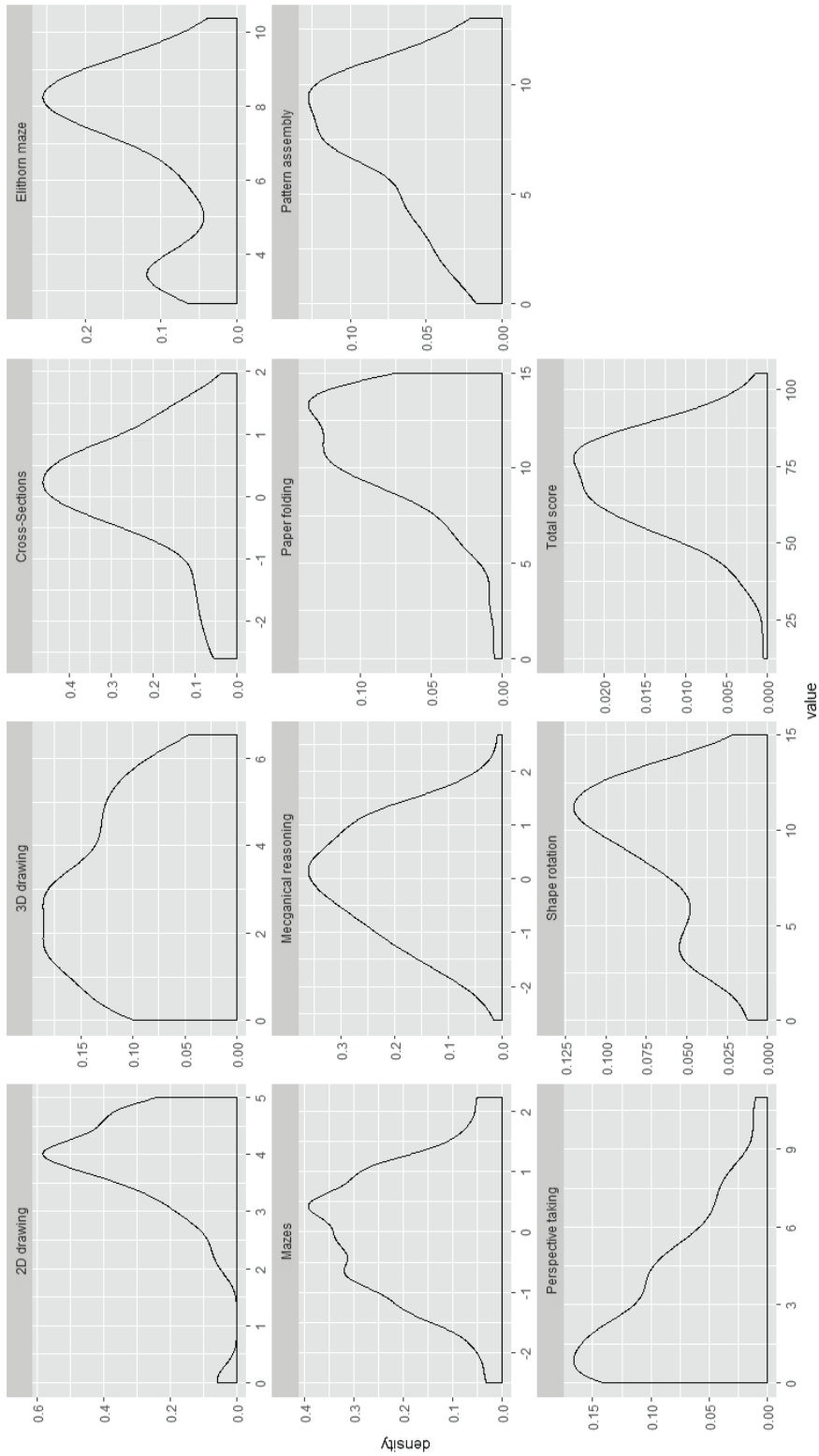


Figure 2. Density plot distributions for all tests (Chinese sample)

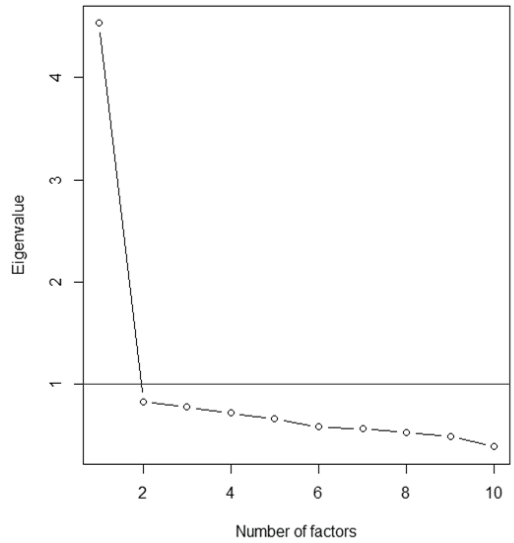


Figure 3. Screeplot for PCA on KC subtests for Russian sample

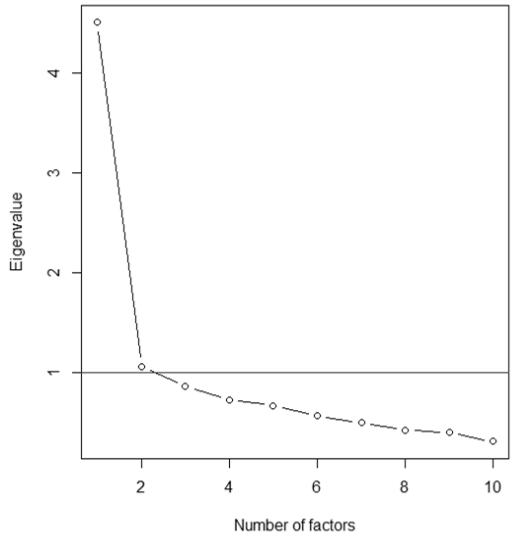


Figure 4. Screeplot for PCA on KC subtests for Chinese sample (oblimin rotation used)

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

Russian Psychological Society
Lomonosov Moscow State University

Language Proficiency in Preschool Children with Different Levels of Executive Function

Aleksander N. Veraksa*, Daria A. Bukhalenkova, Maria S. Kovyazina

Faculty of Psychology, Lomonosov Moscow State University, Moscow, Russia

* Corresponding author. E-mail: veraksa@yandex.ru

Background. According to numerous studies, people's development of executive function is a predictor of their successful acquisition of literacy skills. However, the data on the relationship between the development of verbal language and executive function in preschool aged children are insufficient and contradictory.

Objective. The goal of our research was to study the connection between the three main EF components (working memory, inhibition, and cognitive flexibility) and various spoken language skills in children of senior preschool age. It is the first stage of a longitudinal study aimed at understanding the relationship between executive function and language development starting from ages 5–6, and proceeding through elementary school.

Design. Our study sample included 279 children aged 5–6 years (M = 5.6 years) attending a senior group in Moscow kindergartens (139 boys and 140 girls). The study used NEPSY-II diagnostic complex subtests and the Dimensional Change Card Sort (DCCS) test to measure the level of executive functions (working memory, cognitive flexibility, and inhibition). Language development (vocabulary, phonemic awareness, and word generation) was measured by neuropsychological methods (Akhutina, Pylaeva, 2015).

Results. The results of the study showed significant associations between all EF components and language skills development in preschool children. Oral language skills were more closely related to the level of development of verbal working memory and cognitive flexibility than they were to inhibition or visual working memory. Children with low levels of EF development were significantly less able to cope with tasks such as understanding prepositional structures, understanding similar sounding words, and showing verbal fluency, than children with a high EF level. Furthermore, children with normal and high levels of EF development displayed no significant differences in language development. Thus, the study showed that children with a low level of EF have difficulties with language development.

Conclusion. Our results provide important details about understanding the relationship between executive functioning and language development in children of senior preschool age.

Keywords: Preschool age, executive function (EF), language, vocabulary, phonemic awareness, word generation.

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

doi: 10.11621/pir.2018.0408
http://psychologyinrussia.com

Introduction

The levels of executive function (EF) and language development are major indicators of development in preschool children (Vygotsky, 1984), forming the basis for school readiness (Elkonin, 2006; Matthews, Ponitz, & Morrison, 2009), and serving as predictors of further academic achievement (Blair & Razza, 2007; Duff et al., 2015; Duncan et al., 2007; Willoughby et al., 2012).

According to numerous studies, EF development is a predictor of the successful acquisition of reading and writing skills (Aaron et al., 2008; Cutting et al., 2009; Cantin et al., 2016; Yeniad et al., 2013). However, the data on the relationship between verbal speech and EF in children of preschool age are insufficient and contradictory (Pazeto, 2014). There are ongoing efforts to do theoretical and empirical research in this area, as evidence is building up in favor of the view that the level of EF predicts language development (Blair et al., 2012, Henry, Messer, & Nash, 2012; Verhagen & Leseman, 2016). Conversely, there are longitudinal studies which have shown that expressive language in childhood predicts later development of EF (Kuhn et al., 2014; Petersen et al., 2013).

Executive function

One of the most significant models for understanding EF development has been proposed by Miyake and his colleagues (Miyake et al., 2000). According to this model, the neuropsychological basis for mastering one's behavior is formed by a group of cognitive skills that enable targeted problem-solving and adaptive behavior in new situations. According to this approach, EF is divided into the following three main components: 1) working memory; 2) cognitive flexibility, which is associated with the ability to switch from one rule to another; and 3) inhibitory control, which is supposed to inhibit one's dominant response to a situation in favor of what a task requires. Although originally based on the results obtained for adults, the applicability of this model to describe child development has been confirmed in a number of studies (Diamond & Lee, 2011; Lehto et al., 2003, Visu-Petra et al., 2013; Almazova et al., 2016).

Connection between EF and language skills

The existence of a link between EF and language development in preschool age children has been shown in a number of correlation studies (Blair & Razza, 2007; Matthews, Ponitz, & Morrison, 2009). For example, Blair and Razza (2007) analyzed the relationship between two EF components (inhibition and switching), on the one hand, and various academic skills, including speech, on the other. The study involved 170 children aged 3-5 years. The development of inhibition proved to be significantly related to a child's vocabulary, phonological processing abilities, and letter knowledge, while switching turned out to be associated with vocabulary only. The researchers suggested that these outcomes may support the idea that inhibition is instrumental in the development of academic skills for children of preschool age (Blair & Razza, 2007).

A study by Matthews and colleagues (2009), which was conducted on a sample of 268 3–5 year-old children, revealed a correlation between self-regulation (a multicomponent variable which depends on cognitive skills, including working mem-

ory, attention control and switching, and inhibitory control) and sound awareness, whereas its link to vocabulary development (Picture Naming Test) was not established (Matthews, Ponitz, & Morrison, 2009). Many studies have also shown a link between executive function, and reading and writing skills in the senior preschool and junior school years (Cutting et al., 2009; Diamond, 2006; Duncan et al., 2007)

The relationship between EF and language development in preschool-aged children has also been revealed by longitudinal studies. Fuhs and Day (2011) studied the dynamics of EF and language development in 132 children aged 4-5 years for one academic year. At the beginning and end of the school year, they diagnosed the development of two EF components (inhibition and switching) and assessed the children's school readiness. When processing the autumn and spring data separately, they found a significant connection between both EF components and the receptive vocabulary. However, none of the EF components were associated with the results of the tests measuring expressive vocabulary (as shown in the Picture Naming subtest of the WPPSI-III). The study also showed that children with a receptive vocabulary and a high level of phonological awareness (according to the WPPSI) are significantly more successful in developing inhibition and switching abilities than children with average and low levels of verbal ability (Fuhs & Day, 2011).

In a study by Fuhs and colleagues (2014), 562 four-year-old children were assessed at the beginning and end of their prekindergarten year, and followed to the end of kindergarten. Researchers report that strong bidirectional associations were found for the generalized indicators of EF and oral comprehension skills, but not for literacy skills in the pre-K year. In addition, after controlling for pre-K gains in both EF and achievement, EF skills continued to be a moderate predictor of kindergarten language gains (Fuhs, Nesbitt, & Farran, 2014).

Evidence of such association discovered later in a study by Pazeto and colleagues (2014) analyzed the dynamic relationship between EF and language development in children of preschool age. The study involved 90 children aged 4-5 years who were tested in their final year of kindergarten and in the first grade. The study revealed significant changes in the level of spoken language competence and reading and writing skills, while the level of development of all EF components, except for voluntary attention, remained more or less unchanged. This result shows that language skills develop faster than EF does in children of preschool age. Correlation analysis showed that the levels of EF and spoken language competence are significantly associated with the development of reading and writing skills, whereas EF and spoken language skills appear to be relatively more independent of each other.

Similar results showing the dynamic relationship between EF and language development were obtained in longitudinal studies by Bohlmann and colleagues (2015) and by Slot and Suchodoletz (2018). Bohlmann and colleagues revealed a link between the vocabulary and self-regulation skills in 250 children aged 3-5 years. The preschool children underwent three measurement waves during a period of two years. The results showed that there is a bidirectional relationship between the variables studied: the larger the child's vocabulary had been before the study began, the greater the results he or she achieved in the development of cognitive control during the study period; likewise, the level of executive control development determined the intensity of vocabulary growth. The study by Slot and Suchodoletz (2018) was aimed at assessing the dynamics of language and EF development in

227 children aged 3-4 years. When they entered kindergarten, all the children had their actual level of language and EF development measured, and these levels were later compared with the end-of-school year data. A bidirectional association was found to exist between the development of language and of EF. The authors concluded that language development is important for stimulating the development of EF (Slot & Suchodoletz, 2018).

A number of studies have shown a connection between EF and children's vocabulary (Blair & Razza, 2007; Fuhs & Day, 2011; Bohlmann, Maier, & Palacios, 2015). However, some of the results are contradictory (Matthews, Ponitz, & Morrison, 2009). It is interesting to note that the picture-naming task often tends to be unrelated to EF components (Matthews, Ponitz, & Morrison, 2009; Fuhs & Day, 2011). Other studies found a statistically significant connection between listening comprehension and EF at the preschool age (Blair & Razza, 2007; Matthews, Ponitz, & Morrison, 2009; Fuhs, Nesbitt, & Farran, 2014), unlike a large number of studies showing the relationship between EF and reading and writing skills (Cutting et al., 2009; Diamond, 2006; Duncan et al., 2007). However, the mechanism of this association has not been studied sufficiently (Pazeto et al., 2014). It is also noteworthy that many studies (e.g., Blair & Razza, 2007; Fuhs & Day, 2011) consider only two EF components–inhibition and switching–whereas the connection between working memory and language skills has been studied to a lesser extent.

Consequently, most of the longitudinal studies have established a bidirectional relationship between EF and language development in preschool children (Zelazo et al., 2003; Fuhs & Day, 2011; Slot & Suchodoletz, 2018). These data are also supported by the research of Henry and colleagues (2012), which showed that children with language impairments had much lower EF developmental outcomes (verbal and nonverbal working memory, inhibition and planning) than those with normative language development parameters. Researchers note that difficulties remained even after the diagnostic tasks were adapted to match the children's speech abilities (Henry, Messer, & Nash, 2012).

Another group of studies found a significant predictive relationship between children's EF measured in the fall of preschool, and their receptive vocabulary measured in the following spring (Blair & Razza, 2007; Montgomery, Magimairaj, & Finney, 2010; Rose, Feldman, & Jankowski, 2009; Verhagen & Leseman, 2016; Weiland et al. 2014). For example, the study by Rojas-Barahona and colleagues (2015) found statistically significant changes in the rate of language development taking place after children were given classes designed to develop their working memory.

The third point of view is based on the model that sees language and EF development as mutually stimulating processes (Bohlmann, Maier, & Palacios, 2015), a view which is consistent with the cultural-historical approach to understanding language development and self-regulation in child development (Vygotsky, 1962; Luria, 1976; Reshetova, 2017).

The data presented in this paper are based on the first wave of a longitudinal study aimed at understanding the EF-language relationship in children starting from ages 5-6, and continuing through elementary school. The goal of this research was to study the connection between all three main EF components (working memory, inhibition, and cognitive flexibility) and various spoken language skills in senior preschool age.

Methods

The sample and procedure of the study

Our sample included 279 children aged 5-6 years (M=5.6 years) who were attending a senior group in Moscow kindergartens (139 boys and 140 girls). All tasks were carried out by the children in the spring of 2017, individually, in a quiet room. The parents of the preschool children gave their written informed consent for their children's participation in the study. The procedure was approved by the Board of Ethics, Faculty of Psychology, Lomonosov Moscow State University.

Measures of EF

The test battery consisted of tasks that explored three main EF components in children: working memory, cognitive flexibility, and inhibitory control. Most of methods used in the study were Russian versions of subtests from the neuropsychological complex NEPSY-II, which is designed to evaluate children's mental development within the age bracket of 3-16 years (Korkman, et al., 2007). Previous research has shown the possibility of using these tests on Russian preschool children (Almazova, Bukhalenkova, & Veraksa, 2016).

To assess verbal working memory, a subtest of the NEPSY-II *Sentence Repetition (SR*) was used. This subtest is designed to assess the ability to repeat sentences of increasing complexity and length. A series of sentences was read to a child, who was then asked to reproduce each sentence immediately upon its presentation. Each correct repetition of the sentence scores two points. If the child makes one or two mistakes, the sentence is scored one point; if more mistakes are made, the sentence is scored 0 points.

To estimate visual working memory, a subtest of the NEPSY-II *Memory for Designs (MD)* was used. It is designed to assess spatial memory for novel visual material. The child is shown a grid with four to eight designs on a page, which is then removed from his or her view (subtests 2-5). The child selects the designs from a set of cards, and places the cards on the grid in the same location as was shown previously. Points are scored separately for correctly remembering the spatial (max=48) and content (max=24) characteristics of the images. Bonus points are scored for correctly remembering the integral configurations of the objects and image details (max=48).

To assess the level of cognitive flexibility, the *Dimensional Change Card Sort (DCCS)* method was used (Zelazo, 2006). The children are required to sort a series of bivalent test cards, first according to one dimension (color), and then according to another (shape). On the third sorting, a child had to sort the cards according to an additional factor (cards with borders/cards without borders). The test measures the child's ability to regulate his/her behavior according to complex rules, thereby assessing the level of the child's cognitive flexibility.

To assess the level of inhibition development, the *Inhibition* subtest of the NEPSY-II was used. It is designed to assess the ability to inhibit automatic responses in favor of novel responses, and the ability to switch between types of responses. During the first step, the child looks at a series of shapes (squares and circles) and names the shape as quickly as possible (Naming).Then, having named the shapes, he/she has to switch to an alternate response ("circle" instead of "square" and vice

versa) (Inhibition). The researcher calculates the number of the children's corrected and uncorrected mistakes, and the time it takes them to complete each phase.

Measures of language development

We measured the preschoolers' vocabulary (naming objects and actions, understanding of logical and grammatical constructions), phonemic awareness (understanding similar sounding words), and word generation (the ability to actualize words).

The *Picture Naming* subtest of the WPPSI (2006) (Rzhanova et al., 2018) is aimed at measuring the preschoolers' vocabulary. This technique consists of 18 pictures depicting various images that are shown to the child one by one, which he/she should identify. The child is awarded one point for each correctly performed task. At the end, the total score is calculated for each task (the maximum amount being 18 points).

The *Naming of Actions* (Akhutina, 2016) aims at measuring the preschoolers' vocabulary (the nominative function of speech). It consists of 15 pictures depicting different actions. The child is shown pages with images and asked to name what it is depicted there with one word. The child is awarded three points for giving the exact name; two points for answering by using several words; one point for similar verbal substitutions or distortion of the sound structure of the word; and zero points for an incorrect answer or a noun substitution for a verb. Then the total score is calculated for all the tasks (45 points at most).

The method of *Understanding of Logical Grammatical Constructions* (ULGC) (Akhutina, 2016) begins with examining the children's understanding of active and passive constructions with direct and reverse word order. The child is offered a sheet with pairs of pictures depicting variants of reversible situations (for example, "a girl is caught by a boy"). The researcher reads the sentence and asks the child to show the picture that matches it. Overall, seven sentences are presented, and one point is awarded for every correctly performed task. The second part examines the understanding of prepositional constructions. The material consists of a set of pictures depicting different mutual spatial arrangements of a box and a keg. The child is asked to show which picture the arrangement of the objects matches (for example, "there's a keg behind the box"). Overall, six prepositional constructions are presented.

The method of *Understanding of Similar Sounding Words* (USSW) (Akhutina, 2016) was used to measure a child's phonemic awareness and verbal working memory. The child is given two sheets that have ten pictures depicting objects with similar sounding names. At first, the child is asked to name all the objects. Then the child is given the names of several objects, which he or she is to memorize and show in the same order in which they were presented. The number of items named gradually increases from two to six. Children aged 4-5 years are given tests that grow in complexity until they make mistakes in three tasks in a row. In each assignment, the child is awarded one point for every correctly identified picture (productivity). Additional points are also awarded for duplications (if the child gives the right word and a similar sounding one, he gets one point), for changed word order (one point), omissions (one point), and superfluous words (one point).

The *Verbal Fluency Test* (VFT) (Akhutina, 2016) is aimed at measuring the word actualization process and verbal fluency. It consists of three tasks: the first one is based on free associations (a child is asked to give as many words as possible within a minute). Tasks two and three involve directed associations: in two, the child is asked to name a series of actions for a minute, and in three, to name a series of animals (While the original test, aimed at elementary school age children, called for naming a series of plants, we replaced plants with animals, as more appropriate for preschool children). A similar task is found in NEPSY-II (Korkman et al., 2007). In each task, we calculated the number of productive associations, i.e., all words without repetition and inadequate words.

Results

Descriptive statistics

Since the validation of language assessment was carried out on first-graders, let us first consider the results of these tests for 5-6 years old children (See *Table 1*).

Table 1

Descriptive statistics of language skills measurements in senior preschool age

	Mean	Standard deviation
Picture Naming	13.09	1.970
Naming of Actions	8.73	3.073
ULGC, active, passive constructions	4.57	1.430
ULGC, prepositions	2.68	1.420
USSW, productivity	17.45	7.295
VFT , free associations	17.06	7.344
VFT , actions	7.91	3.517
VFT, animals	11.89	4.525

We observe big standard deviation in the Understanding of Similar Sounding Words test (USSW) implementation that indicates the differences in the development of the phonemic awareness or verbal working memory of the preschoolers. The results of the Picture Naming test and the Understanding of Logical Grammatical Constructions method (ULGC) indicate about the same level of development of children's vocabulary. At the same time, the children vary in performing word generation tasks: naming animals turns out to be easier than naming actions.

Analysis of the correlation between the indicators of EF and language skills development

Since one of the tasks at this stage was to examine which EF components were associated with different language skills, we turned to the correlation analysis of the data (See *Table 2*).

Table 2

Results of the correlation analysis of EF and language skills measurements (Spearman's coefficient and significance level)

	Visual WM				Flexibility	Verbal WM	Inhibition				
	MD Content	**MD Spatial**	**MD Bonus**	**MD Total Score**	**DCCS**	**SR**	**Naming Uncor. Errors**	**Naming time**	**Inhibition Uncor. Errors**	**Inhibition Cor. Errors**	**Inhibition time**
Picture Naming	.110	.012	.029	.054	**.147***	**.363****	–.066	–.090	**–.125***	–.084	
	.072	.847	.633	.380	**.014**	**.000**	.281	.143	**.042**	.171	
Naming of Actions	–.076	.019	.027	–.004	.098	.075	.003	–.078	**–.142***	.114	
	.211	.751	.662	.942	.102	.209	.961	.203	**.021**	.062	
ULGC, active, passive constructions	**.139***	.074	.093	.113	**.252****	.317**	–.165**	–.180**	–.139*	.068	**–.180****
	.023	.224	.128	.064	**.000**	**.000**	**.007**	**.003**	**.023**	.271	.003
ULGC Prepositions	.132*	.140*	.126*	.148*	.270**	**.354****	–.077	–.052	**–.162****	**–.124***	**.132***
	.031	**.022**	**.038**	.015	**.000**	**.000**	.209	.398	**.008**	**.043**	**.031**
USSW, productivity	.177**	**.166****	**.157***	**.185****	**.321****	**.526****	**–.190****	–.110	**–.210****	.001	–.094
	.004	**.006**	**.010**	.002	**.000**	**.000**	**.002**	.072	**.001**	.989	.128
VF , free associations	.074	**.128***	**.142***	**.129***	**.133***	**.237****	–.084	**–.219****	.032	.097	–.106
	.248	**.047**	**.028**	**.046**	**.035**	**.000**	.197	**.001**	.624	.135	.103
VFT, actions	.123	.093	.110	.114	.049	**.279****	–.063	**–.135***	–.054	**.160***	**–.132***
	.052	.142	.083	.072	.436	**.000**	.324	**.034**	.398	.012	.038
VFT, animals	.078	.067	.072	.082	**.167****	**.314****	–.041	**–.160****	–.009	–.052	**–148***
	.211	.283	.249	.190	**.006**	**.000**	.510	**.010**	.881	.406	**.018**

*Note. ** correlation is significant at the level of 0.01, * correlation is significant at the level of 0.05*

The results of both naming tasks (Picture Naming and Naming of Actions) are significantly associated with the level of development of verbal working memory (SR), and they are inversely correlated with the number of mistakes in the Inhibition method (the fewer mistakes children make, the better they cope with the naming task), which indicates the link with the ability to self-correct. It is interesting to note that only the Picture Naming subtest has a significant correlation with cognitive flexibility (DCCS).

A large number of significant correlations were found between the understanding of logical grammatical constructions (ULGC) and tasks on cognitive flexibility (DCCS), verbal working memory (SR), and inhibition. It is noteworthy that the understanding of active and passive constructions correlates only with the aspect of image detail memorization (MD Content) rather than the overall score. At the same time, the task of understanding prepositions significantly correlates with all indicators for visual working memory (content, spatial, and bonus scores).

The children's productivity in the Understanding of Similar Sounding Words task (USSW) was significantly associated with good results in the working memory tasks (both visual and verbal), cognitive flexibility, and the number of uncorrected errors in the Inhibition subtest, which reflects the information-processing speed. Such indicators as substitutions and omissions did not have significant correlations with the EF tasks. However, the mistakes associated with the naming of superfluous words in this task were significantly correlated with cognitive flexibility ($r = 0.134$, $p = 0.025$), and those of duplications were linked to the time of inhibition task performance ($r = 0.165$, $p = 0.007$). In other words, the longer it takes a child to perform the task, the more duplication he has in the phonemic hearing test.

A child's productivity on all three tasks of the Verbal Fluency Test (VFT) (free, action, and animal associations) proved to be significantly associated with verbal working memory and with the time required for the Naming task in the Inhibition method, which reflects the speed of information processing. At the same time, the productivity of free association, unlike the others (animal and action word generation), also correlated with visual working memory. The productivity of free and animal word generation was also significantly associated with cognitive flexibility (DCCS), while action naming was not. However, the productivity of action word generation proved to be significantly inversely associated with the number of corrected mistakes, and was inversely related to the time it took to do the Inhibition task. This result indicates a connection between the ability to generate action words, and the level of inhibition development.

In sum, oral language skills have the greatest number of significant correlations with verbal working memory and cognitive flexibility.

Language performance by children with different EF levels

The cluster analysis (using the K-means method) of the results of the children's performance on the tests diagnosing cognitive regulation singled out three groups of children who differed on their level of EF development (*Table 3*). All differences between the clusters are significant, except for the number of corrected mistakes in the Inhibition task (Kraskell-Wallis criterion, $p \leq 0{,}05$).

It is important to note that children with high levels of EF development significantly differ from those with low levels of EF development on all indicators, except for the number of corrected mistakes in the Naming task of the Inhibition subtest (Mann-Whitney's criterion, $p \leq 0.005$). In addition, children with low EF levels significantly differ in almost all parameters from those with medium EF level (Mann-Whitney's criterion, $p \leq 0.005$). Only on the correct memorization of image details (MD Content) and the number of mistakes in both tasks on the Inhibition subtest are there no differences.

Table 3

Children with different levels of EF development (centers of clusters, cluster analysis using the K-means method)

	Low level of EF	**Medium level of EF**	**High level of EF**
Visual Working Memory (MD), Content	34.96	35.60	43.79
Visual Working Memory (MD), Spatial	15.87	18.08	22.73
Visual Working Memory (MD), Bonus	10.76	12.68	34.08
Visual Working Memory (MD),Total Score	61.60	66.36	100.60
Verbal Working Memory (Sentence Repetition)	16	*19*	*20*
Cognitive Flexibility (DCCS)	17	*19*	19
Naming, uncorrected errors	*0.9*	*0.8*	0.5
Naming, corrected errors	*1.2*	*0.9*	1.3
Naming, time	61.82	43.15	43.25
Inhibition, uncorrected errors	5.2	4.2	2.1
Inhibition, corrected errors	2.4	1.9	2.6
Inhibition, time	85.09	55.59	56.70
Number of children	55	130	73

Children with high and medium EF levels differ less significantly from each other: nor do they differ in the successful performance of the task on cognitive flexibility (DCCS) and verbal working memory (SR), or the time it takes to complete the Inhibition tasks (Naming and Inhibition time). However, at the same time, children with medium EF levels make more mistakes than those with high EF levels. Thus, children with a low level of executive function significantly differ from the other two groups in terms of the development level of all the EF components. Therefore, we compared only the two extreme groups–those with low and high level of EF–on the successful performance of language tasks.

Children with a high level of EF development were more successful in tasks involving understanding of prepositions, understanding of similar sounding words, and verbal fluency (*Table 3*). It is important to note that comparison of children with high and average EF levels did not produce significant differences. Thus, the results of the cluster analysis showed that it is precisely the low level of EF development which correlates with a low level of language development.

Table 4

Differences in the performance of language tasks in children with high and low EF levels (the Mann-Whitney test)

	High level of EF	Low level of EF	Level of significance
ULGC, understanding of prepositions	2.9	2.2	0.010
USSW, productivity	18.9	15.5	0.005
VFT, free associations	18.5	15.8	0.039
VFT, actions	8.6	7.4	0.048
VFT, animals	12.2	10.6	0.039

Discussion

The purpose of this stage of the study was to analyze the interrelationship between EF components and indicators of language development in senior preschool children.

The results of the correlation analysis showed that the level of language development is significantly associated with the scores on the *verbal working memory test*. This relationship is logical and natural since the performance of a working memory task requires the involvement of language functions. Most tasks aimed at language measurement require one to keep words in one's memory, which makes language methods sensitive to the ability to process and retain auditory information in working memory. Thus, the Sentence Repetition subtest may be a good indicator of not only verbal working memory development, but of language development as well. Another explanation of this connection is provided by a theory that the development of EF, and in particular that of working memory, allows children to gradually identify and memorize individual words from the flow of speech interaction, which in turn helps them to increase their vocabulary (Blair et al., 2012; Montgomery, Magimairaj, & Finney, 2010; Rose, Feldman, & Jankowski, 2009; Verhagen & Leseman, 2016; Weiland et al., 2014).

Tasks aimed at measuring the understanding of logical and grammatical constructions, and those on phonological processing abilities, have also proved to be associated with *visual working memory*. Among other reasons, this can be explained by the specifics of this task: the analysis of the large number of pictures the child has to choose from requires a certain level of development of visual working memory and, in particular, the analysis of the image details.

Most of the language indicators are significantly correlated with *cognitive flexibility*, which shows the need to switch from one word, picture, or task to others in the course of their performance. However, the results on the action-naming tasks (based on naming of pictures and verbal fluency test) were unrelated to this EF component. Furthermore, the task of action associations (VFT) was related to the level of development of the inhibitory process, in contrast to other tasks on verbal generation. We can assume that naming actions is more difficult than object naming for 5-6 year old children, and involves some other higher mental functions.

The level of development of *inhibitory processes* has been significantly associated with the successful performance of the task of understanding logical and grammatical constructions (both on understanding active and passive constructions, and understanding prepositions). Perhaps children need to suppress wrong answers in the course of this task. This might also explain the appearance of redundant words and duplications when children perform the tasks on phonemic hearing: when choosing the right pictures, it is difficult for children to inhibit those words on which they reflected during the task, or have met in previous tests. Another explanation for this relationship may that of Zelazo and his colleagues (Zelazo et al., 2003), who believe that children need words in order to imagine carrying out a task before creating the hierarchical structure of the rules needed to solve the problem. It is quite possible that the level of vocabulary, and the ability to compose phrases, affect the emergence of EF. For example, when a child is learning to distinguish between the words "cat" and "frog," while being aware that both names are classified as "animals", he/she has to learn to organize concepts in a hierarchical order (Hall & Waxman, 1993). This understanding of the rules resulting from the use of language may influence the development of the ability to organize the information necessary for cognitive processes, which includes such EF components as cognitive flexibility and inhibition (Zelazo et al., 2003).

Thus, oral language skills are more closely related to the level of development of verbal working memory and cognitive flexibility then to inhibition, which somewhat contradicts the assumption of Blair and Razza (2007) that inhibition is the most significant executive function for language development at preschool age.

The results of cluster analysis showed that children with low levels of EF development are significantly less able to cope with the tasks of understanding prepositional structures, understanding similar sounding words, and verbal fluency, compared to children with high EF levels; whereas children with normal and high levels of EF development display no significant differences in language development. Thus, a low level of EF development has a lot to do with language difficulties.

Conclusion

The results of our study revealed significant associations between all EF components and language skills development in preschool children. These results are consistent with most research in this area (Fuhs & Day, 2011; Pazeto et al., 2014; Fuhs, Nesbitt, & Farran, 2014, Bohlmann, Maier, & Palacios, 2015; Slot & Suchodoletz, 2018), and shows the need to further investigate the mechanisms behind this connection. Furthermore, we found that children with a low level of EF also have a low level of language development.

Limitations

The children's expressive skills were not fully analyzed in our study. For example, a correlation analysis of EF with the ability to construct long phrases and sequences was not included. Furthermore, we did not measure the children's articulation skills. Letter-reading abilities were not included in the study at this stage because they are not a part of the curriculum for 5-6 year old children in Russian kindergarten.

Acknowledgements

This work was supported by a grant from the Russian Foundation For Basic Research No. 17-29-09112.

References

Aaron, P.G., Joshi, R.M., Gooden, R., & Bentum, K.E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading. *Journal of Learning Disabilities*, *41*(1), 67–84. https://doi.org/10.1177/0022219407310838

Akhutina, T.V. (2016). *Methods of neuropsychological examination of children 6–9 years*. Moscow: V. Sekachev.

Akhutina, T.V. & Pylaeva, N.M. (2015). *Overcoming learning disabilities: a vygotskian-lurian neuropsychological approach* (2nd ed.). Moscow: Academia.

Almazova, O.V., Bukhalenkova D.A., & Veraksa, A.N. (2016). Proizvol'nost' v doshkol'nom vozraste: sravnitel'nyi analiz razlichnykh podkhodov i diagnosticheskogo instrumentariia [The voluntariness in the preschool age: a comparative analysis of various approaches and diagnostic tools]. *Natsional'nyi psikhologicheskii zhurnal* [National Psychological Journal], *4*(24), 14–22. https://doi.org/10.11621/npj.2016.0402

Blair, C. & Razza, R.P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Blair, C., Protzko, J., & Ursache, A. (2012). Self-regulation and early literacy. In S.B. Neuman, & D.K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 20–35). New York: Guilford.

Bohlmann, N.L., Maier, M.F., & Palacios, N. (2015). Bidirectionality in self-regulation and expressive vocabulary: Comparisons between monolingual and dual language learners in preschool. *Child Development*, 86, 1094–1111. https://doi.org/10.1111/cdev.12375

Cantin, R.H., Gnaedinger, E.K., Gallaway, K.C., Hesson-McInnis, M.S., & Hund, A.M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology*, 146, 66–78. https://doi.org/10.1016/j.jecp.2016.01.014

Rojas-Barahona, C.A., Förster, C.E., Moreno-Ríos, S., & McClelland, M.M. (2015) Improvement of working memory in preschoolers and its impact on early literacy skills: A study in deprived communities of rural and urban areas. *Early Education and Development*, 26, 5–6, 871–892. https://doi.org/10.1080/10409289.2015.1036346

Cutting, L.E., Materek, A., Cole, C.A.S., Levine, T.M., & Mahone, E.M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia*, *59*(1), 34–54. https://doi.org/10.1007/s11881-009-0022-0

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F.I.M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). New York, NY: Oxford University. https://doi.org/10.1093/acprof:oso/9780195169539.003.0006

Diamond, A. & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science*, 333, 959–964. https://doi.org/10.1126/science.1204529

Duff, F.J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry, 56*(8), 848–856. https://doi.org/10.1111/jcpp.12378

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*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Elkonin, D.B. (2006). *Child psychology: studies textbook for students of higher educational institutions*. Moscow: Academy.

Fuhs, M.W. & Day, J.D. (2011). Verbal ability and executive functioning development of preschoolers at Head Start. *Developmental Psychology*, 47, 404–416. https://doi.org/10.1037/a0021065

Fuhs, M.W., Nesbitt, K.T., Farran, D.C., & Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, 50, 1698–1709. https://doi.org/10.1037/a0036633

Hall, D.G. & Waxman, S.R. (1993). Assumptions about word meaning: Individuation and basic-level kinds. *Child Development* 64, 155–1570. https://doi.org/10.2307/1131552

Henry, L.A., Messer, D.J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry*, 53, 37–45. https://doi.org/10.1111/j.1469-7610.2011.02430.x

Korkman, M., Kirk, U., & Kemp, S.L. (2007). *NEPSY II. Administrative manual*. San Antonio, TX: Psychological Corporation.

Kuhn, L.J., Willoughby, M.T., Wilbourn, M.P., Vernon-Feagans, L., Blair, C.B., & The Family Life Project Key Investigators (2014). Early communicative gestures prospectively predict language development and executive function in early childhood. *Child Development*, 85, 1898–1914. https://doi.org/10.1111/cdev.12249

Lehto, J., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *The British Journal of Developmental Psychology*, 21, 59–80. https://doi.org/10.1348/026151003321164627

Luria, A.R. (1976). *Cognitive Development: Its Cultural and Social Foundations*. Harvard University Press.

Matthews, J.M., Ponitz, C.C., & Morrison, F.J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, *101*(3), 689–704. https://doi.org/10.1037/a0014240

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Montgomery, J.W., Magimairaj, B.M., & Finney, N.C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology*, *19*(1), 78–94. https://doi.org/10.1044/1058-0360(2009/09-0028)

Pazeto, T.C.B., Seabra, A.G., & Dias, N.M. (2014). Executive functions, oral language and writing in preschool children: *Development and correlations*. Paidéia (Ribeirão Preto), *24*(58), 213–221. https://doi.org/10.1590/1982-43272458201409

Petersen, I.T., Bates, J.E., D'Onofrio, B.M., Coyne, C.A., Lansford, J.E., Dogde, K.A., ... & Van Hulle, C.A. (2013). Language ability predicts the development of behavior problems in children. *Journal of Abnormal Psychology*, *122*(2), 542–557. https://doi.org/10.1037/a0031963

Rojas-Barahona, C.A., Förster, C.E., Moreno-Ríos, S., & McClelland, M.M. (2015). Improvement of working memory in preschoolers and its impact on early literacy skills: A study in deprived communities of rural and urban areas. *Early Education and Development*, *26*(5–6), 871–892. https://doi.org/10.1080/10409289.2015.1036346

Rose, S.A., Feldman, J.F., & Jankowski, J.J. (2009). A cognitive approach to the development of early language. *Child Development*, *80*(1), 134–150. https://doi.org/10.1111/j.1467-8624.2008.01250.x

Reshetova, Z.A. (2017). Psikhika i deiatel'nost'. Psikhicheskii mekhanizm usvoeniia [Mind and activity. Psychic mechanism of learning]. *Natsional'nyi psikhologicheskii zhurnal* [National Psychological Journal], *3*(27), 40–55. https://doi.org/10.11621/npj.2017.0306

Rzhanova, I.E., Alekseeva, O.S., Fominykh, A.Ya., & Parshikova, O.V. (2018). Indeks rabochei pamiati kak odin iz osnovnykh pokazatelei testa Vekslera dlia doshkol'nikov [The working memory index as one of the main scales of the Wechsler preschool and primary scale of intelligence — fourth edition]. *Psikhologicheskie Issledovaniya* [Psychological research], *11*(57), 8. Retrieved from http://psystudy.ru/index.php/num/2018v11n57/1529-rzhanova57.html

Slot, P.L. & Suchodoletz A. (2018). Bidirectionality in preschool children's executive functions and language skills: Is one developing skill the better predictor of the other? *Early Childhood Research Quarterly*, 42, 205–214. https://doi.org/10.1016/j.ecresq.2017.10.005

Verhagen, J. & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, 141, 65–82. http://dx.doi.org/10.1016/j.jecp.2015.06.015

Visu-Petra, L., Cheie, L., Benga, O., & Miclea, M. (2012). The structure of executive functions in preschoolers: An investigation using the NEPSY battery. *Procedia — Social and Behavioral Sciences*, 33, 627-631. https://doi.org/10.1016/j.sbspro.2012.01.197

Vygotsky, L.S. (1962). *Thought and language*. Oxford, England: Wiley. https://doi.org/10.1037/11193-000

Vygotsky, L.S. (1984). Child psychology. Vol. 4. In *Collected works in 6 volumes.* Moscow: Pedagogy.

Weiland, C., Barata, M.C., & Yoshikawa, H. (2014). The co-occurring development of executive function skills and receptive vocabulary in preschool-aged children: A look at the direction of the developmental pathways. *Infant and Child Development*, *23*, 4–21.

Willoughby, M.T., Kupersmidt, J.B, & Voegler-Lee, M.E. (2012). Is preschool executive function causally related to academic achievement? Child Neuropsychology: *A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *18*(1), 79–91.

Yeniad, N., Malda, M., Mesman, J., van IJzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences*, *23*, 1–9. https://doi.org/10.1016/j.lindif.2012.10.004

Zelazo, P.D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. National Protocols, *1*, 297–301. https://doi.org/10.1038/nprot.2006.46

Zelazo, P.D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monograph of the Society for Research in Child Development*, *68*(3). https://doi.org/10.1111/j.0037-976X.2003.00261.x

Original manuscript received August 29, 2018
Revised manuscript accepted October 23, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

The Role of Emotional Schemas in Anxiety and Depression among Russian Medical Students

Natalya A. Sirota, Denis V. Moskovchenko*,
Vladimir M. Yaltonsky, Alexandra V. Yaltonskaya

A.I. Yevdokimov Moscow State University of Medicine and Dentistry, Moscow, Russia

Corresponding author. E-mail: xedin-90@mail.ru

Background. Academic success in a higher education institution requires the ability to process large amounts of information in a relatively short period of time, including having proficiency at a high level of basic knowledge, and an ability to cope with stress. Continual study overload, a competitive environment, and ethical dilemmas (e.g. "How should I deal with human suffering?", "How should I convey the diagnosis?", "How should I tell someone that palliative treatment is the only option?", "What if I make a mistake?") can all result in anxiety and depression. Research has shown that students who show signs of anxiety and depression may have maladaptive cognitive strategies for processing their emotional experiences. In the medical community, the rules concerning one's own emotions are, on one hand, determined by specific ethical standards (e.g., the idea that physicians should not show their emotions), and on the other, by the stressful situation itself, which requires taking responsibility for another person's life. The additional stress point is the need for constant study, which requires a pro-active attitude and learning more and more skills. A significant number of physicians tend to ignore their own emotional experiences, or suppress them. The present study deals with indications of anxiety and depression on the basis of such emotional schemas, which we suggest play the key role in the development of emotional maladaptation in medical students.

Objective. In this study we observe signs of *anxiety and depression* in medical students and their dependence upon the intensity of dysfunctional emotional schemas.

Design. The number of participants was 400, comprised of students from general medicine (n = 300) and dentistry (n = 100) at the Moscow State University of Medicine and Dentistry.

Methods. We took from the Symptom Check List-90-Revised (Russian version, N.V. Tarabrina N.V.) the subscales related to affective and anxiety disorders: anxiety, depression, interpersonal sensitivity, obsessive-compulsiveness, somatization, and phobic anxiety. We also used 28 items from the Leahy Emotional Schema Scale II (the Russian version, adapted by the authors and Y.A. Kochetkov).

Results. The medical students fell into two groups: those with low and those with high intensity of the dysfunctional schemas. The groups were distinguished by which of

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

doi: 10.11621/pir.2018.0409
http://psychologyinrussia.com

Leahy's basic emotional regulation strategies, either normalizing or pathologizing, they used. The pathologizing students followed strict, maladaptive rules concerning their emotional experiences. Students with intense dysfunctional schemas also demonstrated signs of anxiety, depression, obsessive-compulsiveness, and somatization. The students who saw their emotions as normal demonstrated lower levels of dysfunctional emotional schemas. As stated in Leahy's emotional schemas theory, such students tend to see their emotions as a normal, important, and meaningful part of their daily lives. Analysis has shown that these types of students exhibit lower levels of anxiety, depression, obsessive-compulsiveness, somatization, and interpersonal sensitivity. Regression analysis demonstrated that emotional schemas are significantly related to emotional maladaptation in students. The analysis also allowed us to determine the association of different emotional schemas with the development of anxiety, obsessive-compulsiveness, somatization, and interpersonal sensitivity. Adoption of emotional schemas correlated with the symptoms of depressive and anxiety disorders.

***Keywords**:* emotional schemas, anxiety, depression, medical students, emotional self-regulation.

Introduction

During the last few decades, affective disorders have become more common in Russian young adults (Garanyan & Kholmogorova, 1999; Voytsekh, 2006; Dozortseva, 2006; Shifner, 2011; Garanyan & Shchukin, 2014; Matyushkina, 2016). Some studies report a high level of affective disorders in students, with one focus group reporting a rate as high as 20% to 30% (Kholmogorova, 2006). Among medical students, anxiety and depression are the most typical reactions to overload and stress; these reactions negatively impact their quality of life, as well as their social and emotional adaptation (Hope & Henderson, 2014)

Academic success at an institution of higher education requires an ability to process large amounts of information in a relatively short timeframe, a high level of basic knowledge, and an ability to cope with stress (Chandavarkar, 2007). A perpetual study overload, competitive environment, and the treatment of ethical dilemmas (e.g. "How should I deal with human suffering?", "How should I convey the diagnosis?", "How should I tell someone that palliative treatment is the only option?", "What if I make a mistake?") can result in anxiety and depression (Sirota et al., 2016). However, some authors state that personality traits have a more significant impact on the development of affective disorders in students (Brown & Gunderman, 2006)

Medical students face both emotionally intense negative (e.g. suffering, death, unsafety), and emotionally-intense positive (e.g., strong role-models of compassion, empathy, patient-oriented behavior) events (Feudtner et al., 1994; Monroux-eet al., 2014; Rees et al., 2015; Branch et al., 2001).

Medical students report symptoms of anxiety, depression, burn-out, and in some cases even demonstrate signs of post-traumatic stress disorder (Tschernig et al., 2000). Previous studies have also shown that medical students experience decreasing empathy and increasing cynicism; for patients this often leads to less-than-ideal care, and for the students themselves, dropping out of medical school (Thomas et al, 2007). It is crucial to understand the strategies used by medical stu-

dents to regulate their emotions, and the broader impact of those strategies. Yet a relatively small number of studies deal with the students´ emotional regulation (Monrouxe et al., 2015). It has been shown that adequate emotional regulation contributes positively to the productivity and wellness of medical professionals (e.g., they pay more attention to the patient and spend more time on them) as well as leading to decreased burn-out and increases in positive emotions (Zammuner et al., 2003).

A growing number of authors describe the importance of emotional reflection for the development and maintenance of psychological disorders (e.g., Naderi et al., 2015). Some studies show that, in coping with patients' worries, some physicians tend to avoid their own emotions and distance themselves from these emotions (Shapiro, 2010). It has also been shown that physicians who tend to repress their negative emotions about their patients are more likely to develop signs of anxiety and depression (Sung et al., 2009). In other words, the authors suggest that emotional repression has a negative impact and leads to the development of affective disorders. Specific strategies and rules concerning managing one's emotions are manifested by specific emotional schemas (Leahy, 2002). Among the new approaches to the problem of emotional regulation is Leahy's theory of emotional schemas.

Studying dysfunctional emotional schemas can be considered a new approach to understanding the role of personal factors in emotional adaptation/maladaptation. Leahy's theory of emotional schemas is based on the emotional-focus model (Greenberg, 2002), as well as on meta-cognitive theory (Wells, 1995). Emotional schemas are the subjective beliefs, ideas, and structures which people use as a response to activation of any strongly experienced emotion. A cluster of interpretations and expectations about one's own and other people's emotions is defined as an emotional schema: it involves what a person thinks about his/her and other people's emotions, and which behavioral and interpersonal strategies they use in response. (Leahy & Tirch, 2011).

The model of emotional schemas includes two emotional coping strategies: normalizing and pathologizing. The process of emotional normalization starts from the idea that emotions can be accepted and expressed adequately, and that they are temporary and reflect contextual value (emotions receive their value in context of the situation). By contrast, in the process of pathologizing coping, negative emotions are considered unique, long-term, and dangerous for the person who experiences them; the person therefore feels the need to repress or control them (Leahy & Kaplan, 2004).

It has been suggested that the emotional schemas one chooses are based on emotional experiences and interactions with significant others. Studies have shown that maladaptive emotional schemas are strongly connected to depression, anxiety, post-traumatic stress disorder, disagreements over family-functioning, and personality disorders (Leahy, 2003). There is a lack of research on emotional schemas and how they are related to students' emotional vulnerability.

Despite the importance and necessity of finding predictor variables concerning emotional maladaptation in students, there are not enough studies dealing with the cognitive factors engaged in the development of negative experiences. The subject of this study was to assess the dysfunctional emotional schemas which lead to anxiety and depressive experiences in medical students.

Methods

The number of participants was 400, comprised of students from general medicine (n = 300) and dentistry (n = 100) attending the Moscow State University of Medicine and Dentistry. The average age was 20 (14 ± 7,25). We had 249 female and 151 male participants. There were 60 participants from the first year of study (n = 60), 50 from the second year (n = 50), 140 from the third (n = 140), 50 from the fourth (n = 50), and 100 from the fifth (n = 100). All the participants had filled out several diagnostic questionnaires.

In order to study the emotional schemas' intensity and structure, we adapted (2016) the "Leahy Emotional Schema Scale II" (2012). The questionnaire included 28 items, two per each schema, and allowed us to evaluate the participants' dysfunctional cognitive schemas about their emotions. The emotional schema scale is a one-factor scale. It evaluates the presence of dysfunctional styles of interpreting emotions. Cronbach's alfa reached the point of 0.80, which indicated a sufficient level of internal consistency. The Kaiser-Guttmann criterion equaled 0.83. Correlation between the subscales equaled 0.716.

Although this is a one-factor structure scale, it assesses the following 14 emotional schemas:

1) **Incomprehensibility**, e.g. "My emotions do not make any sense!"
2) **Loss of Control**: I can surely lose control of my emotions.
3) **Duration**: I will feel this emotion for a very long time.
4) **Devalued**: I should not feel any of this.
5) **Simplistic View of Emotion**: It is hard to understand complex emotions.
6) **Blame**: It is other people who determine my emotional state.
7) **Low Consensus**: The others are just not able to feel the same as I do;, they feel something else.
8) **Low Expression**: I should not openly express my emotions.
9) **Simplistic View of Emotion**: Complex feelings are hard to accept.
10) **Guilt**: My emotions are shameful, wrong, and inappropriate.
11) **Rumination**: I am constantly thinking about my emotions and where they are coming from.
12) **Numbness**: I do not have any feelings.
13) **Invalidation**: Others will not accept my emotions; they will ignore and devalue my feelings.
14) **Overly Rational**: It is better to stay rational.

In order to study the broad extent of anxiety and depression, we used specific subscales from the "Symptom Check List-90-Revised" (1994) by L. Derogatis, adapted by N.V. Tarabrina (2007). We used the following subscales: anxiety, depression, interpersonal sensitivity, obsessive-compulsiveness, somatization, and phobic anxiety. To evaluate the validity of the Russian adaptation, we used a-coefficient (the variation of Ruder-Richardson 20) as a confidence criterion. The final coefficients were sufficiently distributed between 0.77 (lower coefficient for the psychoticism subscale) and 0.90 (higher coefficient for the depression subscale). Test-retest coefficients were obtained from the data of 94 psychiatric patients' data. These pa-

tients were examined upon intake and then a week after, before the first therapeutic session. Most of these coefficients were between 0.80 and 0.90.

We performed statistical analysis in "Statistica 10.0, SPSS 21.0".

Results

Cluster profiles of emotional schemas in students

In the first part of the study, in order to define the specific structure of the students' emotional schemas, we performed a two-stage cluster analysis that allowed us to divide the group on the basis of the emotional schema questionnaire. We used the total score from the emotional schemas questionnaire and then divided it into two values according to the Akaike Information Criteria.

As a result, two cluster modules were defined. We compared the groups on individual schemas. The mean silhouette measure of cohesion and separation was 0.3. We used k-means to identify which students went into which final cluster. The first cluster included students with a lower intensity level of dysfunctional emotional schemas (n = 256); the second cluster included students with a higher intensity level of dysfunctional emotional schemas (n = 144). The differences between the clusters can be seen in *Table 1*.

Table 1

Cluster profiles of emotional schemas in medical students

Emotional schemas in students	Low level of dysfunctional schemas' manifestation Cluster 1 (n = 256)		High level of dysfunctional schemas' manifestation Cluster 2 (n = 144)		Distribution difference according to Mann–Whitney U test
	M	SD	M	SD	
Invalidation	2.53	0.78	3.33	1.10	0.001
Incomprehensibility	2.64	0.88	3.68	1.01	0.001
Guilt	1.58	0.67	3.09	1.03	0.001
Simplistic View of Emotion	4.30	1.13	4.35	1.12	0.888
Devalued	2.79	1.12	2.67	1.24	0.235
Loss of Control	2.30	1.05	4.10	1.10	0.001
Numbness	2.25	0.96	3.17	1.07	0.001
Overly Rational	3.07	1.15	4.13	1.25	0.001
Duration	2.61	0.96	3.24	1.07	0.001
Low consensus	2.44	1.11	2.87	0.96	0.001
Non-Acceptance of Feelings	2.13	0.77	3.50	0.94	0.001
Rumination	2.80	1.04	4.36	1.02	0.001
Low expression	2.59	0.93	3.19	1.15	0.001
Blame	2.79	1.03	3.36	1.17	0.001

M = mean value; SD = standard deviation.

The results showed that the students from the two clusters demonstrated significant differences in almost all dysfunctional schemas. The only areas where we did not see any difference were in "Simplistic view of emotion" and "Devalued".

Results on anxiety and depression in medical students

The study of the students' emotional maladaptation showed that in the group where dysfunctional emotional schemas were most evident, the signs of anxiety and depression were significantly stronger. In addition, we found that in students from the second cluster, the scores on the subscales of "Somatization," "Obsessive-compulsiveness," and "Interpersonal sensitivity" were significantly higher.

One can see the results in *Table 2*.

Table 2

Emotional maladaptation (by SCL-90R)

Emotional maladaptation (SCL-90R)	Low level of dysfunctional schemas' manifestation (Cluster1; n = 256)		High level of dysfunctional schemas' manifestation (Cluster 2; n = 144)		Distribution difference according to Mann–Whitney U test
	M	SD	M	SD	
Somatization	0.48	0.22	1.22	0.66	0.04
Obsessive-compulsiveness	0.88	0.47	1.56	0.43	0.001
Interpersonal sensitivity	0.75	0.56	1.44	0.67	0.001
Anxiety	0.71	0.56	1.44	0.90	0.05
Phobic anxiety	0.34	0.22	0.53	0.22	0.442
Depression	0.68	0.44	1.08	0.33	0.01

M = mean value; SD = standard deviation.

In order to establish significant predictors associated with the development of anxiety, depression, interpersonal sensitivity, phobic anxiety, and obsessive-compulsiveness, we used multiple linear regression (hierarchically entering the variables into the regression model). All subscales of the emotional schemas inventory were included in the regression analyses, but only the subscales that had statistical significance for multiple regression models were presented in the study results.

Anxiety and emotional schemas

The first run of regression analysis was aimed at defining the emotional schemas significantly associated with anxiety. The results can be seen in *Table 3*.

Multiple regression analysis has shown that a general level of anxiety is associated with the following emotional schemas: "Simplistic view of emotion" (e.g., I prefer to clearly understand what I feel about the other person), "Non-acceptance

of feelings" (e.g., I should avoid some feelings and emotions), and "Invalidation" (e.g., Other people do not care) ($F = 17.80$, $p = 0.001$). This model explains the 43% dispersion variance ($R^2 = 0.437$).

Table 3

Emotional schemas' impact on general level of anxiety

Dependent variable: Anxiety	**$R^2 = 0.437$**	**B**	**t**	**p**
Simplistic view of emotion		0.395	5.56	0.001
Non-acceptance of feelings		0.174	2.54	0.01
Invalidation		0.208	3.20	0.001

Interpersonal sensitivity and emotional schemas

The second run of regression analysis was aimed at defining the emotional schemas most common for interpersonal sensitivity. The results can be seen in *Table 4*.

Table 4

Emotional schemas' impact on interpersonal sensitivity

Dependent variable: Interpersonal sensitivity	**$R^2 = 0.419$**	**B**	**t**	**p**
Rumination		0.337	4.71	0.001
Incomprehensibility		0.267	4.04	0.001
Invalidation		0.218	3.38	0.001
Overly rational		–0.213	–3.27	0.001
Blame		0.145	2.16	0.03

Multiple regression analysis has shown that general level of interpersonal sensitivity is associated with the following emotional schemas: "Rumination" (e.g., When I feel low, I often sit alone and think about how miserable I am); "Incomprehensibility" (e.g., There are things about myself that I just don't understand); "Invalidation" (e.g., The others do not care); and "Blame" (e.g., If other people changed, I would feel a lot better). This contrasts with "Overly rational", which had a lower score ($F = 19.39$, $p = 0.001$). This model explains the 41% dispersion variance ($R^2 = 0.419$).

Depression and emotional schemas

The third run of regression analysis was aimed at defining the emotional schemas most common for depression. The results can be seen in *Table 5*.

Multiple regression analysis has shown that general level of depression is conjointly influenced by the following emotional schemas: "Rumination", "Incompre-

hensibility", "Invalidation," and "Overly rational" (It is important for me to be reasonable and practical rather than sensitive and open to my feelings) ($F = 17.80$, $p = 0.001$). This model explains the 43% dispersion variance ($R^2 = 0.437$).

Table 5

Emotional schemas' impact on interpersonal sensitivity

Dependent variable: Depression	**$R^2 = 0.437$**	**B**	**t**	**P**
Rumination		0.395	5.56	0.001
Incomprehensibility		0.174	2.54	0.01
Invalidation		0.208	3.20	0.001
Overly rational		–0.156	–2.42	0.01

Obsessive-compulsiveness and emotional schemas

The fourth run of regression analysis was aimed at defining the emotional schemas most common for obsessive-compulsiveness. The results can be seen in *Table 6*.

Table 6

Emotional schemas' impact on obsessive-compulsiveness

Dependent variable: Obsessive-compulsiveness	**$R^2 = 0.333$**	**B**	**T**	**P**
Rumination		0.236	2.95	0.003
Incomprehensibility		0.232	3.18	0.001
Loss of control		0.211	2.64	0.009

Multiple regression analysis has shown that general level of obsessive-compulsiveness is conjointly influenced by the following emotional schemas: "Rumination," "Incomprehensibility," and "Loss of control" (If I let myself have some of these feelings, I fear I will lose control) ($F = 11{,}43$, $p = 0{,}001$). This model explains the 33% dispersion variance ($R^2 = 0.333$).

Somatization and emotional schemas

The fifth run of regression analysis was aimed at defining the emotional schemas most common for somatization. The results can be seen in *Table 7*.

Multiple regression analysis has shown that general level of somatization is conjointly influenced by the following emotional schemas: "Rumination," "Incomprehensibility," "Loss of control," and "Guilt" (Some feelings are wrong to have) ($F = 12.47$, $p = 0.001$). This model explains the 27% dispersion variance ($R^2 = 0.276$).

Table 7

Emotional schemas' impact on somatization

Dependent variable: Somatization	$R^2 = 0.276$	β	T	p
Rumination		0.269449	3.15	0.001
Incomprehensibility		0.282413	3.68	0.000
Loss of control		0.184069	2.32	0.021
Guilt		0.177456	2.15	0.032

Phobic anxiety and emotional schemas

The sixth run of regression analysis was aimed at defining the emotional schemas most common for phobic anxiety.

Table 8

Emotional schemas' impact on phobic anxiety

Dependent variable: Phobic anxiety	$R^2 = 0.266$	B	t	p
Incomprehensibility		0.282413	3.68	0.000
Loss of control		0.184069	2.32	0.021
Blame		0.177456	2.15	0.032

Multiple regression analysis has shown that general phobic anxiety is conjointly influenced by the following emotional schemas: "Rumination", "Incomprehensibility", "Loss of control", and "Blame" ($F = 7.26$, $p = 0.001$). This model explains the 26% dispersion variance ($R^2 = 0.266$).

Discussion

The findings of the present study are in line with the existing research showing that medical students are a vulnerable group (Dyrbye & Thomas, 2006).

Students with intense dysfunctional emotional schemas face anxious and depressive experiences more often than others (Kamali & Gharraee, 2013).

Cluster analysis allowed us to divide the respondents into two groups. The first group of students demonstrated a lower level of dysfunctional emotional schemas. According to the theory of emotional schemas, such students tend to see their emotions as normal, important, and meaningful parts of everyday life (Khawaja & Chapman, 2007).

The students from the second cluster had, on average, stricter but maladaptive rules about their emotional experiences. In other words, the students from the second cluster tended to think that others will devalue and ignore their emotions.

Such students tended to think that their emotions mean nothing, and so they do not think about their meaning. They more often had a belief that expressing emotions is wrong and shameful.

In other words, students from the second cluster tended to suppress their emotions. Some studies show that suppression is the most common strategy for professions that require high levels of control, such as serving on a police force, court, or in the military (Tull et al., 2007). Moreover, some studies state that emotional disregard is a necessary adaptive strategy for students because emotions consume one's cognitive resources, leading people to think more about their emotional object than about the actual task (Ellis & Ashbrook, 1988). Therefore, controlling one's emotions are associated with cognitive skills of task management and negatively affect productivity (Meinhardt & Pekrun, 2003). In this case, emotional suppression leads to short-term regulation, so that a student can gain control and manage with the task.

However, despite the fact that emotional suppression can be a pragmatic short-term survival strategy, it has some potentially serious long-term consequences for the doctors, patients, and the system at large. Numerous studies mention increasing depersonalization and burn-out, as well as a deficit of empathy among medical students, and state that distancing one's self from one's own inner life leads to distancing one's self from a patient (Neumann et al., 2011).

Moreover, the participants from the second cluster saw their emotions as less controllable; they more often distance themselves from their emotions, rationalized them, and tended to think that an emotion will last for a very long time. Statistically, such students tended to use maladaptive regulation strategies such as "Non-Acceptance" and "Blame" more often than others. In fact, a number of studies show that medical students tend to avoid places, events, or even patients who trigger intense emotional experiences (Fritz & Sonnentag, 2006).

Finally, the participants from the second cluster tended to think that others rarely experience the emotions they do; they focus and reflect on their own specific emotional experiences. Students in this cluster also had a tendency for low emotional expressiveness: as the cluster analysis has shown, they did not want to experience and express emotions.

We should also mention that students with dysfunctional emotional schemas had a tendency for compulsiveness and intrusive ruminating thoughts. According to cognitive avoidance theory, people who repress their emotional experiences often end up expressing them through negative output (Borkovec, 2004).

The results of the present study correlate with previous studies, which have shown that some emotionally vulnerable students tend to experience constant tension, worry, depression, and difficulties in problem solving; they also tend to catastrophize problems (Barikana, 2007; Rizvi, 2011).

Regression analysis demonstrates that certain emotional schemas are significantly associated with emotional maladaptation in students. The regressions demonstrated associations between different emotional schemas and anxiety, obsessive-compulsiveness, somatization and interpersonal sensitivity.

First, symptoms of anxiety correlate with having problems accepting difficult, ambivalent emotions. Moreover, the higher the general level of emotional sensitiv-

ity, the lower score the "Overly rational" schema has. We suggest that a decreased ability to understand emotions is a form of cognitive vulnerability. The regression model of emotional schemas correlating with depression is similar to the model of schemas correlating with interpersonal sensitivity, which may suggest general similarity of these psychopathological clusters. The overall character of these clusters also correlates with the development of obsessive-compulsiveness.

Second, both obsessive-compulsiveness and somatization may be partly explained by compulsive focusing on certain emotional states (cf. fixation on interpersonal sensitivity), and the search for their causes (cf. externalization in interpersonal sensitivity), as well as a lack of comprehensibility, uncertainty in one's ability to control one's emotions, and a belief that experiencing one's own emotions is shameful, wrong, or awkward (cf. both depression and interpersonal sensitivity).

Third, the belief that emotional experiences are useless may correlate with phobic behavior.

Conclusions

This problem demands further research, including an examination of the link between emotional schemas and psychopathology. Thus, we should analyze the circumstances for the development of emotional schemas both in medical and non-medical students, and define whether there is a link between certain emotional schemas and maladaptation.

Limitations

The present study should not be generalized to all students. The results are correlational, not causal. As the data were not longitudinal, we also did not trace the dynamics of emotional schemas from one educational year to another.

References

Barikana, A. (2007). Stress in Medical Students. *Journal of Medical Education, 11*(1), 41–44.

Borkovec, T.D., Alcaine, O.M., & Behar, E. (2004). Avoidance theory of worry and generalized anxiety disorder. *Generalized anxiety disorder: Advances in research and practice* (pp. 77–108). New York: Guilford Press.

Branch, W.T.J., Kern, D., Haidet, P., Weissmann, P., Gracey, C.F., & Mitchell, G., et al. (2001) Teaching the human dimensions of care in clinical settings, *Journal of the American Medical Association*. *286*(9), 1067–1074. https://doi.org/10.1001/jama.286.9.1067

Brown, S. & Gunderman, R.B. (2006). Viewpoint: enhancing the professional fulfillment of physicians. *Academic Medicine*, *81*(6), 577–582. https://doi.org/10.1097/01.ACM.0000225224.27776.0d

Bucci, W. (1997). *Psychoanalysis and Cognitive Science: A multiple code theory.* New York: Guilford Press.

Chandavarkar, U., Azzam, A., & Mathews, C. (2007). Anxiety symptoms and perceived performance in medical students. *Depression and Anxiety*, 24, 103–111. https://doi.org/10.1002/da.20185

Dozortseva, E.G. (2006). Psikhologicheskaya travma u podrostkov s problemami v povedenii [Psychological trauma in teenagers with behavioral problems]. *Diagnostika i korrektsiya* [Diagnosis and Correction]. Moscow: Genesis.

Ellis, H.C. & Ashbrook, P.W. (1988). Resource allocation model of the effects of depressed mood states on memory, *Affect, cognition and social behaviour*. Toronto: Hogrefe, 1988, 25–43.

Feudtner, C., Christakis, D.A., & Christakis, N.A. (1994). Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Academic Medicine*, *69*(8), 670–679. https://doi.org/10.1097/00001888-199408000-00017

Fritz, C. & Sonnentag, S. (2006). Recovery, well-being, and performance-related outcomes: the role of workload and vacation experiences. *Journal of Applied Psychology*, *91*(4), 936–945. https://doi.org/10.1037/0021-9010.91.4.936

Garanyan, N.G. & Kholmogorova, A.B. (1999). Emotsional'nyye rasstroystva i sovremennaya kul'tura [Emotional disorders and modern culture]. *Moskovskiy psikhoterapevticheskiy zhurnal* [Moscow psychotherapeutic journal], 2, 61–90.

Garanyan, N.G. & Shchukin, D.A. (2014) Chastye sotsial'nye sravneniya kak faktor emotsional'noi dezadaptatsii studentov [Frequent social comparison and emotional maladjustment among students]. *Konsul'tativnaya psikhologiya i psikhoterapiya [Counseling Psychology and Psychotherapy]*, 4, 182–206.

Greenberg, L.S. (2002). *Emotion-focused therapy: Coaching clients to work through their feelings.* Washington, DC, US: APA. https://doi.org/10.1037/14692-000.

Hope, V. & Henderson, M. (2014). Medical student depression, anxiety and distress outside North America: a systematic review. *Journal of Medical Education*, *48*(10), 963–979. https://doi.org/10.1111/medu.12512

Kamali, S., Gharraee, B., & Birashk, B. (2013). The role of emotional schema in prediction of pathological worry in Iranian students. *Procedia - Social and Behavioral Sciences*, *84*(9), 994–998. https://doi.org/10.1016/j.sbspro.2013.06.687

Kholmogorova, A.B. (2006). *Teoreticheskiye i empiricheskiye osnovaniya integrativnoy psikhoterapii rasstroystv affektivnogo spectra* [Theoretical and empirical foundations of integrative psychotherapy disorders of affective spectrum] (Doctoral dissertation). Moscow.

Khawaja, N.G. & Chapman, D. (2007). Cognitive predictors of worry in a non-clinical population. *Clinical Psychologist*, *11*(1), 24–32. https://doi.org/10.1080/13284200601151778

Leahy, R.L. (2002). A Model of Emotional Schemas. *Cognitive and Behavioral Practice*, 9, 177–190. https://doi.org/10.1016/S1077-7229(02)80048-7

Leahy, R.L. (2003). *Overcoming Resistance in Cognitive Therapy.* New York: Guilford Press.

Leahy, R.L. & Kaplan, D. (2004). Emotional schemas and relationship adjustment. *Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy*. New Orleans, LA.

Leahy, R.L., Tirch, D., & Napolitano, L.A. (2011). *Emotion Regulation in Psychotherapy: A Practitioner's Guide*. New York: Guilford Press.

Matyushkina, E.Ya. (2016). Uchebnyi stress u studentov pri raznykh formakh obucheniya [Academic stress of students with different forms of learning]. *Konsul'tativnaya psikhologiya i psikhoterapiya* [Counseling Psychology and Psychotherapy], *24*(2), 47–63. https://doi.org/10.17759/cpp.2016240204

Meinhardt, J. & Pekrun, R. (2003). Attentional resource allocation to emotional events: an ERP study. *Cognition and Emotion*, *17*(3), 477–500. https://doi.org/10.1080/02699930244000039

Monrouxe, L.V., Rees, C.E., Endacott, R., & Ternan, E. (2014). "Even now it makes me angry": health care students' professionalism dilemma narratives. *Medical Education*, *48*(5), 502–517. https://doi.org/10.1111/medu.12377

Naderi, Y., Moradi, A., Hasani, J., & Noohi, S. (2015). Effectiveness of emotional schema therapy on cognitive emotion regulation strategies of combat-related post-traumatic stress disorder veterans. *Iranian Journal of War and Public Health, 7*(3), 147–155.

Neumann, M., Edelhäuser, F., Tauschel, D., Fischer, M. R., Wirtz, M., & Woopen, C. (2011). Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Academic Medicine, 86*(8), 996–1009. https://doi.org/10.1097/ACM.0b013e318221e615

Rees, C.E., Monrouxe, L.V., & McDonald, L.A. (2015). "My mentor kicked a dying woman's bed … analysing UK nursing students' "most memorable" professionalism dilemmas. *Journal of Advanced Nursing*, *71*(1), 169–180. https://doi.org/10.1111/jan.12457

Rizvi, A. (2010). Pre-examination stress in second year medical students in a government college. *Journal of Ayub Medical College Abbottabad*, 22, 152–155.

Shapiro, S.L., Shapiro, D.E., & Schwartz, G.E. (2000). Stress management in medical education: a review of the literature. *Academic Medicine*, *75*(7), 748–759. https://doi.org/10.1097/00001888-200007000-00023

Shifner, N.A. (2011) Rasstroistva adaptatsii u studentov (ikh klinika i dinamika). Avtoref. diss. kand. med. nauk [Adjustment disorders in students (clinical aspects and dynamics)] (Ph. D. Thesis). Moscow.

Sirota, N.A., Moskovchenko, D.V., Yaltonsky, V.M., Yaltonskaya, A.V., & Chubukova Y.S. (2016). Profili trevozhnosti i gotovnost' k depressivnym sostoyaniyam studentok, obuchayushchikhsya v meditsinskom vuze [Profiles of anxiety and readiness to depression in female students studying at medical school]. *Psihicheskoe zdorovie [Mental health], 15(1)*, 24–28. Retrieved from https://elibrary.ru/item.asp?id = 28371977

Sirota, N.A., Moskovchenko, D.V., Yaltonsky, V.M., Kochetkov, Y.A., & Yaltonskaya, A.V. (2016). *Psikhodiagnostika emotsional'nykh skhem: rezul'taty aprobatsii russkoyazychnoy kratkoy versii shkaly emotsional'nykh skhem R. Likhi* [Psychodiagnostics of emotional schemas: the results of trans-cultural adaptation and assessment of psychometric properties of Russian Version of Leahy Emotional Schema Scale II]. Obozreniye psikhiatrii i meditsinskoy psikhologii imeni V. M. Bekhtereva [Review of psychiatry and medical psychology named after V.M. Bekhterev], 1, 76–83.

Sung, A. D., Collins, M.E., Smith, A.K., Sanders, A.M., Quinn, M.A., Block, S.D., & Arnold, R.M. (2009). Crying: experiences and attitudes of third-year medical students and interns. *Teaching and Learning in Medicine*, *3*(21), 180–187. https://doi.org/10.1080/10401330903014111

Tarabrina, N.V. (2007) *Prakticheskoye rukovodstvo po psikhologii posttravmaticheskogo stressa* [Practical guidance on the psychology of post-traumatic stress]. Moscow, Cogito Center.

Thomas, M.R., Dyrbye, L.N., Huntington, J.L., Lawson, K.L., Novotny, P.J., & Sloan, J.A, et al. (2007). How do distress and well-being relate to medical student empathy? A multicenter study. *Journal of General and Internal Medicine*, *2*(22), 177–183. https://doi.org/10.1007/s11606-006-0039-6

Tschernig, T., Schlaud, M., & Pabst, R. (2000). Emotional reactions of medical students to dissecting human bodies: a conceptual approach and its evaluation. *Anatomical Record*, *261*, 11–13.

Tull, M.T., Barrett, H.M., McMillan, E.S., & Roemer, L. (2007). A preliminary investigation of the relationship between emotion regulation difficulties and posttraumatic stress symptoms. *Behavioral Therapy*, *38*(3), 303–313. https://doi.org/10.1016/j.beth.2006.10.001

Voytsekh, V.F. (2006). *Dinamika i struktura samoubiystv v Rossii* [Dynamics and structure of suicides in Russia]. *Sotsial'naya i klinicheskaya psikhiatriya* [Social and clinical psychiatry], *3*, 22–27.

Wells, A. (1995). Meta-cognition and worry: A cognitive model of generalized anxiety disorder. *Behavioural and Cognitive Psychotherapy*, *3*(23), 301–320. https://doi.org/10.1017/S1352465800015897

Yevdokimova, Ye.G. (2008) *Interpersonal'nyye faktory emotsional'noy dezadaptatsii* [Interpersonal factors of emotional maladjustment in students]. Dissertatsiya kandidata psikhologicheskikh nauk [Ph.D. Thesis]. Moscow.

Zammuner, V.L., Lotto, L., & Galli, C. (2003). Regulation of emotions in the helping professions: nature, antecedents and consequences. *Australian e-Journal for the Advancement of Mental Health*, *1*(2), 43-55. https://doi.org/10.5172/jamh.2.1.43

Original manuscript received February 16, 2018
Revised manuscript accepted October 04, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

From Rare Mutations to Normal Variation: Genetic Association Study of Mathematical, Spatial, and General Cognitive Abilities

Maja Rodic[a,c], Kaili Rimfeld[b], Daria Gaysina[d] & Yulia V. Kovas[a,c,*]

[a] *Tomsk State University, Russia*

[b] *Kings College, London MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology& Neuroscience, London, UK*

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

[d] *EDGE Lab,School of Psychology, University of Sussex, Brighton, UK*

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

Background. Behavioral genetic findings suggest that complex traits, such as mathematical ability, general cognitive ability (intelligence; *g*), and spatial ability, are influenced by many common genetic variants of very small effects that operate across the ability continuum. Common genetic variants may also be responsible for cognitive deficits associated with rare genetic syndromes, in which whole genomic regions may be affected. To date, relatively few common genetic variants involved in cognitive traits have been identified, and these only explain a small proportion of variance in these traits.

Objective. The aim of the study was to find associations between mathematics-related traits and single-nucleotide polymorphisms (SNPs) within chromosomal regions involved in Williams and Prader-Willi disorders. Both disorders are characterized by patterns of weaknesses and strengths in cognitive abilities. Two types of analyses were performed (SNP-based and gene-based), using genotypic and phenotypic data available for 3000 participants from the UK.

Results. SNP-based tests indicated that none of the SNPs passed the demanding multiple testing correction level for any of the phenotypes. Gene-based analysis suggested that 2 pseudogenes (i.e., GOLGA8I and WHAMMP3) were significantly associated with intelligence, and 1 gene (i.e., TUBGCP5) was significantly associated with mathematics at 16 years of age.

Conclusion. The results are consistent with other findings demonstrating that cognitive traits are influenced by many common genetic variants with very small effects. The results also suggest that a small number of these variants may be located in the chromosomal regions affected in Prader-Willi and Williams syndrome regions.

Key words. Mathematical ability, spatial ability, intelligence, genetic variation, single-nucleotide polymorphism (SNP)

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

doi: 10.11621/pir.2018.0410
http://psychologyinrussia.com

Introduction

Twin studies consistently find moderate to high heritability of cognitive traits, such as mathematical ability (e.g., Polderman, Benyamin, De Leeuw, Sullivan, Van Bochoven, Visscher, & Posthuma, 2015; Tosto, Petrill, Halberda, Trzaskowski, et al., 2014; Kovas, Haworth, Dale, & Plomin, 2007; Oliver et al., 2004; Alarcón et al., 2000), general cognitive ability (Deary et al., 2006), and spatial ability (Bratko, 1996; Tosto et al., 2014). Furthermore, the recently developed Genome-wide Complex Trait Analysis (GCTA, also known as the GREML method), that derives heritability directly from DNA of unrelated individuals (Yang et al., 2011a), provides further evidence for genetic influences on cognitive traits, including different aspects of mathematics, general cognitive ability, verbal and non-verbal cognitive ability (Trzaskowski, Davis, DeFries, Visscher, & Plomin, 2013; Plomin et al., 2013). However, heritability obtained using GCTA is usually lower than that found in twin studies, as the GCTA heritability is based only on common single-nucleotide polymorphisms (SNPs), genotyped in a study, and estimates only additive genetic effects. In contrast, the twin method captures both additive and non-additive (e.g., gene-gene interactions) effects, of common and rare variants (Plomin et al., 2013).

According to the 'generalist genes' hypothesis (Plomin & Kovas, 2005) genes have 'general' effects on traits in the following 3 ways: (1) largely the same set of genes affects different aspects of cognitive ability; (2) largely the same set of genes contributes to a particular ability across development; and (3) risk variants are continuously distributed across an ability continuum contributing to individual differences.

These conclusions are based on research that found substantial genetic correlations (r_g) — the extent to which the same genes affect different traits (Plomin, DeFries, Knopik, & Neiderhiser, 2016). For example, $r_g > .60$ for diverse cognitive abilities have been found in twin studies (Calvin et al., 2012) and confirmed by the bivariate GCTA analysis (Trzaskowski et al., 2013). Using this population-based approach based on DNA alone, studies found a genetic correlation of .74 between mathematics and reading (Davis et al., 2009); .74 between general cognitive ability (*g*) and mathematics; and .81 — between g and language (Trzaskowski et al., 2013). Studies have also found that stability across development of such traits as mathematics, literacy, science, and (*g*) is largely explained by continuous genetic effects (Kovas et al., 2007; Haworth et al., 2010).

Much support has also been found for the genetic links between ability and disability (Plomin, Haworth, & Davis, 2009; Plomin & Kovas, 2005). For example, mathematical disability is not seen as a distinct category, but as a lower extreme of the distribution of mathematical ability. An individual's position on such an ability continuum stems from a complex combination of effects of many common (and rare) DNA variants and environments (Kovas, Haworth, Dale, & Plomin, 2007; Butterworth & Kovas, 2013). Such genetic effects have been described as quantitative trait loci (QTLs), the individual DNA markers (locations) linked to complex traits. The QTL approach in molecular genetics has been applied in order to identify the multiple DNA loci associated with complex cognitive traits, such as mathematics or *g* (Plomin et al., 2013a).

The most widely used method for the identification of specific genetic variants implicated in complex traits is the Genome-Wide Association Study (GWAS). This is a hypothesis-free approach that allows for search for genetic markers (i.e. SNPs) of small effects across the whole genome (Hirschhorn & Daly, 2005). GWAS employs SNP microarrays (gene chips), which are used to genotype common genetic variants, usually SNPs that are found in more than 1% of the population. Many loci have already been identified as associated with cognitive traits, but individually they explain <1% of the variance (Plomin & Deary, 2015). To our knowledge, to date, there have been few GWA studies on mathematical ability (Docherty et al., 2010; Davis et al., 2014; Baron-Cohen et al., 2014; Zhu, Chen, Moyzis, Dong, & Lin, 2015) and general cognitive ability (e.g., Rietveld, Esko, Davies, Pers, Turley, Benyamin, Chabris, Emilsson, Johnson, Lee, and De Leeuw, 2014; Butcher et al., 2008; Davis et al., 2015) and none on spatial ability. The first GWAS on mathematics reported 10 nominally significant SNPs associated with mathematical variation in a sample of 10-year-olds (Docherty et al., 2010). Another study reported 5 SNPs that were associated with mathematical ability of 16-year-olds (Baron-Cohen et al., 2014). In both studies, the individual SNPs did not reach the genome-wide significance level and the effect sizes of individual SNPs were very small (i.e., .02% and OR = 1.63, respectively).The studies had relatively small sample sizes and were underpowered to detect SNPs of very small effects associated with the phenotype.

GWASs on general intelligence show a similar picture. One of the first GWASs on general intelligence (*g*) reported 6 nominally significant SNPs that individually explained between .1 and .4% of the variance in *g*, however, only 1 survived multiple comparison corrected p-value (Butcher et al., 2008). Other large studies of intelligence differences did not find genome-wide significant associations and found SNPs that explain only .2 to .5% of the variance (Rietveld et al. 2014; Benyamin et al., 2014; Desrivieres et al., 2014). Several recent studies on individual differences in how long a person stays in education (years of education) identified many DNA variants (Okbay, et al., 2016; Lee et al., 2018). These variants, discovered for years of education, are also associated with general cognitive ability and other related traits (e.g. Lee et al., 2018; Selzam et al., 2016; Okbay et al., 2016).

Although individual SNPs explain only a small amount of variance, the effects of these polymorphisms are often additive, which allows for aggregation of multiple variants into polygenic scores (Plomin & Deary, 2014). Such polygenic scores are typically referred in the literature as polygenic risk scores (PRS) or genome-wide polygenic scores (GPS). Several recent studies have shown that polygenic scores explain more variance than each variant individually, and with more SNPs included in the polygenic score, more variance in a trait is explained. For example, in contrast to individual SNPs that explained <1% of the variance, the set of 10 SNPs, ~~found in Docherty et al. study (2010)~~ accounted for 2.9% of the phenotypic variance in mathematical ability (Docherty et al., 2010). Moreover, one-third of the children who carried 10 or more risk alleles were nearly twice as likely to be in the lowest 15% of the distribution. Recently, a study employing

genome-wide polygenic scores, created using GWAS results on number of years of education, explained 9% of individual differences in exam performance at age 16 (Selzam et al., 2016).

Although polygenic scores already explain a substantial amount of variance in some traits, they also highlight the problem of 'missing heritability' - the gap between heritability estimates identified by quantitative genetic research and the GWAS-identified associations (Maher, 2008). Multiple reasons for 'missing heritability' have been proposed (Maher, 2008). For example, variation in complex traits results from many common variants of small effect, other types of common variants, and rare variants of small or large effects.

Although GWASs present a powerful tool for uncovering the genetic architecture of complex traits, they require very large samples in order to have the power to detect very small effects. One strategy to increase power for discovering new variants linked to cognitive traits is to focus on specific genomic regions that have been implicated in disorders characterized by specific patterns of cognitive deficits and relative advantages. This research approach is based on the assumption that common variants within these regions may be related to normal variation in cognitive ability. If this is true, then these can potentially be detected with a relatively small sample because the correction for multiple testing only involves known SNPs in these areas, rather than SNPs across the whole genome.

The current study adopts this approach and focuses on the chromosomal regions known to be affected in 2 genetic disorders: Prader-Willi syndrome (PWS) and Williams syndrome (WS). Both disorders are associated with disproportionately low mathematical skills, as well as intellectual impairment, with mean IQ scores for PWS ranging from 60 to 70 (Cassidy, 1997) and for WS ranging from 40 to 90 (Howlin, Davies, & Udwin, 1998).

PWS is a rare genetic disorder that is caused by the deletion of critical genes on the paternal copy of chromosomal region 15q11-13, with the genes on the maternal copy silenced (inactive). People with PWS show profound deficits in mathematical skills and short-term memory. All mathematical domains are affected, even beyond the impairment expected from the general cognitive level (Sulzbacher, Wong, McKeen, Glock, & MacDonald, 1981; Bertella et al., 2005). Interestingly, reading and visuo-spatial ability show relative strengths (Cassidy, 1997), with particularly good performance in solving jigsaw puzzles (Dykens, 2002).

WS is a genetic disorder caused by micro-genetic deletion on the long arm of chromosome 7 (7q11.23), with 25 to 28 genes missing from this region (Ewart et al., 1993). People with WS have poor mathematical ability, number skills, planning and problem solving, spatial working memory, and visuo-spatial long-term memory. In contrast, their spoken language is well developed (Ansari et al., 2003; Bellugi, Sabo, & Vaid, 1998; Mesulam, 1982).

We hypothesized that the genetic regions implicated in these syndromes might contain common DNA variants that are also associated with normal variation in mathematical ability, spatial ability, and general cognitive ability.

Two types of analyses were performed. We first ran an SNP association analysis across the selected regions, including all available SNPs. We then ran a gene-based analysis in which the effects of multiple SNPs were combined into genes (a functional unit of human genome) and tested for association, improving the power of the SNP-based method because it also takes into consideration the correlations among SNPs within a single gene. We tested whether the genes found in the 2 syndromic chromosomal regions show an association with the available phenotypes.

Methods

Participants

The genotypic and phenotypic data for a sample of 3,152 individuals was drawn from the Twin Early Development Study (TEDS), a longitudinal study of more than 11,000 twins born in England and Wales in 1994, 1995, and 1996 (Oliver & Plomin, 2007). The data were obtained when the participants were 12 and 1 6years of age. The TEDS sample is representative of the UK population in terms of demographic characteristics, based on the comparisons to the UK census data (Haworth, Davis & Plomin, 2013; Harlaar, Hayiou-Thomas, & Plomin, 2005). Data were excluded from the analyses of participants who had severe medical problems, specific chromosomal abnormalities, or had severe complications at birth. To decrease heterogeneity of ancestry (which may lead to the problem of population stratification in association studies), only data from families who declared themselves as white and whose first language was English were used for the analyses. The sample size varied depending on available data for each phenotype: from 1,112 for general cognitive ability at age 16 years to 2,906 for the composite mathematics score at age 12. The data from only 1 twin from each pair was used for the analyses to prevent phenotypic covariance of siblings affecting the estimates (Visscher, Yang, Goddard, et al. 2010).

Written parental consent was obtained before data collection and the project received approval from the Institute of Psychiatry ethics committee (05/Q0706/228).

Measures

Twelve phenotypes were tested for associations with SNPs in the 2 genetic regions: 6 mathematics-related measures (3 at age 12 and 3 at age 16), 2 tests of spatial ability, and 1 test of general cognitive ability at age 12; 2 tests of general cognitive ability and 1 test of spatial ability at age16. The information about phenotypes and total sample sizes can be seen in Table 1 (for more detail, see Appendix 1).

Table 1

Phenotypic and genotypic data available

Phenotypes assessed at age 12 (N)	Phenotypes assessed at age 16 (N)	Available SNPs
Mathematics Test — total score for «understanding numbers», «number and algebra», and «shape, space, and measures» (N = 2,270)	**Mathematics Composite** — composite score of problem verification and understanding number tasks (N = 1,144)	**26, 478 SNPs** — Williams syndrome region (7q11.22)
Mathematical Achievement — teacher rated «using and applying mathematics», «number and algebra», «shape, space,& measures», and «handling data» (N = 1,705)	**Number Sense** — non-symbolic comparison of numerosity (N = 1,132)	**19,389 SNPs** — Prader-Willi syndrome region (15q11–13)
Mathematics Composite — an average score of mathematics test and mathematical achievement (N = 2,904)	**Number Line** — an estimation of numerical magnitudes (N = 1,257)	
Jigsaw Puzzle — spatial reasoning test (N = 2,627)	**Ravens Progressive Matrices** — test non-verbal (fluid) intelligence (N = 1125)	
Spatial Composite — first principal component of Hidden Shapes and Jigsaw (N = 2,627)	**Corsi Block** — of visuo-spatial short-term working memory(N = 1,111)	
General Cognitive Ability 'g' — General Knowledge, vocabulary, Ravens,and Picture Completion (N = 2,328)	**General cognitive ability 'g'** — (composite of the Vocabulary and Ravens scores(N = 1111)	

Note. See Appendix 1 for the brief description of tests.

Available genotypic data

Affymetrix GeneChip 6.0 DNA arrays were used for genotyping nearly 700,000 SNPs extracted from buccal cheek swabs of 3,655 unrelated individuals by Wellcome Trust Sanger Institute (Hinxton, UK) as part of the Wellcome Trust Case Consortium (See Trzaskowski et al., 2013 for full details). In addition, imputation was carried out using the IMPUTE version 2.3.0 software (Howie, Donnelly, & Marchini, 2009) on the genotype data after application of quality control procedures, using reference panel data from 1,000 genomes (Phase I, v3, build 37 (hg19)) (Siva, 2008). Of the imputed SNPs, only those that had an information score ≥.98 were retained (see Krapohl et al., 2016 for full details). All available SNPs in the regions associated with the 2 genetic syndromes were used in the analyses: 26,478 SNPs in 7q11.22 area (WS); 19,389 SNPs in 15q11-13 area (PWS). The standard quality control criteria were applied for genotyping data, excluding those with minor allele frequency less than 1%, those with more than 2% of missing data, and those deviating from Hardy-Weinberg equilibrium ($p < 10^{-3}$).

Analyses

SNP-based analysis

We performed an association analysis for SNPs in 2 genetic regions, using linear regression models under an additive genetic model, using Plink software (Purcell, 2007). Participants' sex and ancestry (first 8 principal components of the genotype data) were included as covariates in the regression model. Phenotypic data were normalized using the van der Waerden transformation (van der Waerden, 1952) due to a slight skew in the distribution of scores. The association analysis was conducted for all 12 phenotypes presented in Table 1: 6 phenotypes at age 12 (3 math tasks, 2 spatial tasks, and 1 general cognitive ability); and 6 at age 16 (3 mathematics-related phenotypes, 2 general cognitive ability, and 1 visuo-spatial ability). The results were corrected for multiple testing using Bonferroni correction, with the corrected p-values of 1.8×10^{-6} (0.05/26,478) for WS region and 2.5×10^{-6} (0.05/19,389) for the PWS region.

Gene-based analysis

Gene-based analysis (VEGAS2; Mishra & Macgregor, 2014) was performed by combining the p-values of the SNPs within a gene (for all genes available in the 2 chosen genetic regions separately), to obtain an overall p-value for the association of the entire gene, accounting for the correlations between SNPs. The linkage disequilibrium (LD) was considered using the 1000 Genome II CEU (NCBI build 37) reference panel for each gene and the 50kb boundary. The statistical significance of this parameter was calculated using simulations.

This analysis leads to an increased power and has the potential to explain a greater proportion of variance (Hill, et al., 2014b; Liu, et al., 2010). In addition, gene-based analysis could be advantageous in that a gene (a functional unit of the human genome) is highly consistent across populations, unlike genetic variants, which differ across diverse human populations in allele frequencies, LD structure, and heterogeneity. Moreover, the problem of multiple-testing is reduced, in our case from several thousand SNPs to only 100 to 200 genes (Hill et al., 2016).

There were 110 genes available in the WS region and 194 genes in the Prader-Willi region. The alpha level was therefore set to be 0.00045 (i.e., 0.05/110) for the WS region and 0.00025 (i.e., 0.05/194) for the Prader-Willi region.

Results

SNP-based results

None of the SNP associations met the demanding significance criterion corrected for multiple testing for any of the phenotypes. According to the power calculation, using the Genetic Power Calculator (Purcell & Sham, 2003), there was 80% power to detect effect sizes between 1.8 and 1.9% in the largest sample available to us (N = 2906), with p-value threshold corrected for the number of SNPs tested in WS and PWS regions. However, none of the SNPs had such (relatively) large effect sizes. This is in line with previous findings for quantitative traits with strongest as-

sociations accounting for less than 1% of the variance in height (Lango Allen et al., 2010), general cognitive ability (Meaburn, Harlaar, Craig, Schalkwyk, & Plomin, 2007), and mathematics (Docherty et al., 2010).

Gene-based results

Gene based analyses showed that 1 gene and 2 pseudogenes from the Prader-Willi region (chromosome 15) reached the significance level (p=0.05/194=0.00025) for association with phenotypes at age 16 (see Table 2).

Two pseudogenes, GOLGA8I and WHAMMP3, were both significantly associated with Ravens and general cognitive ability scores. The TUBGCP5 gene was significantly associated with composite math score.

The complete results of gene-based analyses can be seen in Appendix 2.

Table 2

Gene-based analysis results for the Williams syndrome (7q11.22) and Prader-Willi areas (15q11-13)

Chr	Gene ID	n SNPs	Start	Stop	Ravens	General cognitive ability	Composite Math
15	GOLGA8I	20	23205241	23312743	7.80E-05	4.00E-05	n/s
15	WHAMMP3	23	23137728	23258357	8.30E-05	4.70E-05	n/s
15	TUBGCP5	165	22783394	22923891	n/s	n/s	6.80E-05

Note. Only genes that survived the correction for multiple testing were reported. Two genes were significantly associated with Ravens and 'g', and 1 was significantly associated with composite math score, all at age 16. Start and end positions do not include the ± 50 kb boundary.

Discussion

Molecular genetic studies to date have identified only a few SNPs, explaining the small amount of variance in mathematical and spatial ability. Moreover, much of the genetic variance in general cognitive ability also remains unidentified. To reach the statistical power necessary to detect very small effect sizes of individual SNPs, GWAS studies require very large samples. The present study aimed to identify genetic polymorphisms (SNPs) associated with cognitive traits by examining the variations within chromosomal regions implicated in Williams and Prader-Willi disorders. This approach was based on the hypothesis that genetic regions implicated in these syndromes contain genetic variants that are significantly associated with mathematical, spatial, and general cognitive abilities.

Firstly, we tested these genomic regions for single SNP associations. Secondly, we sought to increase the power to detect SNPs implicated in these cognitive traits by grouping SNPs into gene-based scores and considering the mas a unit for the association, taking into consideration correlations among SNPs within a single gene.

The results from the first method confirm the absence of large effects of common genetic variants on any of the phenotypes included in this study. In line with previous research (Butcher, Davies, & Highton, 2006; Docherty et al., 2010), and similar to other quantitative traits, mathematical abilities, general cognitive ability, and spatial ability are likely to be affected by many genetic variants of small effects (QTLs). Because loci are expected to survive stringent corrections for multiple testing, it is likely that a large number of loci remain unidentified (e.g., considered as false negatives).

One explanation for such small (and currently undetectable) effects is that natural selection eliminates mutations of large negative effects on these fitness-relevant traits. This explanation may be relevant to our study, as all explored phenotypes have been associated with desirable life outcomes (e.g. Parsons & Bynner, 2005; Wai et al., 2009; Gottfredson & Deary, 2004;).

It is also possible that regions which contain rare mutations of large effects on cognitive traits do not contain genetic variants responsible for the normal range (Davis et al., 2015). For example, 1 study found that 40 genes related to non-syndromic autosomal recessive intellectual disability, causing extensive and deleterious effects on intelligence, are not associated with the whole range of intelligence (Reichenberg, Cederlöf, McMillan, Trzaskowski, Kapara, Fruchter, Ginat, Davidson, Weiser, Larsson, & Plomin, 2016). This suggests qualitative as opposed to quantitative differences caused by some genetic variants.

The results of the gene-based analyses indicated that on chromosome 15 (Prader-Willi region), 1 gene (TUBGCP5) was related to mathematics score, and 2 pseudogenes (GOLGA8I and WHAMMP3) were related to Ravens matrices (non-verbal ability) and general cognitive ability at age 16. Gene TUBGCP5 encodes gamma-tubulin complex-associated protein 5, which is essential for microtubule nucleation at the centrosome. The deletion of this gene and other genes (i.e., NIPA1, NIPA2, and CYFIP1) is well recognized in PWS and may be linked to observed cognitive and behavioral deficits (Elert-Dobkowska et al., 2014).

It is commonly found that diseases are associated with genetic variants from protein-coding regions; however, the majority of them (80%) fall outside coding regions, highlighting the importance of including both coding and non-coding regions in the search for genetic variants (Hindorff, et al., 2009).

Although most pseudo genes are mostly non-functional (Pink et al., 2011), 2 to 20% of them may be transcribed (Yano et al., 2004; Harrison et al., 2005; Zheng et al., 2005, 2007). For example, recurrent deletions of chromosome 15q13.3 are associated with intellectual disability, schizophrenia, autism, and epilepsy. Investigations into this region show 5 relatively recent structural configurations (.5 to .9 million years ago) ranging in size from 2 to 3 Mb. They are thought to be a result of human-specific expansions of segmental duplications in 2 independent inversion events, and all inversion breakpoints map near GOLGA8 core duplicons. Some evidence suggests that there is a mechanistic role of this core duplicon and its palindromic architecture in promoting the evolutionary and disease-related instability of chromosome 15 (Antonacci et al., 2014).

Conclusion

In the present study, none of the common genetic variants in PWS and WS regions were related to individual differences in cognition in the general population. However, the results from the gene-based analyses suggested that some variance in the cognitive traits may be explained by common SNPs in these regions.

Future studies may investigate genetic variation in regions implicated in other known genetic syndromes. Until whole genome sequencing becomes common practice, the approach of focusing on the genetic regions implicated in rare genetic disorders may be a beneficial method for genetic studies of complex traits. Future studies with larger sample sizes should also investigate other genetic variants, such as structural variants (e.g., copy-number variants), as well as gene-gene interactions. Future research should also test whether polygenic scores based on the associations observed in this study can explain more variance in cognitive phenotypes (Selzam et al., 2016; Domingue et al., 2015; de Zeeuw et al., 2014). Understanding genetic mechanisms underlying individual differences in cognition is an important step toward optimizing education for all learners.

Limitations

Although the study was adequately powered to detect large effects, the study lacked power to detect variants of very small effect. Future studies with larger sample sizes are needed to address this issue, especially as the individual effect of SNPs are proving to be much smaller (<1%). For example, Visscher et al. (2010) suggest that a sample size of 100,000 will be needed to have the statistical power of 80% to detect genetic variants accounting for as little as 0.04% of the variance in a trait at a "genome-wide significance level" of $p < 5 \times 10^{-8}$.

Moreover, the approach used in the present study does not capture effects of the rarer variants (<1% of the population) of small effect, as well as those of other types of variants (e.g., structural). New gene chips are needed that will tag these variants. Future research may rely on next generation sequencing techniques, which will capture the whole DNA variation, regardless of rarity.

Acknowledgments

We gratefully acknowledge the continuing contribution of the Twin Early Development Study (TEDS) families. We are also very grateful to the Wellcome Trust Case Consortium 2 team for their help with genotyping and quality control of our data. This research was supported by a grant from the Russian Science Foundation (project 14-48-00043) to Tomsk State University.

References

Alarcon, M., Knopik, V.S., & DeFries, J.C. (2000). Covariation of mathematics achievement and general cognitive ability in twins. *Journal of School Psychology, 38*, 63–77. https://doi.org/10.1016/S0022-4405(99)00037-0

Ansari, D., Donlan, C., & Karmiloff-Smith, A. (2007). Atypical and typical development of visual estimation abilities. *Cortex: Special Issue on Selective Developmental Disorders, 6*, 758–768. https://doi.org/10.1016/S0010-9452(08)70504-5

Ansari, D., Donlan, C., Thomas, M.S., Ewing, S.A., Peen, T., & Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuo-spatial contributions to typical and atypical number development. *Journal of Experimental Child Psychology, 85,* 50–62. https://doi.org/10.1016/S0022-0965(03)00026-2

Baron-Cohen, S., Murphy, L., Chakrabarti, B., Craig, I., Mallya, U., Lakatosova, S., & Warrier, V. (2014). A Genome Wide Association Study of Mathematical Ability Reveals an Association at Chromosome 3q29, a Locus Associated with Autism and Learning Difficulties:A Preliminary Study. *Plos One, 9*(5), e96374. https://doi.org/10.1371/journal.pone.0096374

Bellugi, U., Sabo, H., & Vaid, J. (1998). Spatial deficits in children with Williams syndrome. In J. Stiles-Davis, M. Kritchevsky, & U. Belugi (Eds.), *Spatial cognition brain bases and dvelopment* (pp. 273–297): Hillsdale, NJ: Lawrence Erlbaum Associates.

Bertella, L., Girelli, l., Grugni, G., Marchi, S., Molinari, E., & Semenza, C. (2005). Mathematical skills in Prader-Willi syndrome. *Journal of Intellectual Disability Research, 49*, 159–169. https://doi.org/10.1111/j.1365-2788.2004.00634.x

Bortolotto, Z.A., Clarke, V.R.J., Delany, C.M., Parry, M.C., Smolders, I., Vignes, M. & Collingridge, G.L. (1999). Kainate receptors are involved in synaptic plasticity. *Nature, 402*, 297–301. https://doi.org/10.1038/46290

Bratko, D. (1996). Twin study of verbal and spatial abilities. *Personality and Individual Differences*, *21*(4), 621–624. https://doi.org/10.1016/0191-8869(96)00091-8

Butcher, C., Davies, C., & Highton, M. (2006). *Designnig Learning. From moduleoutline to effective teaching.* London and New York: Routlege. https://doi.org/10.4324/9780203968482

Butcher, L.M., Davis, O.S. P., Craig, I.W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. *Genes, Brain and Behavior, 7*, 435–446. https://doi.org/10.1111/j.1601-183X.2007.00368.x

Calvin, C., Deary, I., Webbink, D., Smith, P., Fernandes, C., Lee, S., … & Visscher, P. (2012). Multivariate genetic analyses of cognition and academic achievement from two population samples of174,000 and 166,000 school children. *Behavioural genetics, 42*(5), 699–710. https://doi.org/10.1007/s10519-012-9549-7

Cassidy, S.B. (1997). Prader-Willi Syndrome. *Journal of Medical Genetics. 34*, 917–923. https://doi.org/10.1136/jmg.34.11.917

Chabris, C.F., Hebert, B.M., Benjamin, D.J., Beauchamp, J.P., Cesarini, D., van der Loos, M.J.H.M., & Laibson, D. (2011). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science, 23*(11), 1314–1323. https://doi.org/10.1177/0956797611435528

Cooper, D.N., Chen, J.M., Ball, E.V., Howells, K., Mort, M., Phillips, A.D., … & Stenson P.D. (2010). Genes, mutations, and human inherited disease at the dawn of the age of personalizedgenomics. *Human Mutation, 31*, 631–655. https://doi.org/10.1002/humu.21260

Curfs, L.G., Wiegers, A.M., Sommers, J.R., Borghgraef, M., & Fryns, J. P. (1991). Strengths and weaknessesin the cognitive profile of youngsters with Prader-Willisyndrome. *Clinical Genetics, 40*, 430–434. https://doi.org/10.1111/j.1399-0004.1991.tb03114.x

Davis, O.S.P., Butcher, L.M., Docherty, S.J., Meaburn, E.M., Curtis, C.J.C., Simpson, A., Plomin, R. (2010). A three-stage genome-wide association study of general cognitive ability:

Hunting the small effects. *Behavior Genetics, 40*, 759–767. https://doi.org/10.1007/s10519-010-9350-4

Davis, O.S., Haworth, C.M., & Plomin, R. (2009). Learning abilities and disabilities: Generalist genes in early adolescence. *Cognitive neuropsychiatry*, *14*(4–5), 312–331. https://doi.org/10.1080/13546800902797106

Deary, I.J., Spinath, F.M., & Bates, T.C. (2006). Genetics of intelligence. *European Journal of Human Genetics*, *14*(6), 690–700. https://doi.org/10.1038/sj.ejhg.5201588

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences, 95*(24), 14529–14534. https://doi.org/10.1073/pnas.95.24.14529

Docherty, S. J., Davis, O. P. S., Kovas, Y., Meaburn, E. L., Dale, S. P., Petrill, S. A., … Plomin, R. (2010). A genome-wide association study identifies multiple loci associated with mathematics ability and disability. *Genes, Brain and Behaviour, 9*, 234-247.

Dykens, E. M. (2002). Are jigsaw puzzles skills "spared" in persons with Prader–Willi syndrome? *Journal of Child Psychology and Psychiatry, 43*, 343–352. https://doi.org/10.1111/1469-7610.00025

Dykens, E. M., & Kasari, C. (1997). Maladaptive behavior in children with Prader-Willi syndrome,Down syndrome, and nonspecific mental retardation. *American Journal on Mental Retardation, 102*, 228–237. https://doi.org/10.1352/0895-8017(1997)102<0228:MBICWP>2.0.CO;2

Ewart, A.K., Morris, C.A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., … & Keating, M.T. (1993). Hemizygosity at the elastin locus in a developmental disorder, Williams syndrome. *Nature Genetics, 5*(11), 11–16. https://doi.org/10.1038/ng0993-11

Gottfredson, L.S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science, 13*(1), 1–4. https://doi.org/10.1111/j.0963-7214.2004.01301001.x

Gross-Tsur, V., Manor, O., & Shalev, R.S. (1996). Developmental dyscalculia: prevalence and demographic features. *Dev Med Child Neurol, 38*, 25–33. https://doi.org/10.1111/j.1469-8749.1996.tb15029.x

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the 'number sense': The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental psychology, 44*, 1457–1465. https://doi.org/10.1037/a0012682

Harlaar, N., Hayiou-Thomas, M. E., & Plomin, R. (2005). Reading and general cognitive ability: a multivariate analysis of 7-year-old twins. *Scientific Studies of Reading, 9*, 197–218. https://doi.org/10.1207/s1532799xssr0903_2

Haworth, C.M.A., Wright, M.J., Luciano, M., Martin, N.G., De Geus, E.J.C., Van Beijsterveldt, C.E.M., … & Plomin. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular psychiatry*, *15*(11), 1112–1120. https://doi.org/10.1038/mp.2009.55

Hirschhorn, J.N., & Daly, M.J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature reviews genetics, 6(2)*, 95-108. https://doi.org/10.1038/nrg1521

Hornef, N., Olbrich, H., Horvath, J., Zariwala, M.A., Fliegauf, M., Loges, N.T., … & Omran, H. (2006). DNAH5 mutations are a common cause of primary ciliarydyskinesia with outer dynein arm defects. *American Journal of Respiratory and Critical Care Medicine, 174*, 120–126. https://doi.org/10.1164/rccm.200601-084OC

Howie, B. N., Donnelly, P., & Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *Plos Genetics 5*. doi: e1000529.

Howlin, P., Davies, M., & Udwin, O. (1998). Cognitive functioning in adults with Williams syndrome. *Journal of Child Psychology and Psychiatry, 39*, 183–189. https://doi.org/10.1017/S0021963097001789

Huettner, J. E. (2003). Kainate receptors and synaptic transmission. *Progress in Neurobiology, 70*, 387–407. https://doi.org/10.1016/S0301-0082(03)00122-9

Ibanez-Tallon, I., Pagenstecher, A., Fliegauf, M., Olbrich, H., Kispert, A., Ketelsen, U. P., … & Omran, H. (2004). Dysfunction of axonemal dynein heavy chain Mdnah5inhibits ependymal flow and reveals a novel mechanism for hydrocephalusformation. *Human molecular genetics, 13*, 2133–2141. https://doi.org/10.1093/hmg/ddh219

Kovas, Y., Haworth, C.M.A., Dale, P.S., Petrill, S.A., & Plomin, R. (2007b). Mathematical ability of 10-year-old boys and girls: genetic and environmental etiology of typical and low performance. *Journal of Learning Disabilities, 40*, 554–567.

Kovas, Y., Haworth, C.M.A., Dale, P.S., & Plomin, R. (2007). The genetic and enviromental origins of learning abilities and disabilities in early years. *Monographs of the Society for Research in Child Development, 72*, 1–158. https://doi.org/10.1177/0022219407040006060l

Kovas, Y., Haworth, C.M.A., Harlaar, N., Petrill, S.A., Dale, P.S., & Plomin, R. (2007). Overlap and specificity of genetic and environmental influenceson mathematics and reading disability in 10-year-old twins. *Journal of Child Psychol Psychiatry. 48*, 914–922. https://doi.org/10.1111/j.1469-7610.2007.01748..x

Kovas, Y., Petrill, S.A., & Plomin, R. (2007c). The origins of diverse domains of mathematics: generalist genes but specialist environments. *Journal of Educational Psychology, 99*, 128–139. https://doi.org/10.1037/0022-0663.99.1.128

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., … & Oli-Carsten Oliver (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature, 467,* 832–838. https://doi.org/10.1038/nature09410

Law, M., Stewart, D., Pollock, N., & (1998). Critical review form – Quantitative studies: McMaster University: Occupational Therapy Evidence-Based Practice Research Group.

Lee , J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., et al. (2018).Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics, 50,* 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Li, M. X., Yeung, J. M., Cherny, S. S., & Sham, P. (2012). Evaluating the effective number of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human genetics, 13*(5), 747–756. https://doi.org/10.1007/s00439-011-1118-2

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature, 456*, 18–21. https://doi.org/10.1038/456018a

Marchini, J., Howie, B. N., Myers, S., McVean, G., & Donnelly, P. (2007). A new multipoint method for genome-wide association studies via imputation of genotypes. *Nature Genetics, 38*, 906–913. https://doi.org/10.1038/ng2088

Meaburn, E.L., Harlaar, N., Craig, I.W., Schalkwyk, L.C., & Plomin, R. (2007). Quantitative trait locus association scan of early reading disability and ability using pooled DNA and 100K SNP microarrays in a sample of 5760 children. *Molecular Psychiatry, 13*, 729–740. https://doi.org/10.1038/sj.mp.4002063

Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology, 11,* 592–598. https://doi.org/10.1002/ana.410110607

Miyajima, F., Ollier, W., Mayes, A., Jackson, A., Thacker, N., Rabbitt, P., … & Payton, A. (2008a). Brain-derived neurotrophic factor polymorphism Val66Met influences cognitive abilities in the elderly. *Genes, brain and behaviour, 7*, 411–417. https://doi.org/10.1111/j.1601-183X.2007.00363.x

Miyajima, F., Quinn, J.P., Horan, M., Pickles, A., Ollier, W.E., Pendleton, N., & Payton, A. (2008b). Additive effect of BDNF and REST polymorphisms is associated with improved general cognitive ability. *Genes, brain and behaviour, 7*, 714–719. https://doi.org/10.1111/j.1601-183X.2008.00409.x

Morris, C.A., Dempsey, S.A., Leonard, C.O., Dilts, C., & Blackburn, B.L. (1988). Natural history of Williams syndrome: Physical characteristics. *Journal of Paediatric Medicine, 113*, 318–326. https://doi.org/10.1016/S0022-3476(88)80272-5

Murphy, M. M., & Mazzocco, M. M. M. (2008). Rote numeric skills may mask underlying mathematical disabilities in girls with fragile X syndrome. *Developmental Neuropsychology, 33*, 345–336. https://doi.org/10.1080/87565640801982429

NFER. (2002). NferNelson Publishing Company Ltd.

nferNelson. (1994). *Maths 5-14series*. London: nferNelson Publishing Company Ltd.

O'Hearn, K., & Luna, B. (2009). Mathematical skills in Williams syndrome: insight into the importance of underlying representations. *Developmental disabilities research reviews, 15(1)*, 11–20. https://doi.org/10.1002/ddrr.47

Okbay, A., J. P. Beauchamp, M. A. Fontana, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533* (7604): 539–542. https://doi.org/10.1038/nature17671

Oliver, B., Harlaar, N., Hayiou-Thomas, M.E., Kovas, Y., Walker, S.O., Petrill, S.A., … & Plomin, R. (2004). A twin study of teacher-reported mathematics performance and low performance in 7-year-olds. *Journal of Educational Psychology, 96, 504–517*. https://doi.org/10.1037/0022-0663.96.3.504

Oliver, B., & Plomin, R. (2007). Twins' Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behaviour problems from childhood through adolescence. *Twin Research and Human Genetics, 10*, 96–105. https://doi.org/10.1375/twin.10.1.96

Opfer, J.E., & Siegler, R.S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195. https://doi.org/10.1016/j.cogpsych.2006.09.002

Paglin, M., & Rufolo, A. (1990). Heterogeneous human capital occupational choice and male-female earning differences. *Journal of labor economics, 8*, 123–144. https://doi.org/10.1086/298239

Pagulayan, F.K., Busch, R.M., Medina, K.L., Bartok, J.A., & Krikorian, R. (2006). Developmental Normative Data for the Corsi Block-Tapping Task. *Journal of Clinical and Experimental Neuropsychology, 28*(6), 1043–1052. https://doi.org/10.1080/13803390500350977

Parsons, S., & Bynner, J. (2005). Does numeracy matter more? London: National Research and Development Centre for adult literacy and numeracy.

Paterson, S.J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47*, 190–204. https://doi.org/10.1111/j.1469-7610.2005.01460.x

Plomin, R. (2013). Child Development and Molecular Genetics: 14 Years Later. *Child development, 84*, 104–120. https://doi.org/10.1111/j.1467-8624.2012.01757.x

Plomin, R., DeFries, J.C., Knopik, V.S., & Neiderhiser, J.M. (2016). *Behavioral genetics*. New York: Worth Publishers.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131*(4), 592–617. https://doi.org/10.1037/0033-2909.131.4.592

Polderman, T J., Benyamin, B., De Leeuw, C.A., Sullivan, P.F., Van Bochoven, A., Visscher, P.M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics*, *47*(7), 702–709. https://doi.org/10.1038/ng.3285

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., & Ferreira, M.A.R. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics 8, 81*, 559–575.

Purcell, S. C., S.S. Sham, P.C. (2003). Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics, 19*(1), 149–150. https://doi.org/10.1086/519795

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford: Oxford Psychologists Press.

Rietveld, C.A., Medland, S.E., Derringer , J., Yang, J., Esko, T., Martin, N.W., … & Koellinger, P. D. (2013). GWAS of 126, 559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science, 340*, 1467–1471.

Sahoo, T., del Gaudio, D., German, J.R., Shinawi, M., Peters, S.U., Person, R.E., … & Beaudet, A. L. (2008). Prader–Willi phenotype caused by paternal defciency for the HBII-85 C/Dbox small nucleolar RNA cluster. *Nature Genetics, 40*, 719–721. https://doi.org/10.1038/ng.158

Semenza, C. (2008). Number Processing. In B. Stemmer & H. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 219–227). Oxford: Elsevier. https://doi.org/10.1016/B978-0-08-045352-1.00021-5

Sulzbacher, S., Wong, B., McKeen, J., Glock, J., & MacDonald, B. (1981). Long term therapeutic effects of a three month intensive growth group. *The Journal of Clinical Psychiatry, 42(4)*, 148–153.

Thompson, L.A., Detterman, D.K., & Plomin, R. (1991). Associations between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science,* 158–165. https://doi.org/10.1111/j.1467-9280.1991.tb00124.x

Tosto, M.G., Petrill, S.A., Halberda, J., Trzaskowski, M., Tikhomirova, T.N., Bogdanova, O.Y., Ly, R., Wilmer, J.B., Naiman, D.Q., Germine, L., Plomin, R., & Kovas, Y. (2014). Why do we differ in number sense? Evidence from a genetically sensitive investigation. *Intelligence*, *43*(100), 35–46. https://doi.org/

Trzaskowski, M., Davis, O.S.P., DeFries, J.P., Visscher, P.M., & Plomin, R. (2013). DNA Evidence for strong genome-wide pleiotropy of cognitive and learning abilities. *Behavior genetics, 43*(4), 267–273. https://doi.org/10.1007/s10519-013-9594-x

van der Waerden, B.L. (1952). Order tests for the two-sample problem and their power. *Indagationes Mathematicae*, 453–458. https://doi.org/10.1016/S1385-7258(52)50063-5

Van Herwegen, J., Ansari, D., Xu, F., & Karmiloff-Smith, A. (2008). Small and larger number processing in infants and toddlers with Williams syndrome. *Developmental Science, 11*, 637–643. https://doi.org/10.1111/j.1467-7687.2008.00711.x

Varela, M.C., Kok, F., Setian, N., Kim, S.A., & Koiffmann, C.P. (2005). Impact of molecular mechanisms, including delition size on Prader-Willi syndrome phenotype:study of 75 patients. *Clinical Genetics, 67*, 47–52. https://doi.org/10.1111/j.1399-0004.2005.00377.x

Visscher, P.M., Yang, J., & Goddard, M.E. (2010). A commentary on "Common SNPs explain a large proportion of the heritability for human height" by Yang et al. (2010). *Twin research and Human Genetics, 13*, 517–524. https://doi.org/10.1375/twin.13.6.517

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817–835. https://doi.org/10.1037/a0016127

Whittington, J.E., Holland, A.J., Webb, T., Butler, J., Clarke, D., & Boer, H. (2001). Population prevalence and estimated birth incidence and mortality rate for people with Prader—Willi syndrome in one UK Health Region. *Journal of Medical Genetics, 38*, 792–798. https://doi.org/10.1136/jmg.38.11.792

Yang, J., Manolo, T.A., Pasquale, L.R., Boerwinkle, E., Caporaso, N., Cunningham, J.M., … & Vischer, P.M. (2011a). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics, 43(6)*, 519–525. https://doi.org/10.1038/ng.823

Original manuscript received August 17, 2018
Revised manuscript accepted October 15, 2018
First published online December 30, 2018

Appendix 1

Description of phenotypic measures

Age 12 phenotypes

Mathematics Test — administered on the web, based on nferNelson (1994) Maths 5-14 Series (for details of the test see Haworth, Harlaar, Kovas, Davis, Oliver, Hayiou-Thomas, Busfield, McMillan, Dale, and Plomin, 2007).

Mathematical Achievement — rated by teachers, reflected children's school performance according to the UK National Curriculum criteria (for details of the test see Haworth, Harlaar, Kovas, Davis, Oliver, Hayiou-Thomas, Busfield, McMillan, Dale, and Plomin, 2007).

Mathematics Composite — the 3 components of the web test and the 4components of the teacher-rated achievement were previously found to be highly correlated (e.g. Kovas, Haworth, Petrill & Plomin, 2007)

JigsawPuzzle — administered on the web, taken from the NferNelson Spatial Reasoning series of books (NFER, 2002, Tosto, Hanscombe, Haworth, Davis, Petrill, Dale, Malykh, Plomin & Kovas, 2014).

Hidden Shapes — administered on the web, taken from the NferNelson Spatial Reasoning series of books (NFER, 2002, Tosto, et al., 2014).

SpatialComposite — based on the Hidden Shapes and Jigsaw Puzzle items.

General Cognitive Ability 'g' — based on the scores of 4 tasks administered on the web: General Knowledge (Kaplan, Fein, Kramer, Delis, & Morris, 1999), Vocabulary (Kaplan, et al., 1999), Ravens (Raven, Court, & Raven, 1996), and Picture Completion (Wechsler, 1992) tests. For more information on the tests see Haworth, et al., 2007).

Age 16 phenotypes

Mathematics Composite — created from Problem Verification and Understanding Number Tasks (Murphy & Mazzocco 2008; nferNelson, 1994).

Number Sense — administered on the web, assessed the ability to discriminate numerosities (Halberda & Feigenson, 2008).

Number Line — administered on the web, assessed estimation of numerical magnitudes and was programmed and implemented online from a description obtained from (Opfer & Siegler, 2007).

Ravens Progressive Matrices test — administered on the web, a computerized test of non-verbal (fluid) intelligence, adapted from (Raven, Raven, & Court, 1998).

General Cognitive Ability 'g' — administered on the web, a composite of the Vocabulary (Kaplan, et al. 1999) and Ravens scores (Raven, Court, & Raven, 1996).

Corsi Block — administered on the web, a test of visuo-spatial short-term working memory (Pagulayan, Busch, Medina, Bartok, & Krikorian, 2006).

Appendix 2

Gene-based analysis results for the William syndrome (7q11.22) and Prader-Willi areas (15q11-13). Nominally significant genes.

Phenotype	Chr	Gene ID	nSNPs	Start	Stop	p-value
Mathematics Test (age 12)	Chr7	POMZP3	48	76189302	76306620	0.001
	Chr7	LOC100133091	92	76128657	76307299	0.004
	Chr7	POM121C	35	74996059	75165565	0.005
	Chr7	SPDYE5	46	75074298	75183628	0.006
	Chr7	PMS2P3	44	75087068	75207453	0.006
	Chr7	WBSCR28	109	73225488	73330223	0.012
	Chr7	CLDN4	116	73195192	73297023	0.015
	Chr7	WBSCR27	111	73198920	73306855	0.015
	Chr7	UPK3B	44	76089739	76207199	0.034
	Chr7	PMS2P9	10	76618796	76732355	0.045
	Chr7	NSUN5P1	18	74989604	75096071	0.045
Mathematics Composite (age 12)	Chr7	POMZP3	48	76189302	76306620	0.001
	Chr7	POM121C	35	74996059	75165565	0.003
	Chr7	LOC100133091	92	76128657	76307299	0.010
	Chr7	SPDYE5	46	75074298	75183628	0.010
	Chr7	WBSCR28	109	73225488	73330223	0.012
	Chr7	NSUN5P1	18	74989604	75096071	0.013
	Chr7	PMS2P3	44	75087068	75207453	0.014
	Chr7	WBSCR27	111	73198920	73306855	0.014
	Chr7	CLDN4	116	73195192	73297023	0.014
	Chr7	TRIM74	17	74974902	75084888	0.017
	Chr7	TRIM73	17	74974902	75084896	0.017
	Chr7	TPST1	409	65620258	65875438	0.035

Mathematical Achievement (age 12)	Chr7	TYW1	577	66411791	66754507	0.024
	Chr7	AUTS2	948	69013904	70308054	0.033
	Chr7	MIR4650-2_1	405	66529308	66629384	0.033
	Chr7	MIR4650-1_1	405	66529308	66629384	0.034
	Chr7	CALN1	570	71194475	71927360	0.035
Jigsaw Puzzle (age 12)	Chr7	LINC00174	210	65791030	65915395	0.020
	Chr7	CALN1	570	71194475	71927360	0.020
	Chr7	RABGEF1_2	427	66155642	66326448	0.022
	Chr7	LOC493754	424	65945291	66107394	0.023
	Chr7	RABGEF1_1	595	66097077	66326448	0.024
	Chr7	LOC100996437	332	66069505	66184589	0.029
	Chr7	SRCRB4D	117	75968645	76089012	0.031
	Chr7	SBDS	145	66402689	66510588	0.034
	Chr7	POMZP3	48	76189302	76306620	0.036
	Chr7	GTF2IRD1P1	282	66224979	66359813	0.036
	Chr7	TPST1	409	65620258	65875438	0.038
	Chr7	KCTD7	332	66043867	66158216	0.039
	Chr7	LOC100507468	147	69011123	69112481	0.042
	Chr7	GNAI1	267	79714139	79898725	0.044
Spatial Composite (age 12)	Chr7	LOC100507468	147	69011123	69112481	0.021
	Chr7	LINC00174	210	65791030	65915395	0.028
	Chr7	LOC493754	424	65945291	66107394	0.034
g (age 12)	Chr7	POMZP3	48	76189302	76306620	0.017
	Chr7	WBSCR28	109	73225488	73330223	0.026
	Chr7	SPDYE5	46	75074298	75183628	0.033
	Chr7	POM121C	35	74996059	75165565	0.036
	Chr7	CLDN4	116	73195192	73297023	0.037
	Chr7	WBSCR27	111	73198920	73306855	0.039
	Chr7	PMS2P3	44	75087068	75207453	0.046

Number Line (age 16)	Chr7	TRIM50	76	72676531	72792085	0.008
	Chr7	NSUN5	47	72666512	72772864	0.008
	Chr7	FKBP6	85	72692154	72822646	0.008
	Chr7	MIR3914-1	311	70722657	70822756	0.012
	Chr7	MIR3914-2	311	70722659	70822754	0.013
Corsi Block (age 16)	Chr7	MDH2	211	75627336	75746827	0.003
	Chr7	STYXL1	245	75575654	75727321	0.000
	Chr7	SRRM3	222	75781210	75966609	0.002
	Chr7	PMS2P9	10	76618796	76732355	0.023
	Chr7	TMEM120A	137	75566154	75673992	0.000
	Chr7	MIR3914-2	311	70722659	70822754	0.000
	Chr7	MIR3914-1	311	70722657	70822756	0.000
	Chr7	SNORA14A	71	75523100	75623234	0.000
Ravens (age 16)	Chr7	LOC101929736	344	66750962	66855012	0.002
	Chr7	MAGI2-AS3	287	79032272	79150524	0.009
	Chr7	STAG3L4	353	66717624	66836513	0.011
	Chr7	FKBP6	85	72692154	72822646	0.035
	Chr7	TRIM50	76	72676531	72792085	0.037
	Chr7	PMS2P4	354	66691117	66817429	0.039
	Chr7	NSUN5	47	72666512	72772864	0.044
	Chr7	LOC101927243	227	76994514	77104760	0.047
g (age 16)	Chr7	LOC101929736	344	66750962	66855012	0.010
	Chr7	MAGI2-AS3	287	79032272	79150524	0.017
	Chr7	LOC101927269	101	79950832	80053755	0.027
	Chr7	STAG3L4	353	66717624	66836513	0.045
	Chr7	PMS2P9	10	76618796	76732355	0.053
Number sense (age 16)	Chr7					n.s.

Mathematics Composite (age 16)	Chr7	MAGI2-AS3	287	79032272	79150524	0.001
	Chr7	PMS2P9	10	76618796	76732355	0.003
	Chr7	MIR3914-1	311	70722657	70822756	0.003
	Chr7	MIR3914-2	311	70722659	70822754	0.003
Mathematics Test (age 12)	Chr15	OR4N3P	3	22363461	22464395	0.049
Mathematics Composite (age 12)	Chr15					n.s.
Mathematical Achievement (age 12)	Chr15	CYFIP1	131	22892648	23003603	0.033
Jigsaw Puzzle (age 12)	Chr15	GABRA5	132	27061865	27244357	0.015
Spatial composite (age12)	Chr15	TJP1	256	29942356	30164706	0.016
	Chr15	GOLGA8I	20	23205241	23312743	0.025
	Chr15	WHAMMP3	23	23137728	23258357	0.026
	Chr15	GABRA5	132	27061865	27244357	0.027
	Chr15	ATP10A	741	25873859	26158349	0.030
g (age 12)	Chr15	NIPA1	186	22993278	23136843	0.015
	Chr15	NIPA2	222	22954683	23084427	0.018
	Chr15	ATP10A	741	25873859	26158349	0.021
	Chr15	CYFIP1	349	22842648	23053603	0.041
Number Line (age 16)	Chr15	TUBGCP5	165	22783394	22923891	0.014
	Chr15	SNRPN	443	25018793	25273729	0.034
Corsi Block (age 16)	Chr15	GOLGA8N	11	32835656	32949511	0.007
	Chr15	ARHGAP11A	63	32857344	32982150	0.019
	Chr15	OTUD7A	556	31725328	31997542	0.047
Ravens (age 16)	Chr15	GOLGA8I	20	23205241	23312743	7.80E-05
	Chr15	WHAMMP3	23	23137728	23258357	8.30E-05
	Chr15	CHRNA7	311	32272685	32512384	0.019
	Chr15	ATP10A	741	25873859	26158349	0.252

g (age 16)	Chr15	GOLGA8I	20	23205241	23312743	4.00E-05
	Chr15	WHAMMP3	23	23137728	23258357	4.70E-05
	Chr15	CHRNA7	311	32272685	32512384	0.018
	Chr15	MIR211	209	31307234	31407344	0.024
Number Sense (age 16)	Chr15					n.s.
Mathematics Composite (age 16)	Chr15	TUBGCP5	165	22783394	22923891	6.80E-05
	Chr15	CYFIP1	349	22842648	23053603	0.001
	Chr15	GOLGA8I	20	23205241	23312743	0.021
	Chr15	WHAMMP3	23	23137728	23258357	0.022
	Chr15	FMN1_1	710	33007744	33536934	0.043
	Chr15	FMN1_2	359	33391984	33536934	0.049

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

Russian Psychological Society
Lomonosov Moscow State University

Elaboration of Screening Scales for Early Diagnosis of Developmental Delay in Four- to Five-Year-Old Children in Russia

Andrey D. Nasledov, Sergey A. Miroshnikov*, Liubov O. Tkacheva

St. Petersburg State University, St. Petersburg, Russia

* Corresponding author. E-mail: sergeyamir@gmail.com

Background. While various screening systems are used worldwide for early detection of developmental delay (DD), Russia still does not have such a screening system in place, even though a good prognosis for the cognitive development of a child with DD depend strongly on the time of the diagnosis.

Objective. The objective of this study was to create a system to rapidly monitor the mental development of four- to five-year-old Russian children, a system that allows for the use of modern information technologies to obtain reliable results.

Design. This study was carried out with a sample of 1,232 children. For data collection, the multifactor study of mental development tool was used as a part of a software complex for longitudinal research. This tool included a much more extensive set of tasks than in traditional tests of abilities, allowing for a wider variation of the factor structure. For the 4-year-olds, 236 tasks were used and 349 for the 5-year-olds. Factor and discriminant analysis were carried out to construct scales for each age group (6–7 points in each), which most accurately predict the diagnosis (Norm/DD). Structural equation modeling (SEM) was used to verify the prediction model.

Results. Two scales were elaborated, which coincided with the type of variables combined in each of them regardless of age (for 4-years — simpler and for 5-years — more complex): logical reasoning, motor skills, and general awareness (listed in descending order of contribution to the prediction). SEM confirmed that the selected scales are indicators of the general ability factor, which is the main predictor of the diagnosis.

Conclusions. Two short scales for the rapid diagnosis of DD in preschool children were constructed, allowing the use of computer technology to timely identify the risk group among 4- and 5-year-olds with high sensitivity and specificity of the forecast (not lower than 95%).

Keywords: Screening, developmental delay, construction of scales, SEM.

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

doi: 10.11621/pir.2018.0411
http://psychologyinrussia.com

Introduction

Traditionally, the diagnostic tools used in developmental psychology had to meet high psychometric requirements. However, presently, the efficiency and flexibility of their application are becoming more and more relevant, especially when the tools are used in the diagnosis of developmental problems in preschool children. Efficiency is understood to be the possibility of fast screening in large groups to identify children whose development is slightly distorted, but this needs to be addressed even before there are obvious signs of delays.

Traditional tests that are designed for manual execution and processing cannot meet these requirements. The main reason of such imperfection is that traditional methods sacrifice the accuracy of the results when collecting data using forms and manual processing of the results: the important details of the tasks are missed (for example, the number of levels of the elements of a complex task are not taken into consideration). Nevertheless, in Russia, there is even a lack of traditional methods for diagnosing preschool children, which correspond to the standard psychometric requirements for diagnostic tools (reliability and validity, the presence of reasonable norms, legality of distribution). These diagnostic obstacles lead to belated and ineffective attempts to correct a variety of developmental problems, among which the most common are developmental delays (DD) in various forms. Developmental delay can be defined as a significant delay in two or more developmental domains: gross and fine motor skills; speech and language; cognition; personal and social development; or activities of daily living (Majnemer, Shevell, 1995). Recently the number of children with DD in Russia has been increasing (Isaev, 2017), mainly due to medical advances and assistance for premature and somatically weakened children. It is also necessary to note that children who are initially diagnosed as at risk are always in need of further special support, otherwise the positive effect of early medical care may be lost (Collins et al., 2017). It is well known that diagnosis at an early age is significant for intervention: the younger the child, the more flexible his/her brain and the richer the compensatory possibilities (Dennis, 2013). DD has a high potential for rehabilitation, but when only when timely diagnosed and corrected (Barnett, 1995): wasted time significantly increases the probability of an adverse outcome (Kok, 2016).

It is customary to distinguish the many factors and markers that influence mental development at an early age. Among them are motor development (Frick & Mohring, 2013), speech development (Tsao, Liu & Kuhl , 2004), working memory (Cowan & Alloway, 2009), sustained attention (Voelke & Roebers, 2016), spatial thinking (Lauer & Lourenco, 2016), cognitive control mechanisms (Engel de Abreu, Conway & Gathercole, 2010), and social intellect (Wellman et al., 2008. However, the role of these factors as predictors of DD at a particular age is still unclear. While various screenings are used in the West, such as Griffiths (Luiz, 2006) and Vineland (Van Duijn, 2009), there is no such system for monitoring the mental development of children in Russia. Presumably, it could be possible to adapt one of the existing Western screenings to the Russian sample; however, this approach seems to be doubtful, given the complexity of the adaptation itself, which is comparable to the development of new tools, and the legal and financial problems of licensing. According to Carroll A. E., the application of computerized cognitive development

screenings to detect DD significantly increased the number of diagnosed children and the timely intervention for children at an early age (Carroll, 2014). *Therefore, we aimed to create a system for screening diagnostics of mental development in 4- and 5-year-olds, specific to the Russian sample and allowing for the usage of technologies to obtain reliable results on the basis of a relatively small set of data.*

Most studies of DD are based on the classical ideas about the factor structure of intelligence that is specified a priori in tests. However, the usage of structured tests with several types of tasks forming the subtests and factors significantly limits the possibility of studying the real factor structure of abilities, partly due to the questionnaire testing tradition and the manual processing of its results (Macmann & Barnett, 1994). For the purpose of our study, a broader set of initial characteristics was chosen, which allowed for a looser grouping of the factors. We also carefully recorded the results of the child's performance on each task and took into account the performance success of each level of the task and the age of the child (up to one day), etc.

Method

Materials

For data collection, we used the multifactor study of the mental development tool as part of a software complex for longitudinal research (Автор, 2001). It is a complex tool made up of formal observations and tests, which includes 502 tasks covering a wide range of abilities (from motor skills to cognitive and social development). The complexity of the tasks was chosen depending on the calendar age of the child (from 2 months to 7 years). The bank of tasks was formed on the basis of a survey of a large number of expert practitioner psychologists (the content of tasks is presented on the Internet resource of the project http://info11.testpsy.net). Most of the tasks used a categorical (dichotomous) response format: Yes = 2 (the child can perform a control action) or No = 1 (cannot). The baseline data included 847 dichotomous points: 236 points for 4-year-olds and 349 points for 5-year-olds were used in this study. Examples of the tasks are presented in Table 1.

Procedure

The diagnosis of DD was established outside the scope of this study by other experts, representatives of advisory centers and commissions with the participation of neurologists, pathologists, and psychiatrists (1 — Norm, 2 — DD). Data collection was conducted by psychologists involved in psychological and pedagogical support of children in regular and specialized preschool educational institutions (Saint-Petersburg, Murmansk, Belgorod, and others in the period from 2015 to 2017). The diagnostics was performed as a part of routine testing of children after receiving written parental consents.

The sample

In total, 628 4-year-old (575 Norm, 53 DD) and 604 5-year-old children (527 Norm, 77 DD) were surveyed.

Statistical data analysis

The 236 variables (points) for 4-year-olds and 349 variables for 5-year-olds were processed separately for children of each age (4- and 5-year-olds) to: a) identify a set of scales most accurately predicting the diagnosis (belonging to the norm/DD group), which have a simple factor structure and sufficient reliability (according to Cronbach's alpha); b) explain the relationships between predictors and the relative contribution of the scales to the prediction of the diagnosis; and c) develop an algorithm for the rapid assessment of the risk of DD based on the results of the screening application. The analysis was carried out separately, but in the same sequence, for the sample of 4- and then 5-year-olds. All statistical analysis was performed using IBM SPSS software and AMOS version 25.

Results

Selection of variables

A discriminant analysis (DA) of the variables with a stepwise method was conducted to select a set of variables that would best predict group allocation (Norm, DD). More detailed statistical analysis is presented in the article (Автор, 2018). The standard procedure for discriminant analysis was used (IBM SPSS Statistics 25), stepwise method, probability of F: entry 0.05; removal 0.1.; grouping variables: diagnosis (Norm, DD), independent: 236 variables for 4-year-olds, 349 for 5-year-olds. The result: list of variables included in the analysis at the last step of stepwise method application (56 variables for 4-year-olds and 52 for 5-year-olds).

The formation of scales

Using factor analysis (FA) and, thereafter, DA on the calculated factors, variables were selected according to the following requirements: each of the remaining variables was included in only one factor with a load of at least 0.4; each factor included at least five of these variables; calculated factors together provided the maximum accuracy of the diagnosis prediction (DA); the variables included in each factor formed a fairly reliable scale for internal consistency (Cronbach's alpha); the factor validity of the scales was verified by confirmatory factor analysis (AMOS).

For each age (4 and 5 years), a set of three factors, satisfying all the requirements, was obtained: 20 points for 4-year-olds, 19 points for 5-year-olds (6 to 7 points for each factor). The factors coincided by the type of variables united in each of them regardless of age (for 4-year-olds — simpler, for 5-year-olds — more complex): 1) "general awareness" (the breadth of the child's knowledge of the world), seven items for 4-year-olds, six items for 5-year-olds; 2) "motor skills" (sufficient development of gross- and fine-differentiated motor skills), seven items for 4-year-olds, six items for 5-year-olds; 3) "logical reasoning" (the ability to make a logical conclusion based on the comparison operation and consideration of the conditions of the task), six items for 4-year-olds, seven items for 5-year-olds. Examples of the tasks (items) are presented in Table 1.

Table 1

Task examples (items)

Age	Task
Scale (Factor) General awareness	
4-year-old	Can give the names of wild animals (3–5).
	Can give the names of some cities (1–2).
5-year-old	Can navigate in a sequence of days of the week.
	Can answer the question: "How old will you be in one year?"
Scale (Factor) Motor skills	
4-year-old	Can draw a person.
	Can draw a triangle by looking at the example.
5-year-old	Can cut out paper with scissors on a simple contour.
	Can navigate in a sheet of checkered paper, performing tasks according to the instructions.
Scale (Factor) Logical reasoning	
4-year-old	Can distinguish objects by height.
	Can distinguish between true and false depicted in pictures.
5-year-old	Can explain what the car needs brakes for.
	Can answer the question: "What do an axe and a hammer have in common?"

Verification of reliability, relative contribution of scales and prognosis accuracy

To verify the stability of predictions with a given set of predictors, the whole sample was divided into two age groups according to the median of age (in days) of the DD groups: younger and older (median for 4-year-olds was 1,677 calendar days, median for 5-year-olds was 2,041 calendar days). The reliability of the scales (Cronbach's alpha) was separately defined for 4- and 5-year-olds and for parallel subgroups of the younger and older groups for each age. A sufficiently high reliability of each scale (from 0.779 to 0.922) was confirmed in the parallel subgroups. The scales' values were calculated as the sum of their variables (for each age specifically). Checking the reliability of the scales (Cronbach's alpha) and the subsequent verification of the accuracy of the forecast (discriminant analysis) were carried out six times, three times for each age group (4 - and 5-year-olds): for the entire sample, for younger, and for older groups. In all six cases, the factors "logical reasoning," "motor skills," and "general awareness" made the biggest contribution to the accuracy of the diagnosis prediction (Norm/DD) in descending order: the higher their value, the stronger the probability of belonging to the norm. The sensitivity (accuracy of prediction of DD) and specificity (accuracy of prediction of norm) of the prognosis for 4-year-olds were not lower than 96.2%, and for 5-year-olds, a sensitivity not lower than 89.5%, and a specificity not lower than 97.5%. The sequence of analysis and the results for 4-year-olds are presented in more details in the article (Автор, 2018).

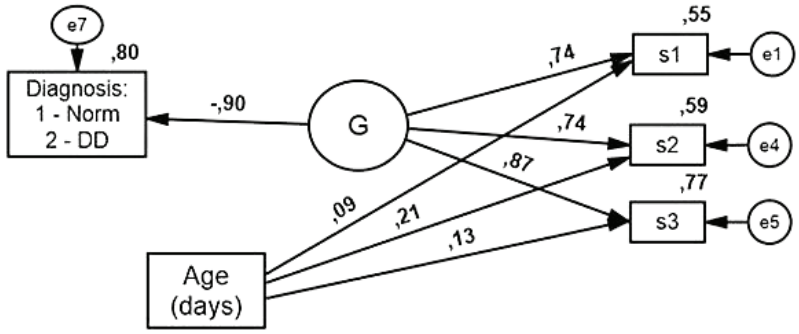


Figure 1. Structural diagnosis prediction model.

Note. Numbers at arrows — standardized regression coefficients; numbers at contours of variables — squares of multiple correlation; S1 — "general awareness", S2 — "motor skills", S3 — "logical reasoning."

Structural equation modeling (SEM) was used to test the following assumptions: the selected scales are indicators of the general factor (G-factor), which is the main predictor of the diagnosis; age has an indirect impact on the diagnosis: the mediator of this effect is the G-factor. The verification was also carried six times on the samples of each age and their subgroups (younger and older).

All six models confirmed the initial assumptions on the indexes of agreement and statistical significance of parameters (Byrne, 2010); example one model is shown in Figure 1 (for 4-year-olds). The models differed only in the statistical significance of the influence of age on the separate scales: this effect was stronger for 4-year-olds, but weaker for 5-year-olds, until the loss of statistical significance for the subgroup of older 5-year children. The fit indices for the models are presented in Table 2.

Table 2

Fit indices for models

Sample	CMIN	df	p	GFI	CFI	RMSEA	Pclose
All, 4-years old	5.500	3	.139	.995	.978	.036	.611
Age<1677 days	1.863	2	.394	.997	1.000	.000	.651
Age≥1677 days	1.439	4	.837	.996	1.000	.000	.944
All, 5-years old	3.467	3	.325	.998	1.000	.016	.772
Age<2041 days	4.135	4	.388	.995	1.000	.010	.731
Age≥2041 days	6.450	5	.265	.990	.998	.034	.576

All models are almost identical and correspond well to the original data. The model for younger 4-year-olds additionally included a negative relationship between errors e1 and e4 ($\beta = -.21$), and in the model for older 4-year–olds, the relationship age — >s1 ($p > .1$) is not statistically significant. In the model for all

5-year–olds, the relationship age — > S3 ("logical reasoning") is not statistically reliable. For younger 5-year-olds, the influence of age on S2 ("motor skills") loses its statistical significance, and in the model for older 5-year-olds, age does not have a statistically significant effect on the other variables of the model.

Development of test scales was aimed to divide groups into Norm and DD with maximum accuracy. The scales were constructed separately for 4 - and 5-year-olds. For this purpose, the results of DA were used to determine the discriminant function (DF) — an axis passing through the centroids of the separable classes (Norm, DD) and to estimate the probability of belonging to DD for each discriminant score (DS) of this function (Klecka, 1980). As a result, nonstandardized DF coefficients (separately for 4-and 5-year-olds), included in the linear equation for calculating DS by predictor values (age in days, S1, S2, S3), were extracted:

$$DS_i = A - B_1 \times Age_i + B2 \times \mathrm{S1}_i + B3 \times \mathrm{S2}_i + B4 \times \mathrm{S3}_i$$

where: A — intercept; $B1$ — $B4$ — DF coefficients; i — number of the child; DS_i — its discriminant assessment; $S1_i$, $S2_i$, $S3_i$ — values of the corresponding scales for the child.

The calculated DS for the entire sample for each age group were the raw scores to be scaled. Both DS distributions differ significantly from the normal distribution with pronounced right-side asymmetry and a long tail toward negative values. Therefore, different variants of the nonlinear transformation of the scales were compared to develop norms. The 50-point percentile scales appeared to be the most accurate, in the construction of which the entire range of raw scores (DS) was divided into 50 intervals (2% of the sample per interval), and the upper limits of DS for each interval were calculated.

The developed scales demonstrate high accuracy of the prognosis. For 4-year-olds, the upper limit of the 12th percentile (P12) at 100% sensitivity (accuracy of prediction of DD) provided 96% specificity (accuracy of prediction of norm). And for 5-year-olds, the upper limit of P16 corresponded to 94.8% sensitivity and 95.4% specificity. It is believed that good screening should have a specificity of 70% to 80%, and sensitivity in the range of 70% to 80% is considered acceptable (Glascoe, 2005). In most existing screenings, sensitivity is 82% and specificity is 78% (Tonelly, 2016), that is, much lower than the expected accuracy of the scales developed in our study.

Discussion

The constructed scales differ significantly from standard ability tests because their contents are not determined by individual differences (development vectors) of normally developing children. The procedure of elimination of the scales determined contents: it was a vector that polarized to the best extent children at risk for DD from their peers without developmental problems.

The greatest difference between groups norm and DD was obtained on the following factors, specified in descending order of contribution: "logical reasoning," "motor skills," and "general awareness." The factor "logical reasoning" was the most powerful predictor of DD. However, it should be emphasized that the prognostic ability of indicators is valid only in their totality.

Thus, the orientation of the scales is generally consistent with previous studies. For example, it was shown that the level of primary understanding of the concept of numbers in early childhood affects the development of logical reasoning in preschool age children (Clements, Sarama, 2007; Hollister Sandberg, 2010). Also, it has been shown that 8- to 12-year-olds with borderline intellectual functioning have a relatively low capacity for logical reasoning capacity and for autonomy and social responsibility in everyday activity (Baglio et al., 2016). The role of logical reasoning as a predictor of development for 4- to 5-year-old children is extremely important for Russia — in 67% of cases initially diagnosed children with DD in Russia have a speech delay (Isaev, 2017), which naturally inhibits the formation of logical reasoning. The importance of motor development at different ages of early childhood and its impact on further cognitive development has been emphasized in many studies (Farber, Beteleva, 2005; Dornelas, et al., 2016; Hernandez, Cacola, 2015; Baglio et al., 2014; Cheng et al., 2014; Mithyantha et al., 2017). It is known that at the age of 4, a typically developed child still maintains a close interaction between visual perception and motor activity, inherent at an earlier age; at this age, manual manipulations with objects are considered as an essential factor of visual identification (Farber, Beteleva, 2005). Likewise, in a longitudinal study, it was shown that a delay in motor development in 2-year-old children as the most significant marker of DD, increases the likelihood of maintaining this diagnosis during the school years (Dornelas, et al., 2016). While in another longitudinal study, it was shown that a high level of motor development during infancy determines the level of cognitive development in 4-year-olds (Hernandez, Cacola, 2015). According to the results of an MRI study of 7- to 8-year-old children with DD in comparison with the norm, the most significant differences were recorded in parts of the brain associated with motor skills, perception, and the regulation of behavior (Baglio et al., 2014). General awareness is considered as the width of a child's knowledge of the world; therefore, it is often used as a predictive marker of DD (Cheng et al., 2014; Mithyantha et al., 2017). But it should be noted that general awareness can be trained to a much greater extent than the nonverbal abilities of the child, as was demonstrated in the study of the training for the Bright Start program for young children with developmental language delays (Tzuriel et al., 2017).

The peculiarity of the scales is reflected not only in the process of their elaboration, but also in their effective application, which involves the use of computers, as focused on the fullest access to the obtained information (data on individual specifics of the tasks performance, the levels and complexities of tasks, etc., and the calendar age of a child up to the day). An attainment of the final result on the scales requires more complex calculations than is provided for in ordinary manual tests, but the widespread usage of computers by modern specialists will welcome the application of these types of scales in scientific studies and for practical assesments. Moreover, the use of computerized diagnostics optimizes the work of a specialist at all stages of the examination: an automation of the selection of the tasks and the stimulus, depending on the age of a child, and the preservation and consolidation of data processing with output, not just the numeric results, but also prior interpretations and possible recommendations.

Conclusion

The main result of this study is the elaboration of short scales for rapid (30 minutes) diagnosis (Norm/DD), allowing practitioners to quickly identify the risk group among 4- and 5-year-old children with a very high accuracy: at least 95%. It exceeds far more than existing standards for screening tests (Glascoe, 2005; Tonelly, 2016). Currently, this scale can be used on the basis of data collected by the longitud program complex (Автор, 2001). In the future, as the scales expand to other age ranges, independent software will be created that is more concise in terms of data collection and is designed only for use in screening mode to identify children at high risk of DD. However, more detailed individual psychodiagnostics will be needed to clarify the diagnosis and classify a kind of developmental delay to a certain nosological group and to determine the necessary corrective measures. Thus, the expected effect of the application of the screening is an early detection of the risk of DD to change the potential trajectory of a child's development. The desired outcome of the screening is to optimize the work of psychological services, reducing the number of children requiring a detailed individual diagnosis, which takes a lot longer than screening..

Limitations

The scales are designed to identify a high risk of delayed development in a timely manner, when the lag has not yet become apparent and an accurate nosological classification is still hindered. In this respect, the obtained results in the primary examination (screening) by these scales should be considered as a basis for refining the child's educational route and planning further observations but not as a diagnosis. A specialist examination is always necessary for the final conclusion about the presence of DD and its particular type, regardless of the severity of the signs revealed during the first computerized assesment.

These scales are not intended for the examination of children with behavioral or other specific disorders that may affect the results of observations and the successfulness of the task performance.

Acknowledgements

The reported study was funded by the Russian Foundation for Fundamental Researches, grant № 17-06-00989-OGN "Longitudinal study of the predictive patterns of delayed psychomotor development of preschool children."

References

Baglio, F., et al. (2014). Abnormal development of sensory-motor, visual temporal and parahippocampal cortex in children with learning disabilities and borderline intellectual functioning. *Frontiers in Human Neuroscience, 8*(806*)*. https://doi.org/10.3389/fnhum.2014.00806

Baglio, G., et al. (2016). Social competence in children with borderline intellectual functioning: delayed development of theory of mind across all complexity levels. *Frontiers in Psychology, 7*, 1604. https://doi.org/10.3389/fpsyg.2016.01604

Barnett, S.W. (1995) Long-term effects of early childhood programs on cognitive and school outcomes. *Future Child*, 5, 25–50. https://doi.org/10.2307/1602366

Byrne, B. M. (2010) *Structural equation modeling with AMOS: Basic concepts, applications and programming.* 2nd ed. (Multivariate applications series). New York: Taylor & Francis Group.

Carroll, A.E. et al. (2014) Use of a Computerized Decision Aid for Developmental Surveillance and Screening: A Randomized Clinical Trial. *JAMA Pediatr. 168*(9), 815-821. https://doi.org/10.1001/jamapediatrics.2014.464

Cheng, E.R. (2014) Cognitive Delay and Behavior Problems Prior to School Age. *Pediatrics*, *134*(3), 749–757. https://doi.org/10.1542/peds.2014-0259

Clements, D.H., Sarama, S. (2007) *Early childhood mathematics learning. Second handbook of research on mathematics teaching and learning.* US: Information Age Publishing.

Collins, P.Y. et al. (2017). Global services and support for children with developmental delays and disabilities: Bridging research and policy gaps. *PLoS Med, 14*(9), e1002393. https://doi.org/10.1371/journal.pmed.1002393

Cowan, N., & Alloway, T. (2009). Development of working memory in childhood. In M.L. Courage, & N. Cowan (Eds.), *The development of memory in infancy and childhood.* Hove, East Sussex, UK: Psychology Press, 303–342.

Dennis, M., et al. (2013) Age, plasticity, and homeostasis in childhood brain disorders. *Neuroscience & Biobehavioral Reviews*, 37, 2760–2773. https://doi.org/10.1016/j.neubiorev.2013.09.010

Dornelas, L.F., Duarte, N.M., Morales, N.M., et al. (2016) Functional Outcome of School Children With History of Global Developmental Delay. *Journal of Child Neurology*, *31*(8), 1041-1051. https://doi.org/10.1177/0883073816636224

Engel de Abreu, P., Conway, A., & Gathercole, S. (2010). Working memory and fluid intelligence in young children. *Intelligence, 38*, 552–561. https://doi.org/10.1016/j.intell.2010.07.003

Farber, D.A., & Beteleva, T.G. (2005) Formation of the system of visual perception in ontogeny. *Human Physiology*, 31, 515–524. https://doi.org/10.1007/s10747-005-0091-3

Frick, A., & Mohring, W. (2013). Mental object rotation and motor development in 8- and 10-month-old infants. *Journal of Experimental Child Psychology, 115*, 708–720. https://doi.org/10.1016/j.jecp.2013.04.001

Glascoe, F.P. (2005) Standards for Screening Test Construction. Retrieved from http://www.aap.org/sections/dbpeds/pdf/Standards%20for%20Screening%20Test%20 Construction.pdf.

Hernandez, A.M., Cacola, P. (2015) Motor proficiency predicts cognitive ability in four-year-olds. *European Early Childhood Education Research Journal*, *23*(4), 573–584. https://doi.org/10.1080/1350293X.2014.991094

Hollister Sandberg, E., McCullough, M. B. (2010) The Development of Reasoning Skills. In E. Hollister Sandberg & B.L. Spritz (Eds.), *A Clinician's Guide to Normal Cognitive Development in Childhood*. New York: Routledge. 179–198.

Isaev, D.N. (2017). *Psihiatrija detskogo vozrasta. Psihopatologija razvitija.* [Psychiatry of childhood: psychopathology of development.] Russia, Saint-Petersburg: Spetslit.

Ivanova, A.E., Miroshnikov, S.A. (2001). *Metodicheskie materialy k ekspertnoi sisteme individual'nogo soprovozhdeniya «Longityud»* [Methodical Materials for the Expert System of Individual Tracking «Longitude»]. St. Petersburg: S.-Peterb. gos. universitet.

Klecka, W.R. (1980). *Discriminant Analysis*. Beverly Hills, California: Sage Publications https://doi.org/10.4135/9781412983938

Kok, L., et al. (2016) The effectiveness of psychosocial interventions for children with a psychiatric disorder and mild intellectual disability to borderline intellectual functioning: A systematic literature review and meta-analysis. *Clinical Child Psychology and Psychiatry*, 21, 156–171. https://doi.org/10.1177/1359104514567579

Lauer, J.E., & Lourenco, S.F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychological Science, 27*, 1291–1298. https://doi.org/10.1177/0956797616655977

Luiz D., et al. (2006) *Griffiths Mental Development Scales, Extended Revised: 2 to 8 years.* Association for Research in Infant and Child Development (ARICD), Hogrefe.

Macmann, G.M., & Barnett, D.W. (1994) Some additional lessons from the wechsler scales: A rejoinder to Kaufman and Keith. *School Psychology Quarterly*, 9, 223–236. https://doi.org/10.1037/h0088288

Majnemer, A., & Shevell, M.I. (1995) Diagnostic yield of the neurologic assessment of the developmentally delayed child. *J Pediatr,* 127, 193–199. https://doi.org/10.1016/S0022-3476(95)70294-6

Mithyantha, R., et al. (2017) Current evidence-based recommendations on investigating children with global developmental delay. *Archives of Disease in Childhood*,102, 1071–1076. https://doi.org/10.1136/archdischild-2016-311271

Tonelly, M. et al. (2016) Recommendations on screening for developmental delay.CMAJ, 188, 8, 579–587. https://doi.org/10.1503/cmaj.151437

Tsao, F.M., Liu, H.M., & Kuhl, P.K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development, 75*, 1067–1084. https://doi.org/10.1111/j.1467-8624.2004.00726.x

Tzuriel, D., et al. (2017) Effects of teaching classification on classification, verbal conceptualization, and analogical reasoning in children with developmental language delays. *Journal of cognitive education and psychology*, 16, 107–124. https://doi.org/10.1891/1945-8959.16.1.107

Van Duijn, G., et al. (2009) Vineland Screener 0–12 years research version (NL). Constructing a screening instrument to assess adaptive behavior. *Int J Methods Psychiatr Res.*, *18*(2), 110–7. https://doi.org/10.1002/mpr.282

Voelke, A., & Roebers, C. (2016). Sustained attention and its relationship to fluid intelligence and working memory in children. *Journal of Educational and Developmental Psychology, 6*, 131–139. https://doi.org/10.5539/jedp.v6n1p131

Wellman, H.M., et al. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology, 44*, 618–623. https://doi.org/10.1037/0012-1649.44.2.618

Original manuscript received May 21, 2018
Revised manuscript accepted September 1, 2018
First published online December 30, 2018

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

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

Russian Psychological Society
Lomonosov Moscow State University

PERSONALITY PSYCHOLOGY

Polar Meaning-Building Strategies: Acmeological Characteristics

Irina V. Abakumova[a], Pavel N. Ermakov[a], Mikhail V. Godunov[b*]

[a] *Southern Regional Scientific Center of the Russian Academy of Education, Rostov-on-Don, Russia*

[b] *Co-creation School, Rostov-on-Don, Russia*

[*] Corresponding author. E-mail: godunmv997@gmail.com

Background. Personality is not simply an end product, but rather, it is a process. Therefore, empirical work on personal meaning-building should examine the genesis of meaning and provide a content-based description of personality in terms of personality traits. Such a description suggests a systemic view of personality, where the meaning-based approach is supplemented with the definition of personality traits. The value and meaning potential of personality encompasses three dimensions: worldview, behavior, and cognition.

Objective. The aim of this study is to identify the properties of personality, reflecting the features of polar strategies of meaning formation in acmeological terms by age, gender, and professional characteristics.

Design. The present study considers the influence of various acmeological factors on meaning-building and concentrates on its two polar strategies: adaptive and developing strategies. We developed nine bipolar scales of personal traits with sublevels by applying the semantic differential technique. In total, there were 145 participants in the study. Participants were grouped according to three criteria: age, gender, and profession.

Results. The obtained indices of meaning-building strategies did not coincide in all the differentiated groups, which clearly speaks in favor of acmeological dynamics of the respondents' personal profiles. We stratified the sample according to the mean score of the basic marker of "life meaningfulness," which enabled us to establish differences in characteristics of actual polar strategies of meaning-building. The respondents who did not fall into either of the two groups are "between the poles." They often have an underdeveloped meaning-building strategy as a result of poorly formed ways of organization and actualization of personal meanings or the presence of a transitional form of situational conceptual initiations.

Conclusion. The personal profiles that were identified represent multifactor models of the personal value and meaning dimensions, which can predict actual meaning-build-

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

doi: 10.11621/pir.2018.0413
http://psychologyinrussia.com

ing strategies using semantic differential scales and indicators ("life meaningfulness" from the Purpose-in-Life test) and help researchers to reduce the number of techniques employed in their studies.

Keywords: personality, meaning, meaning-building strategy, development, adaptation, polarity, semantic scale.

Introduction

The aim of this study is to identify the properties of personality, reflecting the features of polar strategies of meaning formation in acmeological terms by age, gender, and professional characteristics. Personality development takes place in a constantly changing society. Continuous changes and life crises create various transformations in personal value and meaning sphere. The genesis of personal meanings depends on both external conditions and internal subjectivity. Various acmeological factors, showing the possibility of achieving the highest level of individual development, influence the individual meaning. They change the trajectories of meaning-building, the actualization of meaning in certain life circumstances, and transmission during various interactions (Abakumova, Ermakov, & Fomenko, 2013). Modern theories of meaning define meaning-building strategies as dynamic characteristics of personal value-motivational activity (D. Leontiev, 2007). It is the strategy of meaning-building that manifests in the system of meaning in consciousness, which entails the actualization of primary meanings and their dominance in specific interaction situations. Meaning-building strategies orient personal development trajectories in various periods of life and determine the stability of self-development and characteristics of self-identification.

Specific characteristics of primary meaning-building strategies depend on the presence and unity of three components: (a) an initiating principle, (b) the content, and (c) the direction of the development of meanings. The meaning-building strategy is a way of forming and developing the system of personal meanings, which is organized in terms of motives, needs, goals, experience, and subjective relations; it also reflects the specific characteristics and dynamics of individuals' actualization of meanings in specific life situations (Abakumova, Godunov, Enin, & Generdukaeva, 2016). Such a definition of meaning-building strategies can assist in our understanding of both possible and actual ways of organizing personal meaning dimensions, which contain various qualitative preferences and views. This offers new opportunities for studying features of meaning regulation during individuals' interactions in specific decision-making situations. The processes of estimation and choice reflect adherence to various meaning-building strategies. The present paper presents an empirical study of the acmeological characteristics of these strategies.

The issues of alternative estimation, contrast, and choice are relevant for the ontological framework of personal development. Besides, these issues represent a teleological tool for all branches of the humanities. This is where psychology considers and supplements its specific approach, which reflects the multidimensionality of meaning as a category. At the same time, the issues of polar estimation and contrast are associated with different levels of such phenomena. As applied to value-meaning categories, individuals often express polar points of view. How-

ever, many psychological constructs have a dichotomous continuum structure at the empirical level and in everyday life. In its most elementary form, this structure reflects the presence of two alternative poles (attractors) as sources for the development of the system. The simplest logical choice is the choice between two states as opposites on a dipole.

In his three-level model of personality, D. Leontiev (1997) demonstrated that personality structure contains personal meanings (core layer) and personal traits (surface layer), which receive their nourishment from the core layer and reflect its essence. This correlation between structural levels of personality may be considered as a correspondence between personal traits and meanings. In this case, there is a mutual correspondence between the system of personal meanings and personal traits in several of their associations. On one hand, this means that personal traits reflect only the meanings that exist in the value-meaning sphere and the world outlook core of personality. On the other hand, personal meanings influence the characteristics of individuals' interactions, as well as judgments and relationships in various situations, which manifests in certain personal traits as distinctive characteristics. Thus, personal meanings are prototypes, and personal traits are isomorphic images as derivatives of meanings. Consequently, the correspondence between personal meanings and personal traits reflects their integrity and identity, which can only be observed in mentally healthy individuals.

In other words, meanings represent personal internal structure. Their properties are external (surface contact layer) and manifest themselves in interactions. The mechanism of correspondence between personal meanings and personality traits consists in personal interactions and individuals' lives in their various manifestations. Meanings are instrumental in manifesting and actualizing personality traits in interactions. The presence of certain personality traits speaks in favor of actual meanings. Allport (2002) notes that personality traits are the driving elements of human behavior. For example, curiosity serves as a mechanism for achieving well-being and finding the meaning of life (Kashdan & Steger, 2007).

In his study of personality development, Lazursky (1924) concluded that it is impossible to create a unified personality typology. He suggested the classification of personal levels of maturity, which reflect individuals' development and activity in society. Lazursky's classification embodied the following components: (a) passive adaptation to reality, (b) active adaptation to reality, (c) individuals' adaptation of reality to themselves, and (d) realization of human ideals and creativity of "knights of spirit."

The foregoing theories illuminate two meaning-building strategies aimed at adaptation and development of meanings. We define these strategies based on their components (Abakumova et al., 2016) as follows:

1) *Adaptive strategy of meaning-building* refers to organizing the dimensions of meanings in a way compensates for defects in individual development through adjustment and monotonous movement in the layer of acquired personal meanings under the influence of the external environment dominating and determining individuals' vital activity. This strategy is based on formal and stereotyped goals.

2) *Developing strategy of meaning-building* refers to transforming the dimensions of meaning in a way that tends to build prospective meanings and transforms their content under the influence of external factors, which individuals assess as

living conditions that can be overcome. This strategy is focused on understanding the motives and generating actual goals, which are important for personal growth.

According to J. Piaget's operational concept of intellect, the process of adaptation of mental activity includes assimilation and accommodation mechanisms (Piaget, 1998). Thus, the suggested strategies for regulating the dimensions of personal meanings reflect the characteristics of cognitive development.

In one of his final interviews, Abraham Maslow stated that the satisfaction of basic needs does not always provide inevitable and unconditional self-actualization. He argued that when individuals achieve the level where their basic needs are met, some continue to move towards self-actualization, while others stop moving (Frick, 2000). The presence of such a phase meaning "crossing the Rubicon" suggests that dimensions of meaning may be organized in accordance with strategies reflecting the actualized personal preferences for either adaptation or development.

D. Leontiev examined the characteristics of life strategies in the form of two primary alternatives, where symbiotic survival as an escape from responsibility towards collective personality can be contrasted with a transcendent autonomy for true personal growth, when a person "needs the most" (D. Leontiev, 2002). Because of the above-stated dichotomies of meaning, individuals give priority to certain things in comprehension. In other words, they make a choice as a divergence in the process of meaning-building. The concepts of divergence and convergence reflect a variative branching into independent tendencies, converging from different sides to develop in a single direction. To study creative abilities and creativity, Guilford (1967) suggested the concept of divergent and convergent thinking, noting a fundamental difference between two types of mental operations: convergence and divergence. Convergent thinking is associated with finding correct solutions to various problems. Divergent thinking is multidirectional. This type of thinking helps individuals vary problem-solving methods and leads to unexpected conclusions and results (Druzhinin, 2009).

Such a dichotomy in the development of meaning dimensions allows us to conclude that the bipolarity of meanings manifests in their transfer to the external layer, where they meet other meaning systems. According to A. Leontiev, such a "decrystallization" transforms meanings into personal ones. However, meanings become personal based on their consonant or dissonant positioning. Such an interpretation of acceptance/rejection (coincidence/noncoincidence) in actual meaning dimensions provides a dyadic space along with bipolar meanings (meaning continuum).

When ordering complex descriptive personality traits, their linear arrangement may be useful. In its simplest form, such a sequential arrangement has one or two poles (attractors) as sources of the development of various systems including "the physical world, although its description contains bipolar characteristics (distant/close) and unipolar characteristics from zero to infinity. Closeness is the least degree of remoteness. However, we use bipolar characteristics when describing others" (Petrenko, 2013, p. 181).

Interpretations of others contain at least two points of view, which reflect their similarities and differences. This corresponds to Kelly's (1955) constructive alternative approach. Such a process of contrasting creates meaning constructs, which

correspond to alternative ways of perceiving the world and lifestyles. In general, meaning constructs are bipolar characteristics that contain opposing relationships as alternatives. Within the framework of personal constructs, Kelly (1955) concluded that when individuals indicate something concrete, they also mean an opposite state or property. In contrast to the "similarity/difference" of Kelly's personal cognitive constructs, the meaning dissonance perspective suggests that the features of meaning-building strategies should be considered using the "acceptance/rejection" bipolar meaning constructs.

Personality is not simply an end product, but rather, it is a process. Therefore, an empirical study of personal meaning-building should examine the genesis of meanings and provide a content description of personality in terms of personality traits. For example, Cattell (1965) supposed that each word denoting a certain personality trait is its potential representation. Revealing a complex ensemble of personality traits, as content that colors subjectivity, has certain difficulties, as "the meaning approach is absolutely adequate. Unlike personality traits or dispositions, variability and dynamism are inherent in the very nature of meaning structures and systems. When researchers describe personality in terms of personality traits, they have difficulties over explaining the mechanisms of personal changes. It is obvious that the language of traits is clearly inadequate for these purposes" (D. Leontiev, 2007, p. 253).

Bipolar scales with intermediate sublevels should help to overcome such difficulties. These bidirectional semantic axes represent two related personality traits as well as directions referring to both positive and negative development. These semantic axes reflect the directions of symmetry in the many-sided meaning space of personality and demonstrate the dynamics of changes. Thus, the above-stated difficulties can be overcome by virtue of a "theory that creates the possibility of changes in its explanatory structures" (D. Leontiev, 2007, p. 253).

It may be appropriate to define personality from the perspective of the interdisciplinary approach to meaning dimensions. Many mental processes, states, and properties differ in their nature. However, together, they constitute the sphere of the human psychic. Their differentiation and separate analyses may be superfluous. It is unnecessary to discuss the need to distinguish the concepts of personality traits and characteristics when they are similar and sometimes coincident (A. Leontiev, 1975). Hence, the concept of "personality trait" is broader than is the concept of "character trait," and broader yet than the concept of "psychic trait."

Individuals' perfection is determined by their harmonious interactions in both internal and external spaces of their personality. These spheres are not isolated ones; they often overlap and supplement each other. The scope of interactions and their intensity manifest themselves in personal individuality: "As a rule, the man's properties can be called personal if they characterize him as a subject of relations with the surrounding world and are formed in the relationship with the world" (Lubovsky, 2007, p. 187).

The complex of personality traits is a unique combination that provides an inimitable "pattern" in the diversity of the world of people, things, and relationships. Therefore, personal harmony assumes the paths of development that help individuals approach and correspond to certain key concepts, which reflect their understanding of happiness and meaning of life.

Method

Instrument

The scales of personality traits establish the features of their manifestation in various dimensions of personal functioning. According to Rokeach (1973), the value and meaning dimensions of personality regulate the choice of goals and means of activity in accordance with generalized representations of possible benefits and ways to achieve them. Moreover, personal meanings (as life values) activate appropriate strategies for personal development. The value and meaning potential of personality encompass three main dimensions (Diakov, 2015; Kotlyakov, 2013; Pakhomova, 2011): world outlook, behavior, and cognition.

Each of these dimensions should be described using the language of personality traits by means of bipolar meaning scales. In these scales, personality traits are key denotations as specific markers of meanings. The findings of our studies enabled us to select nine personality scales (Godunov, 2016), which correspond to the developmental and adaptive strategies of meaning-building. For each scale, the three upper words refer to the developmental strategy of meaning-building (+), the middle level shows a neutral state (0), and three lower words refer to the adaptive strategy (–) (*Table 1*).

Table 1

Personal Property Scales

1) world outlook direction: +3 self–sufficiency +2 meaningfulness +1 responsibility 0 disinterest –1 levity –2 inadvertence –3 disorganization	2) behavioral direction: +3 tranquility +2 civility +1 leniency 0 indifference –1 bravado –2 impatience –3 inadequacy	3) verbal–linguistic: +3 eloquence +2 erudition +1 originality 0 conventionality –1 narrowness –2 categoricity –3 stereotype
4) logical–mathematical direction: +3 abstractiveness +2 systemacy +1 logicality 0 linearity –1 inconsistency –2 fragmentariness –3 banality	5) visual–spatial direction: +3 imagery +2 expressiveness +1 accuracy 0 mediocrity –1 disorder –2 disunity –3 disproportion	6) motor–leading direction: +3 vitality +2 plasticity +1 mobility 0 ordinariness –1 mismatch; –2 sluggishness –3 passivity
7) musical and rhythmic direction: +3 rhythmicity +2 musicality +1 proportion 0 mediocrity –1 narrowness –2 obsession –3 monotony	8) interpersonal direction: +3 sociability +2 trustfulness +1 benevolence 0 lack of interest –1 hesitation –2 distrustfulness –3 isolation	9) intrapersonal direction: +3 confidence +2 calmness +1 attentiveness 0 unpretentiousness –1 emotionality –2 irritability –3 suspiciousness

The developed scales of personality traits represent personal profiles and demonstrate actual strategies of meaning-building.

The system of personal meanings and the strategies for building them represent a multifaceted and multidirectional conceptual sphere. However, no research has addressed the issue of a unified methodology or approach. We supplemented the developed scales of personality traits with a battery of psychodiagnostic tests. Our study used the following techniques: (a) the Who Am I test by Kun and McPartland (modified by Rumyantseva, 2006) to investigate the content characteristics of personal identity; (b) the Purpose-in-Life test (PIL) modified by D. Leontiev to examine the factors of life meaningfulness (D. Leontiev, 2000); (c) the Rosenzweig Picture-Frustration Test to analyze behavioral characteristics (Rosenzweig, 1945); and (d) Gardner's Multiple Intelligences test to describe the characteristics of cognition (Gardner, 1983).

Participants

Our empirical study of the psychological characteristics of polar estimation regarding different strategies of meaning-building involved first- and third-year psychology and history students studying at Southern Federal University as well as secondary school teachers residing in the Rostov region ($n = 145$). The participants were grouped according to three criteria: age, gender, and profession. The sample consisted of 102 young people aged 18–23 years and 43 people aged 26–56 years; 112 women and 33 men; and 80 psychology students, 30 history students, and 35 secondary school teachers.

Procedure

After grouping the participants according to their acmeological characteristics, we calculated Pearson's correlation coefficients for test scores. The stable presence of positive correlations ($ps \leq .05$) with the proposed meaning scales were found only for the "life meaningfulness" index in the PIL test. Consequently, we considered this parameter as the main marker of polar strategies of meaning-building. The "life meaningfulness" measure in D. Leontiev's PIL test refers to conscious self-reflection. There were no statistically significant correlations between the meaning scales and other measures. Differences of mean scores in groups based on acmeological characteristics were determined using a Kolmogorov-Smirnov criterion. For the index "life meaningfulness" as a function of age, sex, and profession, the Kolmogorov-Smirnov criterion did not exceed .044, suggesting significant differences in the sample groups ($ps \leq .05$).

We further divided the samples into two groups based on the mean scores of "life meaningfulness." Respondents with higher scores fell into the group with a developing strategy of meaning-building. Those with lower scores fell into the group with an adaptive strategy of meaning-building. In each group, we evaluated mean scores for semantic differential scales. These indices reflect the parameters of personal psychological profiles for the polar strategies of meaning-building.

Results and Discussion

Figures 1–3 show nine semantic differential scales of personal traits that underlie the construction of personal profiles in terms of polar meaning-building strategies. Such an approach enables us to reveal the dynamics of the investigated meaning-building strategies and changes in their indices in terms of acmeological characteristics.

When we grouped the respondents with respect to their age (see *Figure 1*), both age from 18 to 23 years and age from 26 to 56 years groups showed higher scores on the developing strategy. Furthermore, compared with the young respondents, the middle-aged participants had higher or equal scores on most semantic differential scales, with the exception of scale 3 ("verbal-linguistic"). The middle-aged respondents had higher scores on the adaptive strategy of meaning-building on all scales.

When we grouped the respondents with respect to their gender (see *Figure 2*), both genders showed higher scores on the developing strategy, with the exception of coincidence on scales 2 and 4 in the male group. Compared with male respondents, women showed higher scores on the developing strategy on scales 1, 2 and 8; equal scores on scales 4 and 5; and lower scores on scales 3, 6, 7, and 9. Male respondents had higher scores on the adaptive strategy on all the scales, with the exception of scale 8 ("interpersonal").

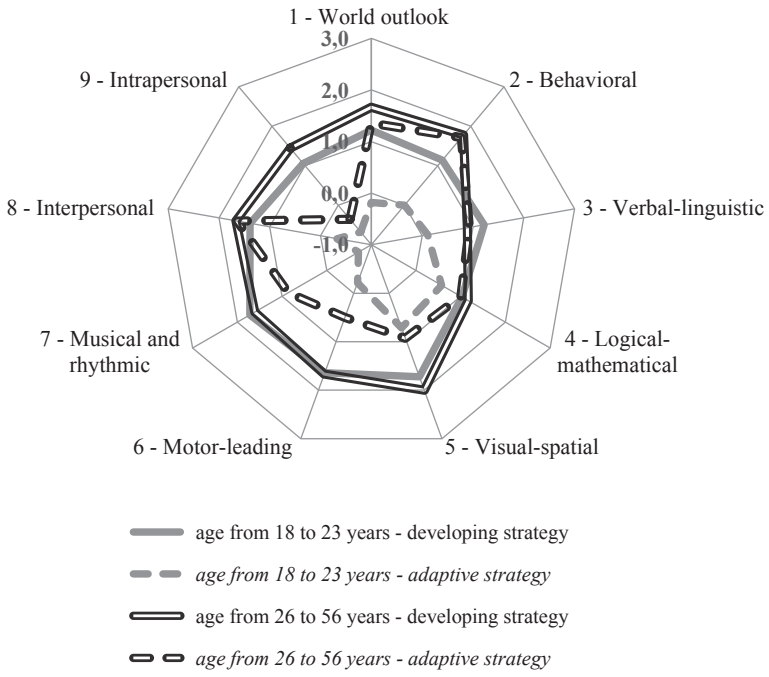


Figure 1. Age changes in meaning-building strategies on semantic differential scales.

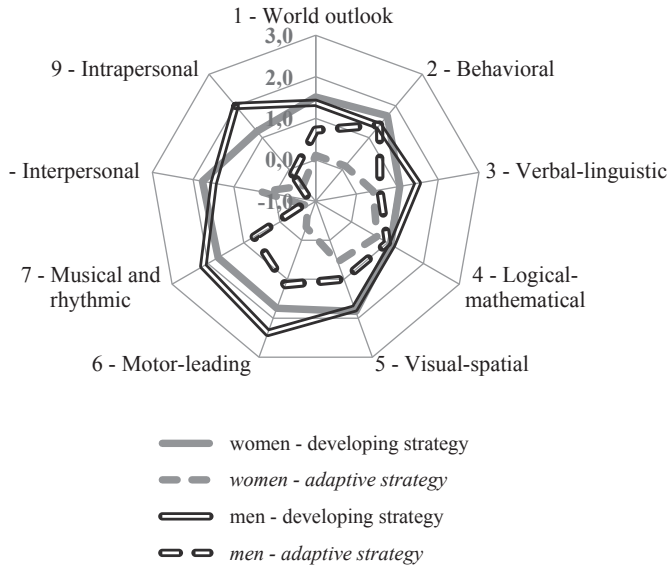


Figure 2. Gender changes in meaning-building strategies on semantic differential scales.

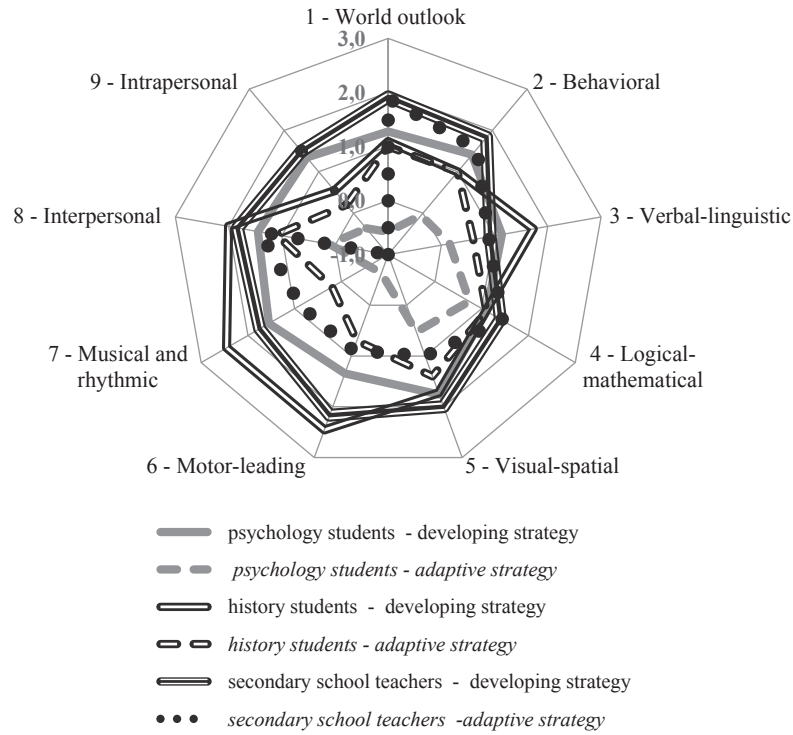


Figure 3. Differences by profession in meaning-building strategies on semantic differential scales.

When we grouped the respondents with respect to their profession (see *Figure 3*), all the professional groups had higher scores on the developing strategy, with the exception of coincidence on scales 2 and 4 among secondary school teachers. The secondary school teachers demonstrated the highest scores on the developing strategy on scales 1, 2, 4, 5, and 9, while the history students had the highest scores of the developing strategy on scales 3, 6, 7, and 8. The secondary school teachers demonstrated the highest scores of the adaptive strategy, with the exception of scales 5 and 9 in the history students group.

Conclusion

The obtained indices of meaning-building strategies did not coincide in all the differentiated groups, which clearly speaks in favor of acmeological dynamics of the respondents' personal profiles. We stratified the sample according to the mean score of the basic marker of "life meaningfulness," which enabled us to establish differences in characteristics of actual polar strategies of meaning-building. The respondents who did not fall into either of the two groups are "between the poles." They often have an underdeveloped meaning-building strategy as a result of poorly formed ways of organization and actualization of personal meanings or the presence of a transitional form of situational conceptual initiations.

The identified personal profiles represent multifactor models of personal values and meaning dimensions, which can predict actual meaning-building strategies using semantic differential scales and indicating markers ("life meaningfulness" from the Purpose-in-Life test) and help researchers to reduce the number of techniques employed in their studies. Further research should pursue a trialectical analysis of meaning-building strategies.

Acknowledgements

This work was carried out as a part of the Southern Federal University internal grant implementation No. 07/2017-01 "Development of technologies for the initiation of meaning-making as a component of modern communication systems for the purpose of ensuring the information security in the Internet."

References

Abakumova, I.V., Ermakov, P.N., & Fomenko, V.T. (2013). *Newdidactics: Book I: Methodology and technology of training: Searching a developing resource*. Moscow: Credo.

Abakumova, I.V., Godunov, M.V., Enin, A.L, & Generdukaeva, Z.Sh. (2016). *Meaning-building strategies: Modern representations in the works of native researchers: textbook.* Moscow: Credo.

Allport, G.W. (2002). *Formation of personality. Selected works.* Moscow: Smysl.

Cattell, R.B. (1965). *The Scientific Analysis of Personality.* Harmondsworth, England: Penguin Books.

Druzhinin, V.N. (2009). *Psychology: A textbook for humanitarian institutions* (2nd ed). St. Petersburg: Piter.

Dyakov, S.I. (2015). Sub'ektnaia samoorganizatsia lichnosti. Psikhosemanticheskii analiz [Subjective self-organization of personality: psychosemantic analysis]. *Vestnik Omskogo universiteta. Seriia:* Psikhologia [Herald Of Omsk University. Series: Psychology], 1, 15–33.

Frick, W. (2000). Remembering Maslow: Reflections on a 1968 interview. *Journal of Humanistic Psychology*, *40*(2), 128–147. https://doi.org/10.1177/0022167800402003

Gardner, H. (1983). *Frames of Mind: The Theory of Multiple Intelligences.* N.Y.: Basic Books.

Godunov, M.V. (2016). *The ensemble of personality traits: Interpreting and diagnosing bipolar semantic scales.* Book 1. Moscow: Credo.

Guilford, J.P. (1967). *The nature of human intelligence.* N.Y.: McGraw-Hill Book & Co.

Kashdan, T.B., Steger, M.F. (2007). Curiosity and path-ways to well-being and meaning in life: Traits, states, and everyday behaviors. *Motivation and Emotion, 31*(3), 159–173. https://doi.org/10.1007/s11031-007-9068-7

Kelly, G. (1955). *The psychology of personal constructs.* N.Y.: Norton.

Kotlyakov, V.Yu. (2013). Metodika "Sistema zhiznennykh smyslov" [The System of Vital Meanings technique]. *Vestnik Kemerovskogo gosudarstvennogo universiteta* [Bulletin of Kemerovo State University], 2 (54), 148–153. https://doi.org/10.21603/2078-8975-2013-2-148-153

Lazursky, A.F. (1924). Klassifikatsiia lichnosti [Personality classification]. In M.Ya. Basov & M.N. Myasishchev (Eds). *Lazursky, A.F.* (3nd ed). Leningrad: State Publishing House.

Leontiev, A.N. (1975). *Deiatel'nost'. Soznanie. Lichnost'* [Activity. Consciousness. Personality]. Moscow: Politizdat.

Leontiev, D.A. (1997). *Ocherk psikhologii lichnosti* [Essay on the psychology of personality]. (2nd ed.). Moscow: Smysl.

Leontiev, D.A. (2000). *Test zhiznennykh orientatsii* [Test life orientations] (2nd ed.). Moscow: Smysl.

Leontiev, D.A. (2007). *Psikhologiia smysla: priroda, struktura i dinamika znacheniia real'nosti* [Psychology of meaning: Nature, structure and dynamics of meaning reality] (3nd ed.). Moscow: Smysl.

Leontiev, D.A. (2002). Simbioz i adaptatsiia ili avtonomiia i transtsendentnost': vybor cheloveka v nepredskazuemom mire [Symbiosis and adaptation or autonomy and transcendence: The person's choice in an unpredictable world]. In Ye.I. Yatsuta (Eds), *Lichnost' v sovremennom mire: ot strategii vyzhivaniia k strategii sozdaniia zhizni* [Personality in the modern world: from the survival strategy to the strategy of life creation] (pp. 3–34). Kemerovo: IPK Graphica, 2002.

Lubovsky, D.V. (2007). *Vvedenie v metodologicheskie osnovy psikhologii* [Introduction to the methodological foundations of psychology]. Moscow: Modek, MPSI.

Pakhomova, E.V. (2011). Metodika otsenki tsennostnykh orientatsii [Value orientations assessment technique]. *Vestnik Samarskoi gumanitarnoi akademii. Seria "Psikhologia"* [Herald of the Samara Humanitarian Academy. Series "Psychology"], *2*(10), 120–134.

Petrenko, V.F. (2013). *Mnogomernoe soznanie: psikhosemanticheskaia paradigma* [Multidimensional consciousness: A psychosemantic paradigm]. 2nd ed. Moscow: Eksmo.

Piaget, J. (1998). O mekhanizmakh assimiliatsii i akkomodatsii [On the mechanisms of assimilation and accommodation]. *Psikhologicheskaia nauka i obrazovanie* [Psychological science and education], 1, 22–24.

Rokeach, M. (1973). *The nature of human values.* N.Y.: Free Press.

Rosenzweig, S. (1945). The picture-association method and its application in a study of reaction to frustration. *Journal of Personality*, *1*(14), 3–23. https://doi.org/10.1111/j.1467-6494.1945.tb01036.x

Rumyantseva, T.V. (2006). *Psikhologicheskoe konsul'tirovanie: diagnostika semeinykh otnoshenii* [Psychological counselling: diagnostics of couple relationships]. St. Petersburg: Rech.

Original manuscript received February 22, 2018
Revised manuscript accepted November 28, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

Gender Differences in Interactions with Avatars of Diverse Ethnic Appearances

Galina Ya. Menshikova*, Olga. A. Tikhomandritskaya, Olga A. Saveleva, Tatyana V. Popova

Lomonosov Moscow State University, Moscow, Russia

*Corresponding Author. E-mail: gmenshikova@gmail.com

Background. Gender differences exist in almost every aspect of our lives. Individuals have an array of different social expectations with regard to behaviors, communication, appearance, attitudes, and social roles, but these expectations tend to be based on whether the individual is male or female. Currently, many social studies have been done with the help of virtual reality technologies. They have been effectively applied to the study of many social phenomena such as nonverbal communication, social skills training, social anxiety rehabilitation, etc. Recently considerable attention has been paid to issue of gender differences during social interaction with the virtual partners, avatars. However, the question of gender differences during interactions with avatars of diverse ethnic appearances has seldom been studied.

Objective. The goal of this study was to investigate the gender peculiarities of interaction with avatars of different ethnic appearances. We used the CAVE virtual reality system to study gender differences in interpersonal distances which were maintained with avatars.

Design. We designed four three-dimensional virtual scenes with avatars of four different ethnic appearances. They were avatars of Slavic, Asian, North Caucasian, and African appearance. All the avatars were male. The participants (who all identified as Russians) were immersed in virtual environments with the help of the CAVE virtual reality system. Their task was to approach the avatar, present herself/himself in any way they wanted, and give instructions for the work. During the task the interpersonal distances between the participants and the avatars were measured. After leaving the CAVE, the participants were asked to fill out a questionnaire assessing the Presence effect.

Results. The results showed gender differences in how much interpersonal distance was maintained: women preferred to keep shorter interpersonal distances with their virtual partners than men did. Moreover, the results showed the impact of ethnic appearance on interpersonal distances. Women approached the avatars of their own ethnic group more closely and kept further away from the avatars of other ethnic groups. Unlike the women, the men stayed the same distance away from the avatars of different ethnic groups. Both gender groups kept equally far away from the avatar of African appearance. Gender differences were also revealed in the participants' estimates of the Presence effect.

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

doi: 10.11621/pir.2018.0414
http://psychologyinrussia.com

Conclusion. We developed a comprehensive method for analyzing the gender differences in interaction with the avatars, including both subjective (the sense of Presence, semantic differential technique) and objective (assessment of interpersonal distance) characteristics. This method allowed us to assess gender variables during the social interaction using both behavioral and psychological responses. The same method can be applied successfully to the study of gender differences regarding other social phenomena.

Keywords: gender differences, interethnic interaction, interpersonal distance, avatar, Presence effect, CAVE virtual reality technology

Introduction

The social perception which occurs in interpersonal interaction is among the most important aspects of communication (Andreeva, 1990). A large number of social, personal, and situational factors have an effect on communicative behavior. One of these factors is gender, which is determined by the behaviors and attitudes that a society considers proper for its males and females (Klecina, 2003; Henslin, Nelson, 1997). Considerable attention has recently been paid to the question of gender differences within social interaction (Andreeva, 2013; Reeder, 2009; Eagly, Wood, 2013). Many studies have shown that women and men communicate very differently in many key aspects (Eagly, 1983; Block, 1976).

In communication, each person takes a certain gender role according to specific gender stereotypes, including behavioral features of men and women which are widespread in his or her society. The content and origins of gender stereotypes differ in different cultures, and vary in intensity and forms across countries. Many modern studies have shown that communication style (nonverbal behavior, linguistic strategies, aggression, leadership, self-esteem, etc.) differs between sexes in their particular cultures (Hofstede, McCrae, 2004; Xiufang Xia, 2013; Zell et al., 2015).

Researchers studying gender pay attention to nonverbal behavioral cues, such as gestures, facial expressions, postures, and appearance. One of the several subcategories in the study of nonverbal communication is proxemics, i.e. how humans use space to interact with others in everyday life. The basic concept of proxemics is interpersonal distance (Hall, 1966), which can transmit a lot of information about the communication process: the degree of sympathy and trust between the partners in the dialogue, the social status of the communicating persons, the duration of their acquaintance, respect, etc. (Henley, LaFrance, 1984).

The spatial behavior of a person takes place in the context of individual psychological, environmental, and socio-cultural factors, which include ethnicity and gender. In particular, the organization of interpersonal space varies according to the partner's gender. Many researchers have found that the space organized by women was more compact than the space organized by men (Kreidlin, 2005). This behavior was considered a manifestation of female communication style: women are more open, friendlier, and more focused on establishing comfortable conditions for cooperation. Studies of spatial behavior have traditionally used methods such as observation, survey, and experiment.

Currently, digital technologies, particularly virtual reality (VR) systems, are in wide use in the studies of social processes (Zinchenko et al., 2015; Fox et al., 2009; Blascovich et al., 2002). Nonverbal behavior is one of the psychological phenom-

ena explored with the help of VR technologies, which were developed in order to study interaction with virtual partners. It has been shown that the participants may perceive a virtual partner in two different ways: as an agent (a virtual partner controlled by a computer program), or an avatar (a partner representing another character). In the latter case the participant's behavior was revealed to be very similar to the behavior that he or she exhibits with human beings (Lucas et al., 2014; Bailenson et al., 2008). This similarity was particularly evident in the proxemics: participants preferred not to violate their personal space when communicating with avatars (Bailenson et al., 2003; Menshikova et al., 2018). These and many other studies have shown that interaction with virtual partners can contribute to social behavior training, and then the skills acquired in a virtual environment extrapolated into real life (Menshikova et al., 2017).

Much of the research applying VR systems to studying social phenomena has considered gender behavior in the process of "user-avatar" interaction. For example, in a 2013 study by Preda and Jovanova, the influence of the *gender* of both the *participant and the avatar* on the efficiency of communication was investigated. Virtual partners (male and female), who differed in the degree of their behavioral realism, were designed. The avatars were constructed highly realistically: Their lip movements could match the phonemes spoken; they could blink their eyes, and turn their heads to follow the participant's glance to keep eye-contact. The results showed that the participants perceived avatars of their own gender as more truthful (Nowak, Rauh, 2008), and that a male agent was perceived as more convincing compared to a female avatar.

The avatar's gender has also been demonstrated to be a factor (Petrakou, 2010; Pals et al., 2014). Researchers investigating the interpersonal distance maintained with the avatars have found that the most important factors affecting whether an avatar would be admitted into the user's interpersonal space were:

- the gender of both the participant and the avatar: if they coincided, the interpersonal distance decreased;
- the direction of the avatar's gaze: in the case of direct eye-contact, the interpersonal distance increased.

The influence of the gender and personality characteristics of participants on their emotional preferences when communicating with avatars was also been demonstrated (Diemer et al., 2015). Diemer et al. showed that women who were introverts preferred to communicate with avatars which were similar to themselves in appearance, whereas women who were extroverts chose to communicate with avatars of different appearance, e.g. having brighter clothes or different hairstyles, or even belonging to other ethnic or racial groups.

From the very beginning of VR application to social studies, the question of whether the participant believed the virtual world as real has been actively discussed (Riva, 1999). The perceptual experience, which enables users to believe that they are "being there" in the virtual world, is important for assessing the impact of the virtual environment. The effectiveness of virtual environments has often been linked to the sense of Presence. Presence is defined as the subjective experience of being in the virtual environment, even when the user is physically situated in the real environment. The Presence effect is considered an indicator of immersion in

the virtual environment and is measured using the Presence questionnaire (Witmer, Singer, 1998).

The application of VR systems has proven effective in the study of gender differences of the "user-avatar" interaction. However, some aspects of this issue have not yet been studied sufficiently. In particular, the interaction between two important factors of communication–the gender and ethnic appearance of the avatars–remains unclear. To measure the role of a participant's gender in interactions with avatars of different ethnic appearances, we evaluated the behavioral (interpersonal distances) and psychological (Presence effect) responses.

Goal

The goal of our study was to investigate gender differences during the interactions with avatars of different ethnic appearances. The following research hypotheses were tested:

- There will be gender differences in the interpersonal distances maintained with avatars: women will keep shorter interpersonal distances from avatars in comparison with men.
- There will be gender differences in the interpersonal distances depending upon the ethnic appearance of the avatars: women will come closer to avatars of their own ethnic group.
- There will be gender differences in the intensity of Presence effect.
- The interpersonal distance will correlate with the intensity of the Presence effect: the higher the degree of Presence, the longer interpersonal distance will be chosen by the participants when interacting with the avatars.

Design

Participants. The sample consisted of 41 participants (22 females, 19 males, age 18 to 26 years). All participants had normal or corrected-to-normal vision and had no vestibular disorders or brain injuries. All participants declared themselves ethnic Russians. All of them live permanently in Moscow and the Moscow region. The small sample size was due to the complexity of our experimental procedure and the large amounts of data they produce.

Stimulation. Four three-dimensional virtual scenes with avatars were designed. Each virtual scene represented a large room connected to the office where the virtual partner was standing. Four types of avatars of different ethnic appearance were created, including avatars of Slavic, Asian, North Caucasian, and African appearance. All avatars were male. The appearance of the avatars slightly differed in the manifestation of masculine-feminine features: the avatar of North Caucasus appearance had more expressed masculine features, the avatar of Slavic appearance showed less pronounced masculine features, and the avatars of Asian and African appearance had moderately expressed features. *Figure 1* shows a scene of interaction of a participant with one of the avatars of African appearance.

Equipment. The virtual scenes were presented using CAVE virtual reality system Barco Ispace 4 with a screen resolution of 1280 x1024 and a field of view of

180°. Special glasses CrystalEyes 3 allowed the participant to form 3D virtual objects and characters. Participants could move through the virtual environment using a manipulator Flystick2. The participant's location was recorded with an 8 Hz update frequency. *Unity 3D* Professional was used for the software development.

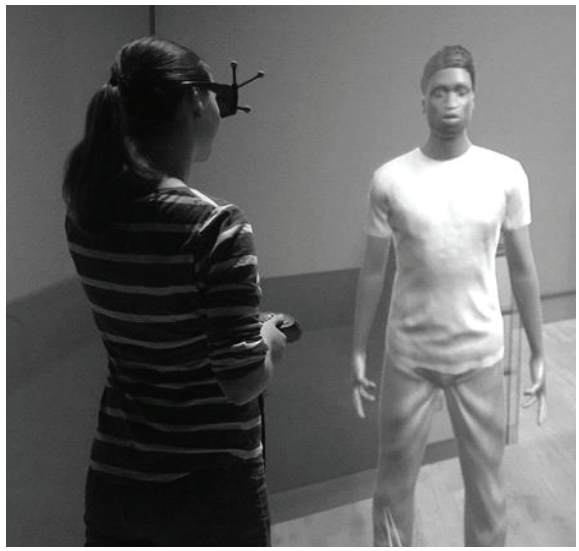


Figure 1. A participant communicating with an avatar of African appearance.

Procedure. The participants first had to be trained to navigate through the virtual space using a Flystick. When they reported that they were ready, the main experiment started. Each participant got into the starting position in the center of the virtual room, and was asked to go into the office room where the avatar was standing. According to the scenario, the participant played the role of an employer who had to hire a designer, an architect, and an engineer (avatar roles) to build country houses. The avatars represented workers who came for an interview. The participant's task was to approach the avatar, introduce herself/himself in any way they wanted, and give instructions for the house construction. There was no time limit. Then the participant interacted with the three avatars of other ethnic appearance in the same way. The order of interaction with the avatars of different ethnic appearance was quasi-random. The spatial coordinates of the participant's movement in the virtual environment were recorded. On average, the interaction process with all avatars lasted for 2–3 hours.

After all the tasks in the virtual environment were completed, each participant was asked to fill out a questionnaire to assess the strength of the Presence effect. The standard questionnaire of B. Whitmer and M. Singer (Witmer, Singer, 1998) was modified in accordance with the objectives of this experiment. The modified version contained 17 questions, which assessed the general impression of the virtual scenes, and the avatar's appearance and behavior. Four questions specifically addressed the realism of the virtual scenes (№ 10, 11, 14, 16); seven questions concerned the emotional impression of "being there" (№ 1, 4, 6, 7, 8, 9, 17); two ques-

tions were about the appearance of avatars (№ 12, 13); and four questions were about any discomforts of nausea and disorientation arising in the virtual space (№ 2, 3, 5, 15). The participants rated the questions on a scale of 1 to 5, where 1 corresponded to the minimum, and 5 to the maximum of the sense of being there, in the virtual environment.

Statistical processing. The data were processed with the use of SPSS Statistics 22 software package. The interpersonal distances that participants maintained with the virtual humans were analyzed. The analysis encompassed only the particular zone within which the participant could have an "eye-to-eye" contact with the avatar, not the entire trajectory of motion. This zone was chosen to reduce the variability of interpersonal distance values due to the compensation effect. This effect showed that participants establish a longer distance from the avatars when they are in the "eye-to-eye" zone, and a shorter distance once they leave this zone (Bailenson et al., 2003). For the virtual scenes the boundaries of this zone were determined as values of the visual angle in the range of ±30° relative to the axis direction of the participant's and the avatar's reciprocal gaze. Within this zone the mean values of interpersonal distance were calculated. The values of the Presence effect were calculated as the average ratings of female and male samples for each ethnic version of the avatar, and for each question separately.

Results

Two types of data were analyzed separately for the female and male samples:

- the interpersonal distance that participants maintained from the avatars of different ethnic appearance;
- the perceived intensity of the Presence effect and its correlation with the interpersonal distances.

Interpersonal distance maintained by participants with avatars of different ethnic appearance. Based on the data, the average interpersonal distances chosen by the participants during their interaction with avatars of different ethnic groups were calculated separately for the female and male samples. The results are given in *Table 1*, which shows the average values of distances (in meters) when the participants were facing avatars of African, Asian, Slavic, and North Caucasian appearance. Standard deviations are given in parentheses.

Table 1

The average interpersonal distance maintained with avatars of different ethnic appearances by the male and female participants

	Avatars of Slavic appearance	**Avatars of North Caucasian appearance**	**Avatars of Asian appearance**	**Avatars of African appearance**
Males	2.21 (0.56)	2.22 (0.69)	2.21 (0.67)	2.35 (0.57)
Females	2.10 (0.59)	2.09 (0.67)	2.13 (0.56)	2.22 (0.64)

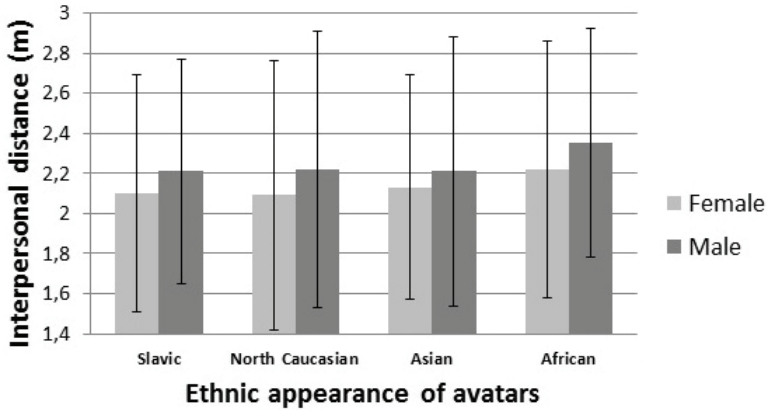


Figure 2. Interpersonal distance (in m) maintained by male (black bars) and female (grey bars) participants when interacting with the avatars of different ethnic appearance

The x-axis identifies the four types of ethnic appearance of the avatars. The y-axis shows the average interpersonal distance in meters. As seen from the *Figure 2*, women preferred to maintain shorter interpersonal distance with the avatars than men, regardless of the ethnic appearance of the avatars. The gender differences between the mean interpersonal distances were calculated in pairwise comparison of the male and female samples. A trend of differences between male and female participants interacting with avatars of any ethnic group was observed: with avatars of Slavic appearance ($t(44) = 3.57$, $p < 0.08$); avatars of North Caucasian appearance ($t(44) = 3.87$, $p < 0.05$); avatars of Asian appearance ($t(44) = 4.15$, $p < 0.04$); and avatars of African appearance ($t(44) = 4.07$, $p < 0.07$). Thus, the data partially confirmed our hypothesis about gender differences: women were more likely to maintain shorter interpersonal distances during interaction with the avatars than did the men.

These results agree with those of previous studies that revealed the same female behavior: women approached their human communication partner more closely than men did (Evans, Howard, 1973; Gifford, 1982). A similar style of behavior was observed when communicating with the virtual partners. (Bailenson et al., 2003; Bailenson et al., 2008). So, the same social rules are applied in the interactions with both real and virtual partners. The reason for this behavior may be a desire to establish a more emotional relationship with a partner, even if this partner is a virtual character (Labunskaya, 1988; Preda, Jovanova, 2013).

The results also revealed a trend of gender differences in maintaining distances from avatars of different ethnic groups. Women most closely approached the avatars of their own ethnic group; a little less the avatars of North Caucasian appearance ($t(44) = 3.65$, $p < 0.09$); and then kept significantly further away from avatars of the Asian ethnic group ($t(44) = 2.44$, $p < 0.05$). The longest distance was maintained from the avatars of African appearance ($t(44) = 3.57$, $p < 0.04$).

Men adhered to a completely different type of behavior. They maintained approximately the same distance from the avatars of Slavic, North Caucasian, and Asian ethnic appearances: The differences between the mean interpersonal dis-

tances with pairwise comparison of all groups were insignificant. The only exception was avatars of African appearance from whom male participants preferred to maintain a significantly longer distance ($t(44)=2.37$, $p<0.05$).

We explain this behavior in the following way: our participants, being residents of Moscow and the Moscow region, rarely communicate with representatives of this ethnic group, which in turn may lead to some emotional tension manifested in the increased interpersonal distance. The studies of gender differences in proxemics reveal that men and women behave the same way in relation to strangers (Sukhova, 2002). Thus, these results again show the general strategies of interaction with the real and virtual partners.

The perceived intensity of the Presence effect.

The average ratings of the Presence effect were calculated for female/male participants separately for each question on the questionnaire. The average scores for each question are shown in *Fig. 3*. The numbers of the questions are on the x-axis, and the y-axis represents the intensity of the Presence effect scored from 1 (minimal intensity) to 5 (maximum intensity). The black bars in *Fig. 3* show the average values for the males, and the gray bars those for the females

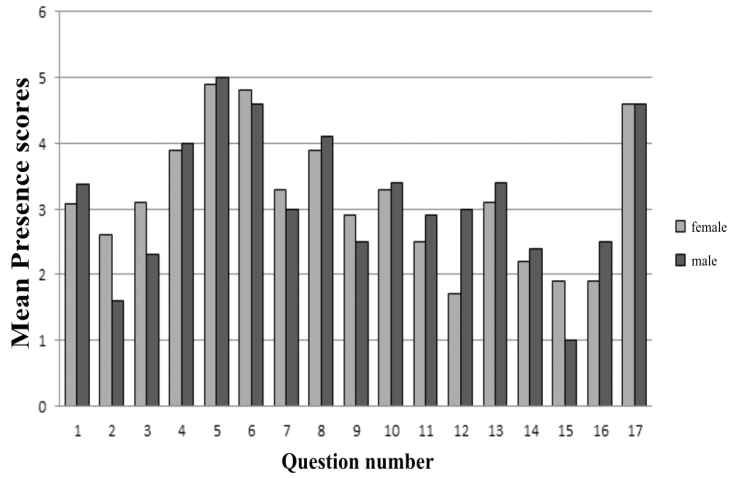


Figure 3. Mean Presence scores versus the question number for the female (grey bars) and male (black bars) participants

for each question of the questionnaire. In general, the estimates of the Presence effect for the virtual environments and avatars were above average, varying from a minimum value of 1.5 points up to a maximum of 5 points. The analysis of the mean ratings showed that the Presence effect differed insignificantly in the male and female samples: many questions were rated almost equally by men and women.

However, significant differences were found for some questions. According to the responses to questions 2 (I lost orientation in space) and 15 (I felt queasy), men experienced this feeling to a much lesser extent than women; as to question 12 (it seemed to me that the avatar could actually exist), men more often believed in the

realism of their virtual partners. This result may be explained by the participants' different experiences in using VR systems. Preliminary interviews with our participants showed that men and women differed in their degree of experience interacting with virtual environments. Fifty percent of males had some experience of being in a virtual environment (HMD, CAVE, SUIT, VR), while among females the level of experience was much lower (about 21%). Thus, as the results show, men are much less likely to lose their orientation in a virtual environment than women, and they are also more likely to perceive avatars as real people. This result coheres with the studies of Maccoby and Jacklin (1974), who revealed the greater ability of men to orient themselves to their environment. This ability also could manifest itself in a virtual environment.

We also calculated the sample correlation coefficient between the values of the mean interpersonal distances and Presence scores, averaged over all 17 questions. The results showed a high negative correlation ($r=0.57$; $p=0.01$), indicating a more pronounced compensation effect if the participant's Presence score was higher. In other words, the more realistic the virtual environment was perceived to be, the larger distances participants kept from the avatars, whatever their ethnic appearance.

In general, the interaction with avatars in a virtual environment was not difficult for our participants; they used the same models of non-verbal behavior as in real communication.

Conclusion

We developed a comprehensive method for estimating gender differences during interactions with avatars of diverse ethnic appearances, including both subjective (the sense of Presence) and objective (assessment of interpersonal distances) measurements.

Variation by gender was evident in the participants' maintenance of interpersonal distances from the avatars: women were more likely to keep shorter distances away from virtual partners than men.

This trend of gender differences in maintaining interpersonal distance also varied according to the ethnic appearance of the avatars. Women came closer to the avatars of their own ethnic group and kept further away from avatars of other ethnic groups. Unlike women, men adhered to the same distance from the avatars of different ethnic groups. Both kept equally far from the avatar of African appearance.

The Presence effect was almost the same for both gender groups, except that men are less likely to lose orientation in a virtual environment than women, and are more likely to perceive avatars as real people.

The interpersonal distance correlated with the intensity of the Presence effect: the higher the degree of Presence, the longer interpersonal distance was maintained by the participants when interacting with the avatars.

In summary, we have shown that nonverbal behavior of men and women during interaction with avatars in a virtual environment did not differ significantly from their behavior in communication with real people. The usefulness of our method for assessing gender differences is that it gives us the ability to explore

participants' nonverbal behaviors when interacting with avatars of different ethnic appearances.

Acknowledgements

This research was supported by Russian Science Foundation, the Grant № 15-18-00109, and with the use of equipment purchased with the funds of the "Program of Development of the Lomonosov Moscow State University".

References

Andreeva, G.M. (1990). *Social psychology.* Moscow: Progress.

Andreeva, G.M. (2013). Sotsial'naya psikhologiya v prostranstve sovremennoi nauki i kul'tury [Social psychology in space of modern science and culture]. *Psikhologicheskie Issledovaniya [Psychological Studies]. 30*(6). 2–11. Retrived from http://psystudy.ru/index.php/num/2013v6n30/861-andreeva30.html

Bailenson, J.N., Blascovich, J., Beall, A.C., & Loomis, J.M. (2003). Interpersonal distance in immersive virtual environments. *Personality and Social Psychology Bulletin. 29*(7). 819–833. https://doi.org/10.1177/0146167203029007002

Bailenson, J.N., Yee, N., Blascovich, J., & Guadagno, R.E. (2008). Transformed social interaction in mediated interpersonal communication. In E. Konijn, S. Utz, M. Tanis, & S. Barnes (Eds.), *Mediated interpersonal communication* (pp. 77–99). New York, NY: Routledge.

Blascovich, J., Loomis, J., Beall, A.C., Swinth, K. R., Hoyt, C.L., & Bailenson, J.N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*(2), 103-124. https://doi.org/10.1207/S15327965PLI1302_01

Block, J.H. (1976). Issues, problems, and pitfalls in assessing sex differences: A critical review of the psychology of sex differences. *Merrill-Palmer Quarterly,* 22, 193–203.

Diemer, J., Pauli, P., & Mühlberger, A. (2015). Virtual Reality in Psychotherapy. In J.D. Wright (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., pp. 138–146). Oxford: Elsevier.

Eagly, A.H. (1983). Gender and social influence: A social psychological analysis. *American Psychologist,* 38, 971–981. https://doi.org/10.1037/0003-066X.38.9.971

Eagly, A.H. & Wood, W. (2013). The nature-nurture debates: 25 years of challenges in the psychology of gender. *Perspectives on Psychological Science*, 8, 340–357. https://doi.org/10.1177/1745691613484767

Evans, G.W., & Howard, R.B. (1973). Personal space. *Psychological Bulletin.* 80, 334–344. https://doi.org/10.1037/h0034946

Gifford, R. (1982). Projected interpersonal distance and orientation choices: Personality, sex, and social situation. *Social Psychology Quarterly,* 45, 145–152. https://doi.org/10.2307/3033647

Fox, J., Arena, D., & Bailenson, J.N. (2009). Virtual reality: A survival guide for the social scientist. *Journal of Media Psychology: Theories, Methods, and Applications, 21*(3), 95–113. https://doi.org/10.1027/1864-1105.21.3.95

Hall, E.T. (1966). *The Hidden Dimension.* New York: Anchor Books.

Henslin, J. M. & Nelson, A. (1997). *Essentials of sociology: A down-to-earth approach.* Scarborough, Ont: Allyn and Bacon.

Henley, N.M. & LaFrance, M. (1984). *Gender as Culture: Difference and Dominance in Nonverbal Behavior: Perspectives, Applications, Intercultural Insights.* New York: Lewiston.

Hofstede, G. & McCrae, R.R. (2004). Culture and personality revisited: Linking traits and dimensions of culture. *Cross-Cultural Research, 38*, 52–88. https://doi.org/10.1177/1069397103259443

Klecina, I.S. (2003). *Workshop on Gender Psychology.* St. Petersburg: Peter.

Kreidlin, T.E. (2005). Men and women in non-verbal communication. Moscow.

Labunskaya, V.A. (1988). *Nonverbal behavior (socio-perceptive approach*). Rostov-on-Don: Phoenix.

Lucas, G.M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: virtual humans increase willingness to disclose. *Computers in Human Behavior,* 37, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Maccoby, E.E. & Jacklin, C.N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Menshikova, G.Ya., Kovalev, A.I., & Luniakova, E.G. (2017) Vliyanie rasovoy prinadlejnosti litza na virajennost effekta vzglyada-podskazki: metod aitrekinga [Studying the influence of race on the gaze cueing effect using eye tracking method]. *Natsionalnyy psikhologicheskiy zhurnal [National Psychological Journal],* 2(26). pp. 46–58.

Menshikova, G.Ya., Saveleva, O.A., & Zinchenko, Yu.P. (2018). The study of ethnic attitudes during interactions with avatars in virtual environments. *Psychology in Russia: State of the Art, 11*(1), 20–31.

Menshikova, G.Ya, Zinchenko, Yu.P., Chernorizov, A.M., Tikhomandritskaya, O.A., Shaigerova, L.A., Isaychev, S.A, Kovalev, A.I., Saveleva, O.A., Kraschilschikova, N.O. & Popova, T.V. (2017). Tekhnologiya virtualnoy i dopolnennoy realnosti v issledovanii fenomena etnokulturnoy identichnosti, mezhetnicheskogo i mezhkulturnogo vzaimodeystviya [The technology of virtual and augmented reality in the study of the phenomenon of ethnocultural identity, interethnic and intercultural interaction]. In Rossiyskaya identichnost. Psikhologicheskoe blagopoluchie. Sotsialnaya stabilnost [Russian identity. Psychological well-being. Social stability] (pp. 260–330), Moscow, Lomonosov Moscow State University (in Russ.).

Nowak, K.L. & Rauh, C. (2008). Examining the perception process of avatar anthropomorphism, credibility and androgyny in static and chat context. *Computers in Human Behavior, 24*(4), 1473-1493. https://doi.org/10.1016/j.chb.2007.05.005

Pals, R., Steg, L., Dontje, J., Siero, F., & Van der Zee, K. (2014). Physical features, coherence and positive outcomes of person–environment interactions: A virtual reality study. *Journal of Environmental Psychology, 40*, 108–116. https://doi.org/10.1016/j.jenvp.2014.05.004

Petrakou, A. (2010). Interacting through avatars: Virtual worlds as a context for online education. *Computers & Education, 54*(4), 1020-1027. https://doi.org/10.1016/j.compedu.2009.10.007

Preda, M. & Jovanova, B. (2013). Avatar interoperability and control in virtual Worlds. *Signal Processing: Image Communication, 28*(2), 168–180.

Reeder, H.M. (2009). A critical look at gender difference in communication research. *Communication Studies,* 47(4), 318–330. https://doi.org/10.1080/10510979609368486

Riva, G. (1999). Virtual reality as communication tool: a socio-cognitive analysis. *Presence Teleoperators & Virtual Environments, 8*, 462–468. https://doi.org/10.1162/105474699566341

Sukhova, N. (2002). Gendernye osobennosti neverbalnogo povedeniya [Gender features of nonverbal behavior]. In Gendernye issledovaniya i gendernoe obrazovanie v vysshey shkole [Gender Studies and Gender Education in Higher Education]. Ivanovo: IvGU. Retrived from file:///Users/anastasia-sidneva/Downloads/2_Gender_Ivanovo_02%20(1).pdf

Witmer, B.G. & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environment, 7*, 225–240. https://doi.org/10.1162/105474698565686

Xiufang Xia (2013). Gender Differences in Using Language. *Theory and Practice in Language Studies, 3*(8), 1485–1489. https://doi.org/10.4304/tpls.3.8.1485-1489

Zell, E., Krizan, Z., & Teeter, S.R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70*(1), 10–20. https://doi.org/10.1037/a0038208

Zinchenko, Yu. P., Kovalev, A.I., Menshikova, G.Ya., & Shaigerova, L.A. (2015). Postnonclassical methodology and application of virtual reality technologies in social research. *Psychology in Russia: State of the Art, 8*(4), 60–71. https://doi.org/10.11621/pir.2015.0405

Original manuscript received April 12, 2018
Revised manuscript accepted November 16, 2018
First published online December 30, 2018

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

Russian Psychological Society
Lomonosov Moscow State University

Development of Parental Competence Through Psychological and Pedagogical Support for Families in the Upbringing of Hearing-Impaired Children

Lyudmila N. Molchanova[a*], Anna V. Chekanova[b]

[a] *Kursk State Medical University, Kursk, Russia*

[b] *Gymnasium No. 1, Severomorsk, Russia*

[*] *Corresponding author. E-mail:* molchanowa.liuda@yandex.ru

Background. The need for psychological and pedagogical support for families in the upbringing of hearing-impaired children makes it imperative to develop innovative methods and an effective model of interaction between the family and a special needs educational institution, to improve parental competence.

Objective. To study the psychological content of parental competence (its cognitive, value-motivational; emotional and behavioral components) and to evaluate parental competence through psychological and pedagogical support for families in the upbringing of hearing-impaired children.

Design. Eighty-seven families with hearing-impaired children from a special needs educational institution in Kursk, Russian Federation, participated in the experimental study. The researchers took measurements at two time points, baseline and followup. At baseline, we made a diagnostic assessment of the psychological content of parental competence. At followup, we evaluated the development of parental competence resulting from the psychological and pedagogical support for these families.

Results. *The cognitive component* was characterized by predominant unanimity between the parents in the upbringing of hearing-impaired children, and a partnership relationship in communicating with them.

The emotional component was represented by the absence of difficulties in understanding the causes of the children's emotional state and an orientation towards the child's emotional state during interactions or physical contact.

Terminal values (such as health, happy family life) and instrumental values (such as responsibility, honesty) were predominant in *the value-motivational component*.

The behavioral component displayed a predominance of the authoritative style in upbringing, whereby parents realized their important role in the development of a child's personality and recognized the right of children to self-development. At the same time, the authoritarian style was still significant.

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

doi: 10.11621/pir.2018.0415
http://psychologyinrussia.com

Conclusion. A model for psychological and pedagogical support of families in the upbringing of hearing-impaired children was developed, tested, and found to be effective.

Keywords: parental competence, psychological and pedagogical support, upbringing, hearing-impaired children

Introduction

Families with hearing-impaired children, as a micro-social system, are characterized by somatic, social, and psychological problems; these can be considered as determinants of the psychological and pedagogical support they receive (Asberg, Vogel, & Bowers, 2008; Hintermair, 2006; Koester & Meadow-Orlans, 1990; Koester, Papoušek, & Smith-Gray, 2000; Molchanova & Sitnikova, 2015; Pipp-Siegel, Sedey, & Yoshinaga-Itano, 2002; Quittner, Glueckauf, & Jackson, 1990; Vaccari & Marschark, 1997).

Quittner et al. (1990) compared the model of social support for mothers of deaf children and an alternative model based on the participation of a mediator, as well as the possibility of using such models to adapt to chronic stress. They found differences between the functions of social support in chronic and acute stressful conditions of the mothers of deaf children, as well as between specific situational stressors and major life events. The researchers established that chronic parental stress was associated with a lowered perception of emotional support and severe symptoms of depression and anxiety.

Vaccari and Marschark (1997) described the role of early communication between hearing parents and deaf children in the children's social and emotional development, since more than 90% of deaf children had hearing parents, mostly without effective means of communicating with them. The authors analyzed the interrelations between early communication of hearing parents and the social and emotional development of their deaf children, which become the basis for development of non-verbal strategies for interaction between parents and children.

Pipp-Siegel, Sedey, and Yoshinaga-Itano (2002) presented potential and additional predictors of parental stress among mothers of young hearing-impaired children. They demonstrated that the mothers who perceived their daily problems as more intensive had higher rates of stress. Additional predictors of maternal stress were social support and annual family income.

Kurtzer-White & Luterman (2003) presented a perspective about chronic parental grief upon the birth of their hearing-impaired child, the parents' coping, and the impact of such chronic grief on the relationship between parents and children .

A study of the correlations among parental resources, social and demographic variables, parental stress, and the social and emotional problems of hearing-impaired children (deaf or hard of hearing), showed that high parental stress was associated with children's social and emotional problems (Hintermair, 2006). The research emphasized the importance of a resource-oriented strategy of counseling and social support in an early intervention process.

The traditional model of psychological and pedagogical support for families in the upbringing of hearing-impaired children currently is considered to be the individual educational route of every child, considering his or her optimal social and psychological adaptation.

Calderon (2000) examined the impact of parental involvement in deaf education programs on the development of children's language, early reading skills, and social and emotional development. Factors that were considered included extent of hearing loss among the children, mother's education, her communication skills, and the use of additional services (other than those offered by the early intervention program or the relevant school program). The results showed that parental involvement in the school educational program was a significant positive predictor for the development of early reading skills, while maternal communication skills and extent of hearing loss among children were the strongest predictors for the development of language. And although the participation of parents in the school educational program of their deaf children had a positive impact on academic performance, the ability of the child to communicate with the parents is a more significant predictor of positive language and academic development.

We offer a model of psychological and pedagogical support in which families raising hearing-impaired children act as a full and competent subject in the upbringing and educational process carried out by a special needs educational institution.

Our review of Russian and foreign studies of parental competence shows that there is uncertainty both in understanding its essence and offering effective methods for its formation.

The terms "parental competence" (Léonard & Paul, 1995; Glăveanu, 2015), "psychological competence of parents" (Shagraeva, 2010), "pedagogical competence" (Topilina, 2016), "parental effectiveness", "social and psychological competence of parents", and "competent parenthood" (Cere, 2013; Teti & Candelaria, 2002) are widely used in the psychological and educational literature. They can be considered as synonyms for "psychological and pedagogical competence of parents" (Korobkova, 2011).

The literature presents several different approaches to an understanding of the essence of "parental competence":

According to the cognitive approach, parental competence is considered as the knowledge, skills, and ways of performing pedagogical activities, as well as the knowledge, skills, and experience of parents in raising children (Glăveanu, 2015; Mizina, 2010). Glăveanu (2015) writes that parental competence is the system of knowledge, skills, abilities, and habits that allow parents to successfully perform their parental duties and prevent/resolve crisis situations, in a way that promotes the development of their children.

The activity approach interprets parental competence as unanimity in the theoretical and practical readiness of parents to perform pedagogical activities, and the ability to understand the needs of children and create conditions to satisfy them (Gladkova, 2009).

The competence approach explains parental competence as a set of personal and activity characteristics manifested in the readiness and ability to accept the children as having value, and successfully socializing them (Selina, 2003; Teti & Candelaria, 2002). Teti & Candelaria (2002) define competent parenting only with respect to successful socialization of children. This refers to the parents' mastery of

competencies based on having the knowledge and skills to actualize them according to the social or educational situation.

The structure of parental competence is also being actively studied, pointing the way to prospects for future research (Gorlova, 2010; Kovalenko 2016), and presenting the main methods and forms of activities related to the development of parental competence (Korobkova, 2011).

Many researchers traditionally distinguish the following components in the structure of parental competence:

1. Value-motivational, personal, cognitive, communicative, creative, and reflective (Mizina, 2010);
2. Motivational and personal, gnostic, communicative, and activity-oriented (Minina, 2013);
3. Motivational and personal, gnostic, communicative, and activity-oriented, competence experience (Sergeeva & Arakelyan, 2016).

In our view, special needs education supplements each component of parental competence in specific ways, depending on the interaction of its subjects, and the content of solvable psychological and pedagogical tasks posed.

We support the view of Kovalenko (2016) in our model of psychological and pedagogical support for families in the upbringing and education of hearing-impaired children. He writes that competent parents know how to establish a relationship of trust with their children, sense the state they are in, and understand their behavior. Moreover, parental competence is a set of interrelated qualities related to the parent's personality, qualities which are necessary for effective childrearing, including the cognitive, value-motivational, emotional, and behavioral components.

Lectures, conferences, practical exercises, meetings of teachers and parents, training, etc. (Ovcharova, 2006) are considered effective methods for developing parental competence in the upbringing and education of hearing-impaired children. However, acquaintance with age-related patterns in the development of children and methods of self-help are not always provided in work with families of hearing-impaired children.

In this regard, research aimed at specifying which methods for the development of parental competence are the most effective, is an appropriate and popular issue in psychological and pedagogical practice. We assume that the most effective way of working with families raising hearing-impaired children is to combine pedagogical and psychological methods within their support system.

Thus the aim of the present study was to look at the psychological content of parental competence (its cognitive, value-motivational emotional, and behavioral components) and to evaluate the degree to which it is achieved through the psychological and pedagogical support of families raising hearing-impaired children.

The target of the study was families raising hearing-impaired children. The subject was the development of parental competence through psychological and pedagogical support of these families.

Methods

Participants

Eighty-seven families with hearing-impaired children from Kursk Boarding School, a special needs educational institution, participated in the study. During the school year, the parents interact with their children approximately one day a week or every two weeks. The age of all parents was between 25 and 46 years. Eighteen of them had a higher education, 25 had secondary vocational education, 26 had primary vocational, and 18 had secondary education. The group of mothers was aged $\dot{X}_1 \pm \sigma = 34.6 \pm 5.1$; with higher education — 13, secondary vocational — 16, primary vocational — 7, and secondary — 5. The group of fathers was aged $\dot{X}_1 \pm \sigma = 39.5 \pm 5.6$; with higher education — 5, secondary vocational — 9, primary vocational — 19, and secondary — 13. The 87 children were aged $\dot{X}_1 \pm \sigma = 9.5 \pm 1.8$ and comprised of 39 boys and 48 girls. Twenty children had the first degree of hearing impairment, 26 — the second degree, 19 the third degree, and 22 the fourth degree.

Instruments and Measurement Procedures

The experimental study involved two stages: baseline and followup. The baseline stage assessed the psychological content of parental competence (its cognitive, value-motivational, emotional, and behavioral components), while the followup stage evaluated the development of parental competence through psychological and pedagogical support for the families raising hearing-impaired children.

The baseline and followup stages used the Russian version of the Rokeach Value Survey (Rokeach, 1973), a questionnaire on emotional relationships in families (Zakharova, 1997), and a questionnaire for parents that identifies the specifics of the communication between parents and children, as well as levels of upbringing activities (Zvereva & Krotova, 2009).

Statistical processing of the results was performed with Statistica 11.0. Descriptive and comparative statistics were used in the quantitative and qualitative evaluation of the data (statistical criterion for the significance of the differences (φ^*) of the angular Fisher transformation).

Procedure

The study was performed in several stages.

The first stage included motivation of the parents and giving them information about the tasks, objectives, expected results, and practical importance of the study. It ended with their signing a statement of informed consent. Instructions based on the psychological and diagnostic materials were given to all participating parents.

The second stage was the experimental study.

The third stage involved the statistical processing of the primary results. The psychological content of parental competence (its cognitive, value-motivational, emotional, and behavioral components) was studied with the help of descriptive statistics. Parental competence was evaluated by comparative analysis (statistical criterion for the significance of the differences (φ^*) of the angular Fisher transformation). Fisher's exact test makes it possible to evaluate the reliability of the dif-

ferences between the percentages of two samples based on the significance of the effect of interest to the researcher (Sidorenko, 2006).

The fourth stage included analysis and generalization of the results and the drawing of conclusions.

Results

Diagnostics of the cognitive component of parental competence used the questionnaire for parents (Zvereva & Krotova, 2009). The objective of the questionnaire was to identify the parents' understanding of the need to develop communication skills, to observe ethical standards in communicating with their child and in interpersonal communication in the child's presence. The specifics of communication between parents and children were also correlated with the level of upbringing activities of parents and the level of their own communication skills (high/medium/low). Three such indicators as the parents' understanding of the need to develop skills for interpersonal communication, adherence to ethical communication with the child, and norms of interpersonal communication among themselves in the child's presence are the criteria used to evaluate them.

Zvereva and Krotova's (2009) questionnaire for parents did not reveal anyone with a high level of upbringing activities or communication skills (0%); rather the level of upbringing activities and communication skills of the 61 parents was average (70%).

Unanimity in approach to upbringing was not always seen in such families. Both a dominating position and communication on equal terms were noted in the parents' communication with their child. Hearing-impaired children were affected not only by instructions, but also by explanations, suggestions, and requests.

Twenty-six parents (30%) had a low skill level for upbringing activities and communication skills. No unanimity between the parents was found in their upbringing. A dominating position was seen in those parents' communication with their children. The hearing-impaired children were affected mainly by instructions; requests were rarely used, and explanations almost never. The parents often pretended that they were listening to their children, rarely understood their mood, and did not take it into account when choosing activities. Also, the parents indicated that their families had no need to develop interpersonal communication skills, including norms for communicating with each other.

Thus, the predominance of middle and low levels of upbringing activities of the parents with hearing-impaired children testified to an insufficient cognitive component of parental competence, a predominant lack of understanding of the need to develop interpersonal communication skills and to observe ethical means of communicating with the children and each other.

A repeat diagnosis identified a high level of parental upbringing activities and communication skills among 24 parents (28%). Sixty-three parents (72%) had an average level of upbringing and communication skills.

In these families there was mostly unanimity between the parents in their upbringing and they communicated with their children mostly on an equal footing. They affected the hearing-impaired child through explanations and requests. Skills and norms of interpersonal communication were developed. A low level of up-

bringing and communication skills was identified in one parent (1%) at followup (in comparison with 26 parents (30%) at baseline).

Thus, the predominance of high and medium levels of upbringing activities of parents with hearing-impaired children indicates sufficient development of the cognitive component of their parental competence, the predominant understanding among the parents of the need to develop skills of interpersonal communication, and the observance of ethical communication with the children and each other.

The method of the Rokeach Value Survey (Rokeach, 1973) was used in diagnosing the value-motivational component of parental competence. According to Rokeach, "terminal values" are the ultimate goals of individual existence toward which it is necessary to strive (values-goals), and "instrumental values" are the means of achieving those goals (values-means).

The parents in our study were presented with two lists of values (18 in each list): terminal and instrumental. A rank number was assigned to each value according to its significance.

The ratings of the terminal values of parents at baseline and followup are presented in Table 1.

Table 1

Distribution of terminal values of parents with hearing-impaired children

Indicators of value	**NF at baseline**	**NF at followup**	**PF at baseline**	**PF at followup**	**φ***
Health	27	29	31	33	10.59*
Happy family life	11	14	13	16	41.34*
Interesting work	9	8	10	9	0.07
Financial security	9	8	8	7	0.07
Friends	7	6	8	7	0.08
Love	8	7	9	8	0.08
Life wisdom	6	5	7	6	0.10
Self-confidence	5	4	8	7	0.12
Happiness of others	5	4	6	5	0.12
Freedom	2	2	2	2	0.00
Total	87	87	100	100	

*Note. NF — Number of families. PF — Percentage of families. φ — Fisher test: * $p \leq 0.001$*

The results at baseline showed that the leading terminal values of parents with hearing-impaired children were: health, happy family life, interesting work, and financial security.

The triad of significant terminal values, as shown by the results at followup, still included the values of health, happy family life, and interesting work.

The values health ($\varphi^* = 10.59$ with $p = 0.000$) and happy family life ($\varphi^* = 41.34$ with $p = 0.000$) consistently led the frequency distribution (see *Table 1*). The predominance of the health value was predictable among families raising hearing-impaired children, as concern about health care was associated with its psychological and physiological characteristics.

At the same time, the significance of interesting work and financial security decreased slightly compared to the baseline stage. The low significance of freedom in the lives of these families may have been determined by excessive demands on the parents at their jobs.

The instrumental values of parents with hearing-impaired children, considered as the means to achieve the terminal values and as characterological features necessary to realize the instrumental values, are represented in Table 2:

Table 2

Distribution of instrumental values of parents with hearing-impaired children

Indicators of value	NF at baseline	NF at followup	PF at baseline	PF at followup	φ^*
Responsibility	21	14	24	16	32.51*
Honesty	13	16	15	18	36.10*
Rationality	11	11	12	13	0.00
Education	10	10	11	11	0.00
Accuracy	8	6	9	7	0.31
Tolerance	7	10	8	11	52.90*
Cheerfulness	7	5	8	6	0.36
Strong will	5	5	6	6	0.00
Sensitivity	3	8	3	9	68.60*
Independence	2	2	2	2	0.00
Total	87	87	100	100	

*Note. NF — Number of families. PF — Percentage of families. φ — Fisher test: * $p \leq 0.001$*

As for the results at baseline, a high distribution of responsibility, honesty, and rationality in the list of instrumental values proved significant in harmonizing the marital relationship. Parents with hearing-impaired children tried to create the best material conditions for them and worried about their health and well-being. The parents tried to protect their children from various difficulties, blocking their contact with people outside the family, so that the children get their life experience outside the home. The parents' behavior was characterized by constant intrusions into their children's world.

The triad of significant terminal values at followup also included responsibility, honesty, and rationality.

The values of responsibility ($\varphi^* = 32.51$ with $p = 0.000$) and honesty ($\varphi^* = 36.10$ with $p = 0.000$) consistently led the frequency distribution, although the value of

responsibility, which might be interpreted as hyper-responsibility, significantly decreased. The high rating of the value of honesty demonstrated its importance and the need to harmonize the marital relationship.

The significance of tolerance ($\varphi^* = 52.90$ with $p = 0.000$) and sensitivity ($\varphi^* = 68.60$ with $p = 0.000$) increased in statistical significance. The value of independence had low positions (2%) on the list of instrumental values (see Table 2). This might be explained by excessive demands on the parents at their jobs.

Such characteristic features as responsibility, honesty, rationality, tolerance, and sensitivity — both in the marital relationship and in the relationship between parents and children — made it possible to preserve the integrity of the family as a micro-social system, to care for their children's health, and to ensure a happy and financially secure family life and a successful career.

The questionnaire on the family's emotional relationship (Zakharova, 1997), indirectly revealing the expression of such values as sensitivity, emotional acceptance, and behavioral manifestations of emotional interaction, was used to diagnose the emotional component of parental competence.

That emotional component at baseline was characterized by the following sensitivity: parents' average ability to perceive the child's state ($X_{av} = 3.77$, $X_{cr} = 4.22 \pm 0.52$) and average empathy ($X_{av} = 3.17$, $X_{cr} = 3.39 \pm 0.59$), but insufficient understanding of the causes of that state: ($X_{av} = 2.21$, $X_{cr} = 3.85 \pm 0.65$), and insufficient emotional acceptance of the child: mother's feelings during interaction with her child ($X_{av} = 3.64$, $X_{cr} = 3.9 \pm 0.6$), unconditional acceptance ($X_{av} = 3.47$, $X_{cr} = 3.84 \pm 0.64$), acceptance of oneself as a parent ($X_{av} = 3.67$, $X_{cr} = 3.78 \pm 0.68$), predominant emotional ambiance of the interaction ($X_{av} = 3.61$, $X_{cr} = 3.66 \pm 0.66$).

In behavioral manifestations of emotional interaction, parents experienced difficulties making physical contact with their hearing-impaired children ($X_{av} = 3.22$, $X_{cr} = 4.03 \pm 0.73$) and in orienting to the child's state during the interaction ($X_{av} = 2.03$, $X_{cr} = 2.95 \pm 0.65$). However, they did affect the child's state ($X_{av} = 3.85$, $X_{cr} = 3.8 \pm 0.6$) and tried to provide the child with emotional support ($X_{av} = 3.31$, $X_{cr} = 3.47 \pm 0.67$).

The parents had no difficulty in understanding the reasons for the child's state, however (baseline: $X_{av} = 2.21$, $X_{cr} = 3.85 \pm 0.65$; followup: $X_{av} = 3.36$, $X_{cr} = 3.85 \pm 0.65$); they tried to make physical contact (baseline: $X_{av} = 3.22$, $X_{cr} = 4.03 \pm 0.73$; followup: $X_{av} = 3.7$; $X_{cr} = 4.03 \pm 0.73$) and to orient to the child's state during interactions (baseline: $X_{av} = 2.03$, $X_{cr} = 2.95 \pm 0.65$; followup: $X_{av} = 2.42$, $X_{cr} = 2.95 \pm 0.65$).

The behavioral component of parental competence at baseline and followup was diagnosed using the Stepanov Styles of Parental Behavior questionnaire (Stepanov, 2000). This instrument allowed us to identify the parents' strategy in raising their hearing-impaired children.

We defined "upbringing activities" as those determining the main lines of child-drearing personal qualities, expansion of their horizons, acquainting them with cultural values, direction and support of the children's activities.

At baseline, the predominant style of upbringing was authoritarian. Fifty-six parents (64%) displayed exactingness and strictness, and had quite clear ideas about how their children should grow up. The liberal style was found among 17 parents (20%). Those parents trusted their children, considered their weaknesses

forgivable, communicated with them on equal grounds, and were not inclined to use prohibitions and restrictions. An indifferent style was seen in 13 parents (15%). They were busy with their work, and problems of raising their children were not their priority. The children were left alone, did not receive support and help in difficult situations, made their own decisions, and resolved their own problems as well as they could. An authoritative style of upbringing was scarcely found at all (1%).

This suggests that the parents were not fully aware of their important role in the development of their child's personality and did not recognize the child's and their own right to self-development. They were not always willing to reconsider their views.

Thus the results showed that a destructive strategy of unconditional subordination of children to their parents and the children's complete dependence on them dominated in the behavioral component of parental competence. This prevented the development of an active personal identity and independence. At the same time, a significant proportion of parents was prone to the liberal and indifferent styles, in which prohibitions and followup were reduced or the children were left to themselves.

In the followup phase of the study, the frequency of the authoritarian style was statistically and significantly decreased ($\varphi^* = 19.37$ with $p = 0.000$). This style was found among 56 parents (64%) at baseline and 27 parents (31%) at followup. This confirmed a significant reduction in hyper-protection and attempts to intrude into the child's world.

The authoritative style of upbringing became significantly more prevalent ($\varphi^* = 46.88$ with $p = 0.000$); one parent (1%) was identified at baseline, and 34 parents (39%) at followup. The frequency of the liberal style significantly increased ($\varphi^* = 22.75$ with $p = 0.000$). Seventeen parents (20%) with this style were found at baseline, and 23 parents (27%) at followup.

This indicates that the parents trusted their children and communicated with them on an equal footing. The frequency of an indifferent style ($\varphi^* = 24.70$ with $p = 0.000$) significantly decreased; that style was found only among 3 parents (3%) (rather than 13 parents [15%] at baseline), which indicated that the children felt they had constant support and help in difficult situations.

Discussion

An important task in modernizing education in the Russian Federation is to ensure the accessibility of quality education, its individualization and differentiation, which implies: psychological and pedagogical assistance and support to families of at-risk children (those requiring special attention); creation of a single space for upbringing and education on the part of families as well as educational institutions, which is only possible with competent parenthood and a partnership between educational institutions and families.

A bibliometric analysis of published works with a search for "parental competence", as shown in the Russian citation index (2008 to 2017), revealed that 0.001% of the total number of studies found there were similar in nature. Brief analysis of some of them allowed us to see the following semantic contexts in the study of pa-

rental competence: stage in the development of the adult's personality and the family's resources, as a component of the relationship between parents and children; and characteristics of the relationship between parents and children, considering the children's psychological age and stage in age-related psychological development (Gorlova, 2013).

There are a small number of studies of parental competence in the upbringing of hearing-impaired children, which support the relevance of our study and its results, as well as its scientific novelty. This is also indicated by the results of the bibliometric analysis performed at elibrary.ru with the search phrase: "parental competence in the upbringing of hearing-impaired children". The bibliometric analysis covered 10 years (2008 to 2017), and the total number of publications was two, which corresponds to 7,034 — 6%.

We considered the process of modeling the interaction between families with hearing-impaired children and special needs educational institutions as analogous to modeling the interaction of families with a primary school. The process consisted of three blocks: purpose-oriented, organizational-substantive, and evaluation of effectiveness (Kovalenko, 2016; Sergeeva & Arakelyan, 2016).

Our results in the evaluation of parent competence (its cognitive, value-motivational emotional, and behavioral components) with the psychological and pedagogical support of families raising hearing-impaired children were consistent with the results of studies related to the interaction of families and the primary school (Kovalenko, 2016; Sergeeva & Arakelyan, 2016).

In this regard, the presence of general trends and patterns in the evaluation of effectiveness can be the basis for developing a model of the interaction between families and schools in an inclusive educational environment.

However, comparing our results with those of Kovalenko (2016) shows some discrepancies in the content of the value-motivational, as well as behavioral, components of parental competence when raising hearing-impaired children, in the followup phase of our formative study.

Thus in the value-motivational component of parental competence, the dominant values, in addition to those of "health", "happy family life", and "interesting work", which are typical for parents of primary schoolchildren, were the values "financial security", "honesty", and "rationality".

According to our study, the parents were trying to create the best material conditions for their children, and feared for their health and well-being. The parents raising hearing-impaired children showed responsibility, honesty, rationality, tolerance, and sensitivity in the marital relationship and the relationship between parents and children, in order to preserve the integrity of the family and ensure happy, financially secure family life.

The predominance of the authoritarian style, in addition to the authoritative style of child-rearing, was noted in the behavioral component of parental competence, manifesting exactingness and strictness.

In our view, these values in parental competence in raising hearing-impaired children are due to the psychological and physiological characteristics of the children and hyper-responsibility for their life and health.

Conclusion

At baseline, the study of the psychological content of parental competence in the psychological and pedagogical support of families raising hearing-impaired children made it possible to draw the following conclusions.

First, parental competence was represented at the cognitive, value-motivational, emotional, and behavioral levels by the set of interrelated qualities of the parent's personality necessary for effective child-rearing. The structure of parental competence allowed us to determine the content, identify the forms of effective work on its development in accordance with the assigned tasks, the professional competencies of specialists, and the planned results.

Second, regarding the cognitive component of parental competence, we identified the predominant underestimation by the parents of the need to develop interpersonal communication skills, and to observe ethical communication with the children and each other. The emotional component of parental competence was represented by their difficulties in understanding the causes of the children's emotional state, orienting towards their emotional state during the interaction, and making physical contact.

The values of health and happy family life, as well as those of professional self-realization, were predominant in the value-motivational component of parental competence. The behavioral component of parental competence was dominated by the authoritarian style: exactingness and strictness.

The results of the followup stage of our study allow us to draw the following conclusions:

First, the psychological content of parental competence has significantly changed compared to baseline.

The cognitive component of parental competence is characterized by predominant unanimity in the upbringing of hearing-impaired children and a relationship of partnership in communication with them.

The emotional component of parental competence is represented by the absence of difficulties in understanding the causes of the children's emotional state, an orientation towards their emotional state in interaction with them, and making physical contact.

The values of health, happy family life, interesting work, and financial security, honesty, and rationality are still dominant in the value-motivational component of parental competence. The value of health is predominant. The importance of hyper-responsibility is significantly reduced.

The behavioral component of parental competence is represented by the predominance of the authoritative style of upbringing. Parents perform their important role in the development of the personality of children and recognize the right of the children to self-development. At the same time, the authoritarian style is still significant.

Second, the developed and tested model of the psychological and pedagogical support of families raising hearing-impaired children is effective, since the results demonstrate an increase in parental competence.

The information in this article can be useful to specialists at special needs educational institutions, as well as to parents who are raising hearing-impaired children, and may be included in the process of psychological and pedagogical support.

The results of the study make it possible to identify problems and prospective directions that require further study. Among these are the following: development of competence on the part of parents with hearing-impaired children, and developing scientific and methodological support to provide parents with psychological and pedagogical support.

Limitations

The main limitation of the study is that its participants were exclusively parents of hearing-impaired children in one Russian city and one special needs educational institution, located in Kursk. Care should therefore be taken in generalizing the results to non-participant parents of hearing-impaired children. This limitation can be overcome by studying parents of hearing-impaired children from several cities.

Another limitation is that the study did not consider the relationship between siblings, which could provide more details about the families functioning as a micro-social system and that system's impact on parental competence.

In addition, the corporate culture of the special needs educational institution, the level of education and personal characteristics of its employees, and the relationships among them could affect the development of parental competence and should also be considered.

The reliability of the results and conclusions of the study has been demonstrated by various mathematical statistics methods and comprehensive analysis of the data.

References

Asberg, K.K., Vogel, J.J., & Bowers, C.A. (2008). Exploring correlates and predictors of stress in parents of children who are deaf: Implications of perceived social support and mode of communication. *Journal of Child and Family Studies, 17,* 486–499. https://doi.org/10.1007/s10826-007-9169-7

Calderon, R. (2000). Parental involvement in deaf children's education programs as a predictor of child's language, early reading, and social emotional development. *Journal of Deaf Studies and Deaf Education, 5*(2) Spring, 140–155. https://doi.org/10.1093/deafed/5.2.140

Cere, D. (2013). What is parenthood? Proposing two models. Toward an integrative account of parenthood. In L.C. McClain & D. Cere (Eds.), *What is parenthood? Contemporary debates about the family* (pp. 19–40). New York: New York University Press. https://doi.org/10.18574/nyu/9780814729151.003.0001

Gladkova, Yu.A. (2009). Deyatel'nost' doshkol'nogo uchrezhdeniya po povysheniyu psikhologo-pedagogicheskoi kul'tury sovremennoi sem'i: avtoref. dis. …kand. ped. nauk. [Activities of preschool institutions to enhance the psychological and pedagogical culture of the modern family. Author's abstract. dis. ... candidate of pedagogical sciences]. Moscow: Moscow City Pedagogical University.

Glăveanu, S.M. (2015). The parental competence of single-parent families from vulnerable groups. *Procedia - Social and Behavioral Sciences, 187*(13), 201–205. https://doi.org/10.1016/j.sbspro.2015.03.038

Gorlova, E.L. (2010). Roditel'skaya kompetentnost': podkhody k izucheniyu i razvitiyu [Parental competence: Approaches to learning and development]. *Vestnik RGGU. Seriya: Psikhologiya. Pedagogika. Obrazovanie* [Bulletin of the RSUH. Series: Psychology. Pedagogy. Education], *17*(60), 214–224.

Gorlova, E.L. (2013). Psikhologicheskii vozrast rebenka kak osnovanie dlya proektirovaniya strategii razvitiya roditel'skoi kompetentnosti [Psychological age of the child as a basis for designing strategies for the development of parental competence]. *Natsional'nyi psikhologicheskii zhurnal.* [National Psychological Journal], *2*(10), 134–141. https://doi.org/10.11621/npj.2013.0217

Hintermair, M. (2006). Parental resources, parental stress, and socioemotional development of deaf and hard of hearing children. *Journal of Deaf Studies and Deaf Education, 11*(4), 493–513. https://doi.org/10.1093/deafed/enl005

Kabanchenko, E.A., & Kholodkova, O.G. (2017). K probleme roditel'skoi kompetentnosti v sem'yakh, vospityvayushchikh detei s OVZ [On the problem of parental competence in families raising children with disabilities]. In G.Yu. Lizunova (Eds.), *Tsennostnye orientatsii molodezhi v usloviyakh modernizatsii sovremennogo obshchestva Sbornik nauchnykh trudov.* [Value orientations of youth in conditions of modernization of modern society] (pp. 309–312). Gorno-Altaisk: Gorno-Altaisk State University.

Kliukvina, A.Yu. (2015). Teoreticheskie aspekty issledovaniya problemy kompetentnosti roditelei, imeyushchikh detei doshkol'nogo i mladshego shkol'nogo vozrasta s intellektual'noi nedostatochnost'yu [Theoretical aspects of the study of the problem of the competence of parents with children of preschool and primary school age with intellectual impairment]. *Vestnik Moskovskogo gorodskogo pedagogicheskogo universiteta. Seriya: Pedagogika i psikhologiya* [Bulletin of Moscow City Pedagogical University. Series: Pedagogy and psychology], *1*(31), 86–95.

Koester, L., & Meadow-Orlans, K. (1990). Parenting a deaf child: Stress, strength, and support. In D. Koester & K. Meadow-Orlans (Eds.), *Educational and developmental aspects of deafness* (pp. 299–320). Washington, DC: Gallaudet University Press.

Koester, L.S., Papoušek, H., & S. Smith-Gray (2000). Intuitive parenting, communication, and interaction with deaf infants. In P.E. Spencer, C.J. Erting, & M. Marschark (Eds.), *Essays in honor of Kathryn P. Meadow-Orlans: The deaf child in the family and at school* (pp. 55–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Korobkova, V.V. (2011). Sushchnost' i soderzhanie psikhologo-pedagogicheskoi kompetentnosti roditelei v sovremennykh issledovaniyakh [The essence and content of parents' psychological and pedagogical competence in contemporary research]. *Pedagogicheskoe obrazovanie i nauka* [Pedagogical education and science], *6,* 101–105.

Kovalenko, T.V. (2016). Modelirovanie vzaimodeistviya sem'i i shkoly v usloviyakh sovremennogo nachal'nogo obrazovaniya: avtoref. dis. …kand. ped. nauk. [Modeling the interaction between family and school in the context of modern primary education. Author's abstract. dis. ... candidate of pedagogical sciences]. Omsk: Omsk State Pedagogical University.

Kurtzer-White, & E., Luterman, D. (2003). Families and children with hearing loss: Grief and coping. *Mental Retardation and Developmental Disabilities Research Reviews, 9*(4), 232–235. https://doi.org/10.1002/mrdd.10085

Léonard, N., & Paul, D. (1995). La vie du couple et les sentiments de compétence parentale [The life of the couple and feelings of parental competence]. *Can. Nurse, 91*(9), 42–45.

Minina, A.V. (2013). Struktura i soderzhanie pedagogicheskoi kompetentnosti roditelei v vospitanii samostoyatel'nosti u detei doshkol'nogo vozrasta [Structure and content of pedagogical competence of parents in the education of independence in preschool children].

Obrazovanie. Nauka. Innovatsii: Yuzhnoe izmerenie [Education. Science. Innovations: The Southern Dimension]. Rostov-on-Don: Southern Federal University, *2,* 93–98.

Misina, N.N. (2009). Roditel'skaya kompetentnost': psikhologicheskii aspekt problemy [Parental competence: the psychological aspect of the problem]. *Nauchnykh trudov Severo-Kavkazskogo gosudarstvennogo tekhnicheskogo universiteta. Gumanitarnye nauki* [Scientific papers of the North Caucasus State Technical University. Humanitarian sciences], *7*. Stavropol: North Caucasus State Technical University. Retrieved from: http://science.ncstu.ru/articles/hs/2009_07/psychology.

Molchanova, L.N., & Sitnikova, A.V. (2015). Osobennosti funktsionirovaniya semei, vospityvayushchikh detei s narusheniem slukha [Features of the functioning of families raising children with hearing impairment]. *Korrektsionno-pedagogicheskoe obrazovanie* [Correction and pedagogical education], *3*(3), 16–27.

Ovcharova, R.V. (2006). *Roditel'stvo kak psikhologicheskii fenomen* [Parenthood as a psychological phenomenon]. Textbook. Moscow: Moscow Psychological and Social Institute.

Petukhova, A.V. (2012). Povyshenie urovnya kompetentnosti semei, imeyushchikh detei s rechevymi narusheniyami cherez integratsiyu raboty uchitelya-logopeda i pedagoga-psikhologa [Raising the level of competence of families of children with speech impairments through the integration of the work of a speech therapist and a psychologist]. *Nauchnye trudy SWorld* [Scientific works SWorld], *17,* 3–5.

Pipp-Siegel, S., Sedey, A.L., & Yoshinaga-Itano, C. (2002). Predictors of parental stress in mothers of young children with hearing loss. *Journal of Deaf Studies and Deaf Education, 7*(1), 1–17. https://doi.org/10.1093/deafed/7.1.1

Popova, S.Yu., & Nogovitsyna, N.M. (2017). Rol' obshcheobrazovatel'nykh uchrezhdenii v povyshenii roditel'skoi kompetentnosti [The role of general education institutions in improving parental competence]. *Nauchno-metodicheskii elektronnyi zhurnal Kontsept* [Concept. Scientific and methodical electronic journal], *26,* 192–194.

Quittner, A.L., Glueckauf, R.L., & Jackson, D.N. (1990). Chronic parenting stress: Moderating versus mediating effects of social support. *Journal of Personality and Social Psychology, 59,* 1266–1278. https://doi.org/10.1037/0022-3514.59.6.1266

Rokeach, M. (1973). *The nature of human values.* New York: The Free Press. Russian version: Metodika Rokicha "Tsennostnye orientatsii". Retrieved from: http://psycabi.net/testy/320-metodika-rokicha-tsennostnye-orientatsii-test-miltona-rokicha-issledovanie-tsennostnykh-orientatsij-m-rokicha-oprosnik-tsennosti-po-rokichu.html.

Selina, V.V. (2009). Razvitie pedagogicheskoi kompetentnosti roditelei detei rannego vozrasta v doshkol'nom obrazovatel'nom uchrezhdenii: dis. … kand. ped. nauk [Development of the pedagogical competence of parents of young children in preschool educational institutions: dis. … candidate of pedagogical sciences]. Novgorod: Novgorod State University.

Sergeeva, B.V., & Arakelyan, R.K. (2016). Formirovanie roditel'skoi kompetentnosti v usloviyakh nachal'noi shkoly [Formation of parental competence in the conditions of an elementary school]. *Istoricheskaya i sotsial'no-obrazovatel'naya mysl'* [Historical and social-educational thought], *8, 1*(1), 184–188. https://doi.org/10.17748/2075-9908-2016-8-1/1-184-188.

Shagraeva, O.A. (2010). Kategoriya "psikhologicheskaya kompetentnost' roditelei" v svete osnovnykh polozhenii kontseptsii P.Ya. Gal'perina [The category "psychological competence of parents" in light of the main provisions of the concept of P.Ya. Galperin]. *Psikhologiya i pedagogika: metodika i problemy prakticheskogo primeneniya* [Psychology and pedagogy: methods and problems of practical application], *16*(2), 264–271.

Sidorenko, E.V. (2006). *Metody matematicheskoi obrabotki v psikhologii* [Methods of mathematical processing in psychology]. St. Petersburg: Rech'.

Stepanov, S. (2000). Oprosnik S. Stepanova "Stili roditel'skogo povedeniya" [S. Stepanov "Styles of parental behavior" questionnaire]. Retrieved from: http://skorova.spb.ru/?mode=test19.

Teti, D.M. & Candelaria, M.A. (2002). Parenting competence. In Marc H. Bornstein (Ed.), *Handbook of parenting* (pp. 149 –180). Mahwah, NJ: Lawrence Erlbaum Associates.

Topilina, N.V. (2016). Pedagogicheskie sposobnosti i kompetentsii roditelei v sovremennoi pedagogicheskoi praktike [Pedagogical abilities and competence of parents in modern pedagogical practice]. *Innovatsionnaya nauka* [Innovative science], *2–4,* 136–139.

Vaccari, C., & Marschark, M. (1997). Communication between parents and deaf children: Implications for social-emotional development. *Journal of Child Psychology and Psychiatry, 38*(7), 793–801. https://doi.org/10.1111/j.1469-7610.1997.tb01597.x

Vlaykova, K.V. (2010). Vklyuchenie roditelei v korrektsionno-razvivayushchii protsess kak odna iz form raboty po povysheniyu roditel'skoi kompetentnosti [Inclusion of parents in the correctional-development process as one of the forms of work to improve parental competence]. *Vospitanie i obuchenie detei s narusheniyami razvitiya* [Education and training of children with developmental disabilities], *4,* 55–57.

Zakharova, E.I. (1997). Oprosnik emotsional'nykh otnoshenii v sem'e [Questionnaire of emotional relations in the family]. Retrieved from: https://pro-psixology.ru/psixologicheskoe-obsledovanie-semi/606-oprosnik-yemocionalnyx-otnoshenij-v-seme-ei.html

Zvereva, O.L., & Krotova, T.V. (2009). *Obshchenie pedagoga s roditelyami v DOU: Metodicheskii aspekt* [Questionnaire for parents. Communication of the teacher with parents in preschool: Methodological aspect]. Moscow: Sfera.

Original manuscript received August 04, 2018
Revised manuscript accepted September 30, 2018
First published online December 30, 2018

Russian Psychological Society
Lomonosov Moscow State University

Panel of Referees

We express our sincere gratitude to our reviewers who contributed to the quality of our publications in 2018.

Adriana del Palacio Gonzalez, Aarhus University, Denmark
Aleksander N. Poddiakov, National Research University Higher School of Economics, Russia
Aleksander N.Tatarko, National Research University Higher School of Economics, Russia
Anatoly A. Skvortsov, Lomonosov Moscow State University, Russia
Andrew McNeill, Northumbria University, United Kingdom
Andrey D. Nasledov, St. Petersburg State University, Russia
Andrey K. Krylov, Institute of Psychology of Russian Academy of Sciences, Russia
Andrey L. Ryzhov, Lomonosov Moscow State University, Russia
Anna H. Fam, National Research University Higher School of Economics, Russia
Anna Sircova, Time Perspective Network, Copenhagen, Denmark
Anna V. Leybina, Lomonosov Moscow State University, Russia
Boris B. Velichkovskiy, Lomonosov Moscow State University, Russia
Boris M. Velichkovskiy, Institute of Cognitive Studies "Kurchatov Institute", Russia
Ekaterina Kuzmina, University of Oslo, Norwey
Ekaterina Yu. Matveeva, Lomonosov Moscow State University, Russia
Elena A. Sergienko, Institute of Psychology of Russian Academy of Sciences, Russia
Elena I. Pervichko, Lomonosov Moscow State University, Russia
Elena I. Rasskazova, Lomonosov Moscow State University, Russia
Elena L. Grigorenko, University of Houston, USA
Elena R. Agadullina, National Research University Higher School of Economics, Russia
Elizaveta Berezina, Sunway University, Malaysia
Evgeny N. Osin, National Research University Higher School of Economics, Russia
Irina A. Baeva, Herzen State Pedagogical University, Russia
Katerina N. Polivanova, National Research University Higher School of Economics, Russia

Lucía Jiménez University of Seville, Spain
Marina K. Bardyshevskaya, Lomonosov Moscow State University, Russia
Marina S. Luzyanina, Landmark Center, Russia
Martin F. Linch, University of Rochester, USA
Mia Silfver-Kuhalampi, University of Helsinki, Finland
Michael Cole, University of California, USA
Natalia A. Rozhdestvenskaya, Lomonosov Moscow State University, Russia
Nataliya A. Varako, Lomonosov Moscow State University, Russia
Oleg A. Sychev, Shukshin Altai State Humanities Pedagogical University, Russia
Rifkat J. Muhamedrahimov, St. Petersburg State University, Russia
Seger Handoyo, Airlangga University, Indonesia
Sergey B. Malykh, Psychological Institute of the Russian Academy of Education, Russia
Sofya K. Nartova-Bochaver, National Research University Higher School of Economics, Russia
Stanislav A. Kozlovsky, Lomonosov Moscow State University, Russia
Tamara O. Gordeeva, Lomonosov Moscow State University, Russia
Tatyana V. Akhutina, Lomonosov Moscow State University, Russia
Tsuneyuki Abe, Tohoku University, Japan
Yuriy V. Karpov, Touro College, USA