



Measuring Visual, Spatial, and Visual Spatial Short-Term Memory in Schoolchildren: Studying the Influence of Demographic Factors and Regression-Based Normative Data

A. Toornstra¹ · P. P. M. Hurks¹ · W. Van der Elst² · G. Kok³ · L. M. G. Curfs⁴

Received: 5 May 2019 / Revised: 8 August 2019 / Accepted: 21 August 2019 / Published online: 11 September 2019
© The Author(s) 2019

Abstract

The study aims to establish demographically corrected norms for three computerized tasks measuring different aspects of visual short-term memory (VSTM) in Ukrainian schoolchildren. These tasks measure respectively visual STM (the Pattern Recognition Memory (PRM) test), spatial STM (the Spatial Span (SSP) task), and visual spatial STM (the Paired Associates Learning (PAL) task). All tasks were administered to $n = 186$ children aged 5.10 years to 14.5 years old to evaluate the influence of demographic variables. Relevant demographic factors that influence task scores (VSTM), i.e., age and level of parental education, are identified and in keeping with the current literature. No sex differences were found. Based on these data, regression-based, demographically corrected norms were established per task. This approach to constructing norms differs from how (worldwide) PRM, SSP, and/or PAL norms have been constructed traditionally. In the latter approach, norms are calculated for each age group separately and without correcting for level of parental education, whereas in the regression-based normative method, multiple regression models are used to compute the expected test scores of an individual (rather than the subgroup means that are used in the traditional approach). Consequently, the regression-based norms for the PRM, SSP, and PAL presented in this paper are individualized, taking into account the unique characteristics of the individual that is tested on these tasks. Last, the confidence intervals of the PRM scores of the Ukrainian schoolchildren and the western norm group largely overlapped, except for the youngest age group, which adds to the literature about cultural effects on cognition.

Keywords Visual (spatial) memory · Demographic influences · Continuous norms · Schoolchildren · Ukraine

Introduction

The Cambridge Neuropsychological Test Automated Battery (CANTAB[®]; Cambridge Cognition 2012) is a test battery

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40817-019-00070-6>) contains supplementary material, which is available to authorized users.

✉ A. Toornstra
annet.toornstra@maastrichtuniversity.nl

¹ Faculty of Psychology and Neuroscience, Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, The Netherlands

² Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Hasselt, Belgium

³ Department of Social, Work and Organizational Psychology, Maastricht University, Maastricht, Netherlands

⁴ Maastricht University Medical Center, Maastricht, Netherlands

frequently used in both clinical and behavioral research. Three CANTAB[®] subtasks, i.e., the Pattern Recognition Memory (PRM) task, the Spatial Span (SSP) task, and Paired Associates Learning (PAL) task, are most often used to evaluate visual short-term memory (VSTM) in children. VSTM is typically defined as a storage system that holds limited amounts of information in mind for a brief period of time (Baddeley et al. 2011). VSTM is required in learning (Kessels et al. 2000). Hitch et al. (2001) have shown that indeed VSTM is a predictor of success in school. Separate brain mechanisms are believed to be involved in the short-term storage of information from different modalities, i.e., within the domain of VSTM (e.g., visual, versus spatial information; Smith et al. 1996; Bigelow and Poremba 2012). For instance, neurobiological research shows a ventral pathway in the brain, which is more concerned with processing and storing of colors and shapes (i.e., the features) of an object, and a dorsal pathway, which is involved in the processing and storage of locations of objects and/or their spatial relations (Gruhn and Rauscher 2002; Kravitz et al. 2013). The ventral

(parvocellular) pathway is thought to involve the primary visual cortex and temporal brain areas associated with short- and long-term memory, emotions, reward, and learning (Kravitz et al. 2013). The dorsal (magnocellular) pathway is thought to be an occipital parietal network, which lies between the primary visual cortex and subcortical areas; is associated with, e.g., visually guided action (including the (pre)motor cortex) and spatial navigation; and is involved in processing random and possibly dynamic relations between objects (i.e., their locations) (Kravitz et al. 2013). Simultaneously, features of an object (ventral processing) are thought to be, to some degree, incorporated in dorsal temporal processes to, e.g., guide decisions (i.e., not grasping a prickly-looking cactus). Indeed, the three CANTAB[®] subtasks, i.e., the PRM, the SSP, and the PAL, measure different aspects of VSTM: visual memory, spatial memory, and visual spatial memory, respectively (see below for a detailed description of the tasks). In this study, the CANTAB[®] and specifically these subtasks were used to test these different components of nonverbal VSTM.

The PRM tests recall of abstract-colored patterns. The participant needs to identify the pattern that was previously shown (i.e., the target pattern) in a series of patterns (for an example, see Fig. 1 in the “Methods” section). The SSP is different from the PRM in that it requires the child to remember and to reproduce a *location* of the target and not the (visual) pattern of the target (spatial VSTM; for an example, see Fig. 2). Finally, in the PAL, the child needs to remember and reproduce both the pattern and location of the target (see Fig. 3). Administering and analyzing responses on the three tasks allows for a more complete understanding of brief VSTM in children; e.g., remembering a pattern is different from remembering a location.

These tests, however, also implicate other cognitive functions associated with visual short-term memory, e.g., a general executive (Baddeley et al. 2011). Diamond (2013) defined executive functions (EFs) as a set of independent yet interconnected cognitive skills that, among others, support one to meet demands of novel tasks, consider concepts from several perspectives, and plan and execute solutions. Attention is an EF thought to underlie all other cognitive functions, and maturation of these EFs has been associated with the development of prefrontal brain areas (Anderson 2002). Alloway et al. (2006) found that domain general processes, e.g., the employment of

attention, were more active in 4–6-year-old children in spatial VSTM tasks, which was thought to be related to the relative immaturity of these general cognitive functions. Processes involving more than brief storage, e.g., the manipulation of information in the service of complex cognitive behaviors such as comparing and updating information, are also theorized to be working memory (Fuster 2002). It is beyond the scope of this article to discuss in-depth visual information processing and STM and/or working memory.

Importantly, results obtained on tests such as the PRM, the SSP, and/or the PAL are only meaningful when compared to appropriate norm data for a sample of schoolchildren (Lezak et al. 2012). In the current study, normative data for these CANTAB[®] subtests were prepared for children living in Ukraine. Even though normative data for these CANTAB[®] subtests have been published for schoolchildren living in, among others, countries such as the USA, Australia, and Finland (De Luca et al. 2003; Lehto et al. 2003; Luciana and Nelson 2002), this is not the case for a Ukrainian pediatric sample.

It is relevant to collect normative data for these tests for schoolchildren living in different countries, since cognitive functions, in turn, have been associated with environmental factors such as culture. Thinking, knowledge, values, and beliefs are thought to constitute culture and are formed in social situations (Ardilla 2005; Ardila et al. 2005). These variances in social and cultural contexts are linked to geographical differences. Ukraine, with its long and complex history, may share cultural characteristics with other Slavic and/or former USSR countries (Subtelny 2009). Variances in test performance could, for instance, arise due to different levels of familiarity with testing, including nonverbal tests (Ardila 1995). Comparison of test scores of, for example, Ukrainian children against western norms (e.g., Western European and American norms as is mostly the case with the CANTAB[®]) may help to identify differences and similarities between these two groups and is therefore included in this study [in line with, e.g., Zasenکو et al. 2017 and Toornstra et al. 2019].

As mentioned above, results of these tasks are only relevant when compared to suitable norm data (Lezak et al. 2012). Norm data presently available internationally for these CANTAB[®] subtests include several limits. These norms tend to be calculated based on means and standard deviations for

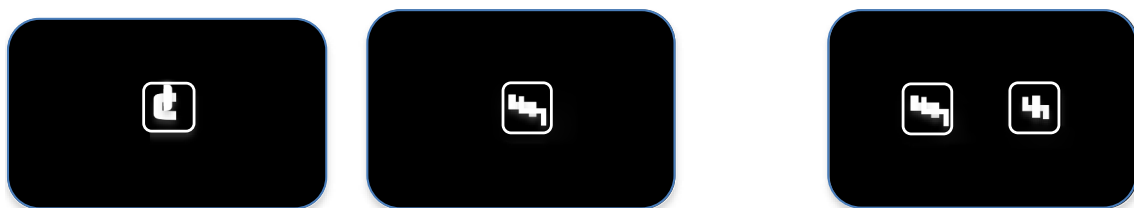
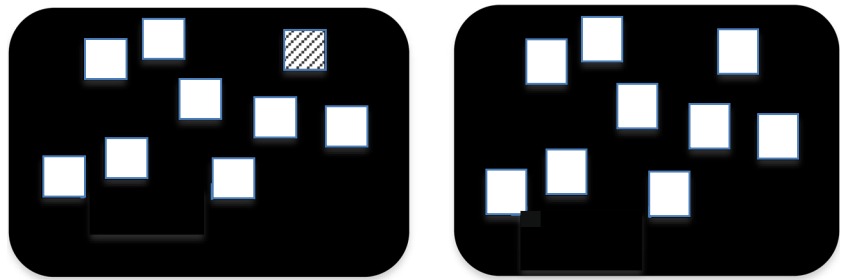


Fig. 1 The PRM on a touch screen. One series of 12 single patterns is followed by a forced choice recognition of the target patterns (this is not a real but similar CANTAB[®] PRM sample)

Fig. 2 The SSP on a touch screen. Showing one box that has changed color in this example trial of 3 boxes in total (this is not a real CANTAB[®] SSP sample)



relevant subgroups of children discretely, e.g., per 2-year age band (De Luca et al. 2003; Lehto et al. 2003; Luciana and Nelson 2002). Including age as a demographic factor in establishing norms is consistent with research. For example, multiple studies have shown that performances on numerous cognitive tests are influenced by the demographic variable “age” (Lezak et al. 2012). There are, to the best of our knowledge, relatively fewer studies on the PRM, SSP, and PAL tasks in children (De Luca et al. 2003; Luciana and Nelson 2002; Robbins et al. 1994; Sahakian and Owen 1992). Studies in children using these CANTAB[®] and comparable subtasks have shown that VSTM can be measured in children aged 5 years and older and age-related improvements can be observed with these tasks (Davidson et al. 2006; De Luca et al. 2003; Green et al. 2019; Luciana and Nelson 2002). In addition, performances on (spatial) VSTM tasks, comparable to the PRM, SSP, and PAL tests, have been shown to increase with age, i.e., from early preschool years well into middle school (De Luca et al. 2003; Luciana and Nelson 2002).

However, calculating norms based on means and standard deviations for relevant subgroups of children separately, e.g., per 2-year age bands, has its limitations (see, e.g., Van der Elst et al. 2011). In calculating normative data, the total sample needs to be subdivided into several different subgroups based on, e.g., age. Four examples of the disadvantages of making these subgroups in the data are that (1) test scores are only applied to one particular (age) subgroup in the sample while (2) children closely in age (i.e., 1 month apart; e.g., 6.9 years versus 7.0 years old) may be in different age subgroups. Additionally, (3) the (age) subgroups themselves may be subject to unusual sample characteristics but result in a norm for that (age) subgroup.

Another (4) limitation is that often other demographic characteristics (e.g., not only age but also, e.g., sex and the level of parental education (LPE)) are not included, or considered,

while constructing norms. Including these demographic characteristics next to age, in the traditional way of norming described above, would result in even more subgroups. These other demographic characteristics have been associated with inter-individual differences in cognitive neuropsychological test performances in, e.g., adults (and participant’s own level of education) including visual spatial memory (Heaton et al. 2003; Norman et al. 2011). For instance, Voyer et al. (2017) conducted a meta-analysis regarding the effects of sex differences on performances on various (spatial) VSTM tests in 3-year- to 86-year-old participants. They included, in total, 180 studies and found, overall, a small but significant effect in favor of males. However, when distinguishing between VSTM (pattern recognition as in the PRM) and memory for location (as seen in the SSP and PAL tasks), they found an effect in favor of females on tasks concerning memory for location. Furthermore, the demographic variable LPE has been studied relatively little in relation to VSTM in schoolchildren. LPE is perceived as an approximation of socioeconomic status (SES) (Davis-Kean 2005), and higher levels of both have been positively linked to an increase in children’s cognitive functions in general and among others to VSTM specifically (Hackman et al. 2014; Kaplan et al. 2001; Noble et al. 2007). In addition, Noble et al. (2015) found, in their large sample (1099 participants from 3 to 20 years old), an association between SES and differences in children’s brain structures (e.g., brain structures associated with memory performance and spatial skills). This showed especially when comparing the brains of children in the lower SES group to children from the other SES groups (adjusted for age and genetic factors). Differences in the environment were hypothesized to be, among others, family stress, limitations in cognitive stimulation and nutrition, and exposure to toxins of which children in lower SES environments might experience more. Taking into account these unique characteristics (i.e.,

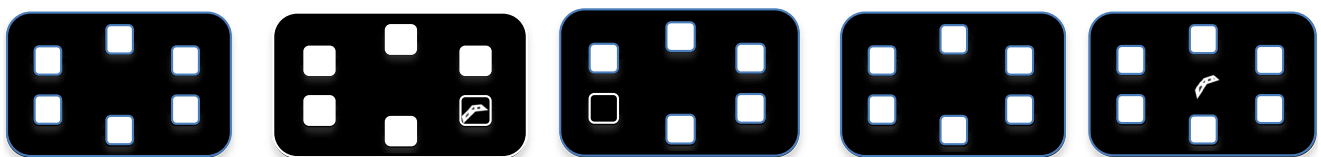


Fig. 3 The PAL on a touch screen. A trial with a target pattern in the middle. The PAL increases in difficulty to 8 patterns (this is not a real CANTAB[®] PAL sample)

including age, but also sex and LPE) of the individual who is tested on the tasks would make the norms more individualized in our opinion.

To circumvent the disadvantages of traditional norming methods mentioned above, we will employ in our study an alternative and promising method called continuous norming to collect PRM, SSP, and PAL normative data (in line with, e.g., Van Breukelen and Vlaeyen 2005; Van der Elst et al. 2011; Zachary and Gorsuch 1985). This different method of norming is based on multiple regressions while weighing several demographic variables, such as age, sex, and LPE. In these regression analyses, both continuous variables (e.g., age in years) and categorical variables (e.g., sex or LPE) can be simultaneously incorporated, without creating subgroups for each variable (Van der Elst et al. 2011). Using age as a continuous variable means that the relation between age and the predicted score is calculated from all data points in the whole sample (i.e., combining children from all age groups in one analysis). This results in more refined estimates of age-based norm scores (Bechger et al. 2009). Additionally, the possibility of incorporating more than one demographic variable in regression models makes this normative method rather individualized, as mentioned above (Van der Elst et al. 2011). These findings underline the relevance to study the effects of the demographic variables age, sex, and LPE, and establish the normal range and norm data for the PRM, the SSP, and the PAL for this sample of Ukrainian schoolchildren ($n = 186$) aged 5.10 years to 14.5 years old.

Methods

Participants

The sample comprised of $n = 186$ children enrolled in two local primary and middle schools in Ukraine. Per school, all children of these schools were asked and 80.5% took part. Characteristics of this sample are displayed in Table 1. Age was defined as a continuous variable ranging in age from 5.10 to 14.5 years. The level of parental education was established as the level of education completed by the parent(s) or guardians in a household; low was defined as education up till grade 9 (primary and middle school) and high was from grade 10 (high school) onward, in line with the Ukrainian educational

system (Ukraine Channel 2017). Education is compulsory up until and including grade 9. This is akin to western ages of compulsory education, i.e., 16 years old. The highest level of education was assigned as the mean for the household in the 6.5% of the cases when there was a difference (high versus low) between parents' (or caregivers') level of education (as in, e.g., Koball and Jiang 2018). Attending school beyond the compulsory years is, in general in Ukraine but even more in these rural areas, linked to socioeconomic status. High school and further education require children to travel further to larger institutes, which necessitates families to have or access means to do so. The kind of schooling parents attended has been linked to test scores in children (Ardila et al. 2005).

Procedure

Researchers informed schools through word of mouth in the area about the study. School management invited the parents to information meetings about the study. Information and consent letters were provided by school management and then taken home by the parents. Testing started after parents and children had provided their teachers with the informed consent letters. Testing was carried out during school hours, and no compensation was provided.

As part of the CANTAB[®] (Cambridge Cognition 2012), the PRM, the SSP, and the PAL were administered individually in a separate room according to the prescribed standardized procedure. The tests were administered on a HP Pavilion TS sleek-book laptop with a 15-in. touch screen. Two CANTAB[®]-trained researchers administered the PRM, the SSP, and the PAL, applying the predescribed script and practice items in the children's native language. Tests and outcome measures are described below.

The research ethics committee of the Faculty of [name and place of university] approved this study. All data were collected according to ethics regulations of the World Medical Association Declaration of Helsinki (2013). In line with these ethics regulations, the first researcher gave each child and his/her parents/guardians a brief individual report (stating the individual's number of items remembered, e.g., an encouragement to remember one more item over the next month and some general ideas to nurture brain development), while management received anonymous summaries of class and school achievements.

Table 1 Demographic data of the sample

	<i>n</i>	Age (mean in years)	Age SD	Sex		Level of parental education (LPE)	
				Female	Male	Low	High
Total sample	186	9.63	2.49	90	96	23	163

Instrument

The PRM is a VSTM task (Cambridge Cognition 2012). Validity in children was established, among others, in schoolchildren aged 4–12 years old (Luciana and Nelson 2002; Teixeira et al. 2011). Furthermore, a moderate test-retest reliability (.64 to .84) was found on CANTAB® tasks (including a test comparable to the PRM, the SSP, and the PAL; Lowe and Rabbitt 1998). For the PRM task, the procedure is as follows (Fig. 1): First, abstract-colored target patterns appear in the middle of the screen, one at a time on a black background. After a series is finished, two patterns (the target and a distractor) appear on the screen. The participant needs to identify which of these two patterns was part of the previously presented target pattern, by touching it on the screen. Auditory and visual feedback is provided to indicate a(n) (in)correct response.

The outcome measure PRM reports accuracy, i.e., how many patterns were recognized correctly. The participant completes two series of twelve trials in total, making the maximum accuracy score 24 correctly identified patterns. Total administration time is, on average, 4.5 min (Cambridge Cognition 2012). In general, in these three CANTAB® tasks, touching boxes outside the designated squares or touching them twice (illegal responses) is not registered but only the first touch of a specific box.

The SSP is a computerized version of the Corsi Block-Tapping task (see Milner 1971), which is used to assess spatial (including temporal) VSTM. Support for the validity of the SSP in pediatric samples was, for example, found by Lehto et al. (2003) and Luciana and Nelson (2002). The procedure of this task is as follows (Fig. 2): A number of randomly located boxes change into a particular color one after the other and then turn white again. The participant's task is to remember the location and exact order of boxes that changed color. He or she needs to touch the previously colored boxes in the same order shown earlier, after which the trial is finished.

Each trial starts on a new black screen. Task difficulty rises with increasing span length by one block at a time, ranging from two to nine boxes. There are several difficulty levels depending on the number of boxes that change color. Per difficulty level, e.g., two boxes, subjects have three attempts to repeat the previously shown string of two. The test terminates in case of three erroneous responses at a certain difficulty level and continues to the next level after a correct response is completed. A trial is scored as correct if all boxes have been touched in the exact same order as previously shown. The outcome measure of the SSP is the total amount of cubes correctly located on the first attempt (i.e., first trial out of the possible three attempts) over all difficulty levels. The maximum SSP score is 44 boxes correct (i.e., 8 trials, starting at 2 boxes and increasing with one box each). The total administration time is about 5.5 min (Cambridge Cognition 2012).

The PAL is said to assess visual spatial STM. Validity of this test was found in schoolchildren (Luciana and Nelson 1998, 2002).

The procedure is as follows (Fig. 3): White boxes are displayed in a circle on the screen, and one or more boxes contain a unique (hidden) pattern. A trial starts with one of the boxes opening and revealing a complex colored pattern. Then, the box closes, and a next box opens to reveal its pattern. Once all boxes are closed again, a target pattern appears in the middle of the screen. The participant needs to touch the box where the target pattern was located (and previously shown) in the trial. Then, another target pattern appears, and the participant needs to locate this one and so forth. The trial ends when all patterns have been successfully located in one attempt. In case of an error, the whole trial (and thus all boxes) is registered as incorrect, and it is shown again with the target patterns; the participant needs to try the same trial from the beginning, which may be repeated up till a maximum of 10 times. As in the SSP, a trial is only scored as correct if all boxes are identified correctly. The test is terminated if the 10th attempt is also incorrect. The test continues to the next level as soon as a correct trial is completed, and a new black screen appears with the number of white boxes for the particular level. The number of boxes opening—and thus, the difficulty level—increases over trials to a maximum of 8 boxes (i.e., the range consists of 1–8 boxes). The outcome measure of the PAL is the number of patterns successfully located in the first attempt across difficulty levels. The maximum PAL score is 32 correctly identified patterns. The total administration time is, on average, 11 min (Cambridge Cognition 2012).

Statistical Analyses

Means and standard deviations were calculated for the PRM, the SSP, and the PAL (see Table 2). Then, Pearson correlations among these PRM, SSP, and PAL outcome measures were computed (Table 3).

The regression analyses included the variables age, sex, and LPE. The categorical variables sex and LPE were dummy coded (0 female and 1 male; LPE low = 0 and LPE high = 1). Also, all two-way interactions were included (e.g., age × LPE as an independent variable) as predictors that contribute to explaining the variance in the PRM, SSP, and/or PAL scores (Van der Elst et al. 2011). Age was centered ($Age_C = \text{calendar age in months} - \text{mean age of the sample of } 115.58 \text{ months}$) to avoid multicollinearity due to inclusion of power terms (Van der Elst et al. 2005). Quadratic and cubic age were included in the analyses to ensure the best possible fit of the model (e.g., curvilinear) as opposed to a linear relation, i.e., allowing for a smaller increase in scores at certain ages and a larger increase in scores during other age periods (Shaw et al. 2006). Multiple linear regressions were fitted to the data, applying a step-down hierarchical method. Predictors with a

Table 2 Means and standard deviations of the Ukrainian raw scores for the PRM, the SSP, and the PAL amount correct

	Mean	Standard deviation	<i>n</i>	Minimum of the sample	Maximum of the sample	Confidence interval	Maximum possible test score
PRM	19.09	3.45	186	8	24	18.59–19.59	24
SSP	7.94	7.50	183	0	38	6.85–9.03	44
PAL	9.61	4.80	186	1	26	8.91–10.30	32

significant *p* value of $< .05$ were included. The observed scores and expected scores (based on the regression models) were plotted and visually evaluated. More complex models allow for a better fitting model, e.g., in case of multiple curves in the data. To effectively model the potentially more complex curvilinear relation between the demographic variables and the test score at hand, fractional polynomials were fitted. Fractional polynomials can be used when initial visual inspection of the data indicates a nonlinear relationship, i.e., the effect of a variable; e.g., age may be greater at a lower age compared to the effect in the middle age group children and change direction once more in older children. Adjusting for the possibility of multiple curves in the data resulted in choosing fractional polynomial models because the parameters for each curve can be more reliably estimated compared to multiple linear regressions (Van der Elst et al. 2016). For a better fit, a restricted set of 7 exponents was used $\{-2, -1, -0.5, 0, 0.5, 1, 2\}$ with exponent 0 as the natural log (Van der Elst et al. 2016). Each of these values was used for the combinations of powers, while this restricted set has proven to reliably result in an optimally fitting regression model and prevents having to consider too many (multiples of hundreds) fractional polynomials (Van der Elst et al. 2016). The best model was selected among a proposed set of several models in the analysis based on the Akaike information criterion (AIC, i.e., lower indicates a better fit). Finally, the fit of the fractional polynomial models was again graphically evaluated against the observed scores, i.e., by adding the expected scores to the plot.

The assumptions of regressions were tested for each model. Regression analysis is based on the assumption of a normal distribution of the residuals (Field 2009). Visual inspection to assess this assumption of normality was carried out using P

Table 3 Pearson zero-order correlations for the PRM, the SSP, and the PAL correct

	PRM	SSP	PAL
PRM	1		
SSP	0.38**	1	
PAL	0.22**	0.30**	1

n = 186

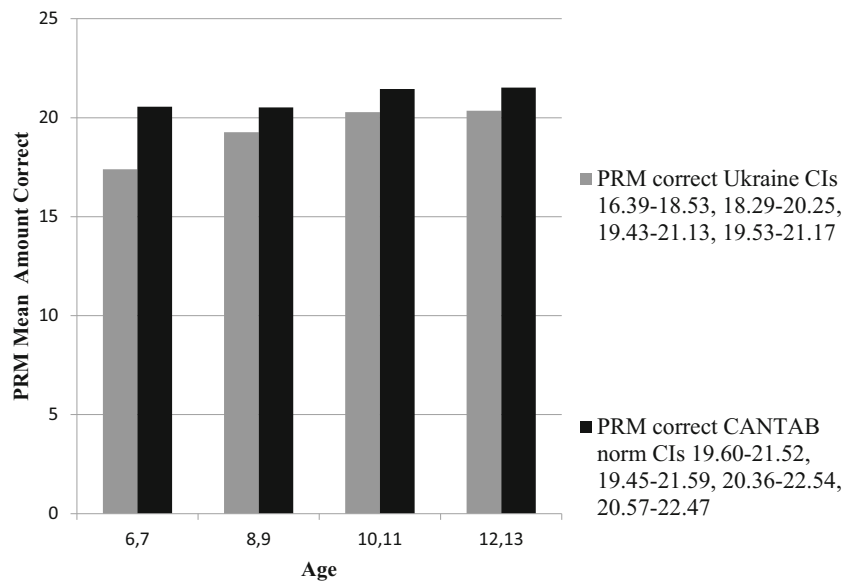
**Correlation is significant at the 0.01 level (2-tailed)

plots of the residuals. Next, the normal distribution was tested using the Kolmogorov-Smirnov test on the residuals. Then, the predicted values were divided in quartiles. These quartile groups were applied to the standardized residuals, and Levene's test was used to assess homoscedasticity (Van der Elst et al. 2011). Multicollinearity was evaluated using variance inflation factors ($VIFs \leq 10$), while Cook's distance and standardized residuals $> |3|$ of the mean were calculated to identify influential cases (Fisher et al. 2014).

Finally, normative data based on the regression (fractional polynomial) models were calculated using the following procedure (Van Breukelen and Vlaeyen 2005; Van der Elst et al. 2016). The expected scores were calculated using the final models. Then, the residuals were computed (= observed score – expected score). Next, the residuals were standardized by calculating residual / standard deviation (SD) (residual) with $SD(\text{residual}) = \text{the standard deviation of the residuals in the normative sample}$ (Van der Elst et al. 2011). Finally, the residuals were converted into percentile values using the normal distribution (with mean = 0 and $SD = SD(\text{residual})$) if the assumption for normality of the standardized residuals was met in the normative sample (otherwise, the empirical distribution of the standardized residuals of the test score in this sample was used). The Appendix shows these calculated scores. The alpha level of .05 was used to avoid type 1 errors. Analyses were carried out in SPSS (version 24) and R.

Last, the Ukrainian sample was compared to the western norm to establish how the groups relate to each other. Individual scores and LPE were not available for the CANTAB[®], so overall means of the two samples could only be compared in a descriptive manner, i.e., by establishing the confidence intervals (CIs) (Field 2009). The mean child CANTAB[®] PRM observed score (Cambridge Cognition 2014), based on its traditionally grouped 2-year age brackets (i.e., 6–7 years old up till 13 years old; see Fig. 4), was compared to the same age group of the Ukrainian PRM sample. Only the PRM outcome measure was comparable to the CANTAB[®] norm group outcome measure. The SSP and PAL outcome measures used here were calculated to only include correct items on the first attempt in contrast to other studies and could thus not be compared to the CANTAB[®] norm group (see materials; Green et al. 2019; Luciana and Nelson 2002; Mattson et al. 2010).

Fig. 4 Observed scores for the mean PRM correct comparing the Ukrainian sample to the CANTAB® traditional mean norms (2-year age groups), including confidence intervals (CIs)



Results

The means, SDs, and CIs for the observed data of the PRM, the SSP, and the PAL are displayed in Table 2. Three children did not start the SSP, because they wished to terminate the test battery early.

The comparison of the overall PRM means (observed scores) of the Ukrainian sample to the CANTAB® norm groups included *only* the age groups which overlapped between the two populations, e.g., 6–7 years old and excluding the youngest and oldest age groups (Fig. 4). The CIs for the means of the two samples overlapped for children from 8 to 13 years old, indicating that there is no significant difference between the means for the compared age groups (see Fig. 4 for all CIs). However, the CIs of the means of the 6–7 years old, of the CANTAB® (19.60–21.52) and Ukrainian (16.39–18.53) samples, did not overlap.

Table 3 shows the correlations between these outcome measures for the Ukrainian sample. The PRM, the SSP, and the PAL outcome measures correlated significantly among each other; a higher accuracy on PRM correlated to a higher accuracy on the PAL for example. The correlations ranged from 0.22 to 0.38, which is not high (i.e., .90 or higher). This indicates a significant relation between all three scores, but at the same time, the PRM, the SSP, and the PAL appear to measure different components as well. This difference may be supported by the overlap in CIs of the SSP and PAL, but the CIs of both tests do not overlap with those for the PRM (Table 2), measuring accuracy on the first attempt (although test differences need to be kept in mind).

Subsequently, the final regression (fractional polynomial) models ($p \leq .03$) are shown in Table 4. Box Cox transformations were applied because initial analyses indicated heteroscedasticity in the observed scores of the PRM and the

SSP (Osborne 2010). Additionally, the PAL and SSP displayed 3 and 2 outliers, respectively (residual standard deviation > |3|), that were removed. After this, assumptions for regressions were mostly met for the final models, i.e., all values of Levene’s statistic $p > .05$ except for SSP (Bartlett $p = .51$). Bartlett’s test of sphericity is based on a likelihood ratio to make an estimate of the distribution (Filed, 2009). This test was more suitable for the SSP due to the distribution of the residuals showing a natural log and cubic age effect (Table 4).

Cook’s distance values $\leq .01$; all variance inflation factors ≤ 6.9 (which is below the VIF cutoff value of 10 that indicates multicollinearity). Kolmogorov-Smirnov values were $p \geq .01$, except for the SSP, $p < .001$. The assumption of normal distribution as assessed in the Kolmogorov-Smirnov test was not met for the SSP, despite visual inspection of residual plots showing reasonably normal distributions. Yet, to meet statistical assumptions, norms for the SSP were based on the distribution of the standardized residuals of the SSP test scores of the Ukrainian sample (i.e., the empirical distribution), as opposed to the theoretical standard normal distribution of residuals (Vander Elst et al. 2012). Sex did not influence test performance, which also shows when comparing the mean observed scores for females and males, respectively, per task, i.e., PRM, $n = 96$ (mean = 18.86) and $n = 90$ (mean = 19.33); SSP, $n = 88$ (mean = 7.943) and $n = 95$ (mean = 7.937); PAL, $n = 90$ (mean = 10.11) and $n = 96$ (mean = 9.14). Also, none of the interactions (e.g., Age_C \times LPE) did reach significance.

For all outcome measures, there was a curvilinear effect (see Fig. 5, based on the final norm calculations and, if applicable, reverted Box Cox transformations; Osborne 2010). The PAL and SSP also showed a cubic age effect, indicating less-steep development with age in the youngest age group. The highest levels for the PRM were achieved from 10.5 years old; for the SSP, from 12 years old; and for the PAL, from 12.5 years old,

Table 4 Final multiple linear regression models for the PRM, SSP, and PAL correct outcome measures resulting from the step-down hierarchical procedure

Test score	<i>n</i>	Variable	<i>B</i>	<i>p</i>	<i>t</i>	Standard deviation (residual)	<i>R</i> ²
PRM (Box Cox transformation, $\lambda = 2.00$)	186	(Constant)	72.27	0.00	9.20	34.13	.16*
		Age_C	0.41	0.00	4.68		
		Age_C ²	-0.01	0.00	-3.08		
		LPE	17.11	0.03	2.21		
SSP (Box Cox transformation, $\lambda = 0.85$)	181**	(Constant)	14.69	0.00	1.78	5.88	.30*
		Age_C ³	0.001	0.00	2.65		
		Age_C ³ (LNAge)	-0.0001	0.00	-2.70		
		Age ²	-0.03	0.00	-2.29		
PAL	176**	(Constant)	8.25	0.00	9.66	3.66	.30*
		Age_C	0.13	0.00	5.12		
		Age_C ²	-0.0004	0.4	-0.93		
		Age_C ³	-0.00004	0.02	-2.37		
		LPE	1.92	0.02	2.29		

Encoding of the predictors: Age_C = calendar age in months – mean calendar age in 115.58 months

LPE level of parental education

***n* is lower due to removal of outliers (residuals > 3 standard deviations); **p* ≤ .03

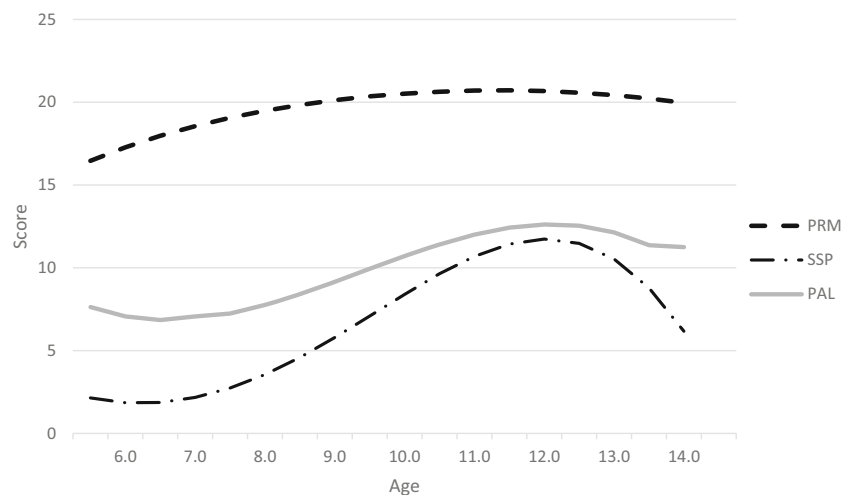
respectively. Accuracy started to level off from 14 years old in the PRM and from 13 years old in the SSP and the PAL.

Last, the more detailed reporting of the SSP and the PAL (correct items on the first attempt only) allowed some comparison between the PRM, the SSP, and the PAL (i.e., visual, spatial, and visual spatial memory). The maximum score on the PRM of 24 items was achieved by 4.8% of the Ukrainian schoolchildren. The maximum possible score was not achieved on the SSP and PAL; 0.6% scored 38 (out of 44) on the SSP, and 1.7% of this sample scored 28 (out of 32) on the PAL. The CIs of the mean observed scores (Table 2) for the PRM (18.59–19.59) do not overlap with the SSP and the PAL (6.85–9.03 and 8.91–10.30, respectively), confirming differences in score ranges.

Normative Procedure

Norms for the PRM, the SSP, and the PAL correct outcome measures (Appendix Tables A1–A3) are established using the procedure described above. For example, assume that a 6-year-old child, whose parents have a high LPE, scored 6 points correct on the PAL. First, in this normative procedure, the expected score for this child is calculated using the regression model presented in Table 4. This is constant + $[0.13 \times (\text{age child in months} - \text{average age sample})] + (-0.0004 \times \text{Age_C}^2) + (-0.00004 \times \text{Age_C}^3) + (1.92 \times \text{LPE})$, that is $8.25 + 0.13 \times (72 - 115.72) + (-0.0004 \times 1911.44) + (-0.00004 \times -83,568.09) + (1.92 \times 1) = 7.068$. The residual is calculated, which is -1.068 (score 6–7.068). Next, the

Fig. 5 Results of the regression, i.e., expected scores (Table 4) by age for the PRM (LPE high), SSP, and PAL (LPE high)



residual is standardized, -0.29 ($-1.068/3.66$). The standardized residual is converted into a percentile value based on the normal distribution. A standardized residual of -0.29 corresponds to a value of the 45th percentile. This means that 45% of the population of 6-year-old children, whose parents have a high LPE, obtains a PAL score that is equal or lower than 6. The PAL score of this child is, therefore, in the lower part of normal limits (5th–95th percentile). The SSP did not meet the normality assumption as evaluated in the Kolmogorov-Smirnov test. Norms were therefore based on the distribution of the standardized regression residuals (Vander Elst et al. 2012). The respective norm tables are in Appendix Tables A1–A3.

Discussion

The aim of the present study was to evaluate the effects of demographic variables on three tests measuring respectively visual, spatial, and visual spatial short-term memory (STM) in school-aged Ukrainian children. Also, normative data were established for this population by applying a promising method called continuous norming on these tests. This different approach to norming included the demographic variables age, sex, and level of parental education (LPE) and resulted in more adequate estimates of norm scores for reasons mentioned in the “Introduction” (Van Breukelen and Vlaeyen 2005; Van der Elst et al. 2011).

First, we found that all outcome measures on visual (PRM), spatial (SSP), and visual spatial (PAL) STM tasks correlated significantly, ranging from .22 to .38. The correlation indicates measurement of the same construct to some extent, yet it equally implies unique constructs, as the correlations are not high (i.e., $\geq .90$). Support for the correlation between visual, spatial, and visual spatial STM compared to, e.g., verbal STM was found by Smith et al. (1996). This model of domain specificity for visual (spatial) information was supported by Luck and Vogel (2013). Similarly, Alloway et al. (2006) found in children from 4 to 11 years old that brief memory processes rely on a central executive thought to be involved in domain general activities such as allocation of attention, while domain-specific visual STM (VSTM) appears as a system that represents and maintains visual, spatial, and visual spatial information briefly (Baddeley et al. 2011; Luria et al. 2009). Simultaneously, the far from perfect correlations find support in research, distinguishing the brief storage of visual information in a visual and a spatial network, i.e., the ventral pathway (involved in processing of shape and colors of objects) and the dorsal pathway (implicated in processing spatial information), respectively (Courtney et al. 1996; Logie and Marchetti 1991; Quinn 1994; Rauscher et al. 1997). Moreover, the SSP, requiring participants to remember a specific sequence of locations of squares, is thought to be a different task (associated with the

dorsal pathway), i.e., a visual spatial-sequential process (Mammarella et al. 2008).

Second, we found changes on test performance with age on all outcome measures, showing, in principle, increasing accuracy scores with maturation, albeit a curvilinear development (i.e., we found quadratic age effects on all outcome measures, Fig. 5). Luciana and Nelson (1998, 2002) found age-related development on the PRM, the SSP, and the PAL in two samples of American children from 4 to 14+ years old. Most increases in accuracy appeared to be achieved in younger children (up until 8 years old; Luciana and Nelson 1998, 2002). Curvilinearity, found in this study, also showed most increases in PRM scores with age in younger children and continued until about 11 years old. The overall mean scores of the western PRM CANTAB[®] norms and the Ukrainian sample showed overlapping CIs for 8–13-year-old children, which indicates a nonsignificant difference in the mean PRM scores between the two samples. However, the pattern of age-related increases in SSP and PAL in this sample equally showed differences compared to some literature (De Luca et al. 2003; Luciana and Nelson 1998, 2002); i.e., we observed cubic age effects (stabilizing scores) (Fig. 5). These may relate to domain-specific processes. Johnson et al. (2008) reviewed the evidence from infancy (i.e., younger children compared to, e.g., Luciana and Nelson (2002)) and concluded that refined development in the dorsal stream (spatial information processing) occurs later compared to the ventral stream, which, in turn, also implicates the development of more complex connections between ventral and dorsal streams (Kravitz et al. 2013). The SSP is thought to test spatial-sequential VSTM, i.e., a particular sequence of locations, which is associated with processing the location of objects in the dorsal pathway and the arbitrary and possibly dynamic relations between those processed objects (Kravitz et al. 2013; Mammarella et al. 2008). This applies to the PAL task as well; i.e., remembering the location of an *increasing* amount of target patterns per trial and related response items and the covering of both target and response stimuli imply a sequential process (Mammarella et al. 2008). Additionally, the PAL involves remembering matching pairs of target stimuli (VSTM for patterns) *and* their respective locations, thus implicating ventral and dorsal processing (Johnson et al. 2008; Kravitz et al. 2013). Furthermore, stabilizing scores in our study in younger children may implicate general executive functions (EFs) which, according to Alloway et al. (2006), affect especially younger children in visual spatial tasks (Anderson 2002). An environmental factor contributing to explaining the above-suggested lower mean PRM score of younger children in Ukraine compared to the western sample and literature may be related to differences in school systems; compulsory education starts from 5 years old in, e.g., England and 6 years old in Ukraine (Ukraine Channel 2017). Last, development with age for spatial VSTM is proposed to continue into young

adulthood (Farrell Pagulayan et al. 2006; Luciana et al. 2005; Luciana and Nelson 2002), indicating a need for research including 15–21 years old.

The third finding concerns the demographic variable LPE, which is associated with socioeconomic status (SES) (Davis-Kean 2005). A lower LPE (and thereby SES) background has been linked to decreased cognitive performances in children (Ardila et al. 2005; Bradley and Corwyn 2002; Koball and Jiang 2018). Higher LPE in this study is a significant positive predictor for PRM and PAL scores. This means that higher LPE is, in principle, associated with higher correct scores on tasks measuring visual and visual spatial STM (i.e., both tasks contain the detailed object features associated with parvocellular ventral processing). There were no effects on spatial STM (i.e., the SSP scores), which may support theory on distinguishing the ventral from the dorsal pathway (Kravitz et al. 2013). This is important because LPE is an environmental variable, and refined findings such as these may help to identify specific areas in stages of development (i.e., the parvocellular object feature processing) that could benefit from intervention. On the other hand, finding no LPE effects on the SSP may contribute to research as well. Spatial VSTM is theorized to be associated with mathematics development in school, which, in turn, has been the subject of inconclusive findings (Alloway and Passolunghi 2011; Mix and Cheng 2012; Raghubar et al. 2010). Equally, learning is thought to become more fine-tuned after infancy and children learn to recognize which features are important to attend to, and allocate attention (i.e., employ executive functions) based on these selections (e.g., top-down processes; De Graaf 2016) to support more complex pattern recognition tasks such as the PAL. It could be conceived that parents encouraging their children in these processes (e.g., playing memory games together) influence development. Researching the effects of LPE on (spatial) VSTM is, thus, relevant not only for current norm scores but also for identifying variables that may be linked to advantageous developmental pattern recognition. Last, the criterion for the high-level LPE group (from grade 10 and higher) may appear relatively low, yet it might have wider applicability. The National Center for Education Statistics (2014), for example, estimated that 7% of the US population did not obtain a high school diploma by the age of 24 years old.

The fourth finding is the absence of sex differences on these outcome measures. This is in line with several studies, which did not find sex differences on tasks measuring spatial VSTM (Farrell Pagulayan et al. 2006; Kessels et al. 2000; Luciana and Nelson 1998). Sex differences on spatial VSTM continue to be debated since Voyer et al. (2017) found in their meta-analysis an overall significant but small effect in visual spatial skills in favor of males (independent of age). They argued, however, that differences found might be

explained by moderator variables such as the task medium (i.e., computerized tests or remembering a pair of matched locations versus a sequence of locations), affecting the size and direction of effects in task subcategories in their review analysis of visual spatial development.

There are some limitations. The first concerns the sample size in this study. Indeed the sample size is relatively small, yet a range of studies used a similar sample size (e.g., the study of Luciana and Nelson (2002) consisted of one school and a self-selected sample with, on average, $n = 65$ in each age subgroup; De Luca et al. (2003), $n = 194$ total, divided into age subgroups). The power of a test is defined as the probability that the null hypothesis of a significance test (like, e.g., t tests to test the null H_0 : beta $k = 0$ in regression models) is rejected when the alternative is true (i.e., beta k does not equal to zero). In our models, significant effects of age and LPE were observed. This implies that the null hypotheses related to age and LPE were rejected and that there was, therefore by definition, sufficient power to conduct these tests. Furthermore, there were no sex differences, which is in line with several earlier studies (measuring visual, spatial, and visual spatial STM (Farrell Pagulayan et al. 2006; Kessels et al. 2000; Luciana and Nelson 1998)) and thus supporting the evidence in this study.

The second limitation concerns sample characteristics. The high-level LPE group was larger in numbers than the lower-level LPE group. However, this is observed in more studies of norm data (Cherner et al. 2007; Norman et al. 2000). Regression-based norms are thought to be a more accurate approximation of norms because the regression equations include the data for all demographic groups (as opposed to being based on a particular subgroup, e.g., lower LPE) (Van Breukelen and Vlaeyen 2005). Not using subgroups but all scores in one regression analysis means statistically significant results can be achieved with fewer data. Furthermore, the sample consisted of public schools in rural areas. Towns fulfill a central role in Ukraine in that children travel to the largest one in an area, so schools are not necessarily smaller compared to urban schools. However, research including urban areas will increase generalizability as, e.g., Barona et al. (1984) found effects for urban-rural regions in children and adults with head trauma on the WAIS-R scores.

The third limitation concerns comparison to western norms, which was only carried out for the PRM (see “Results” and Fig. 4). Calculating for the SSP and PAL scores, the number correct on the first attempt enhanced comparing different forms of VSTM (e.g., not including the additional and varying numbers of attempts allowed and administered in this study per task). This may limit opportunities for comparison to other research or applicability of the norms, yet it may also show how the described procedure of norming could be applied to scores on any of the outcome measures available in the three tasks. Further research is needed (beyond the descriptive measures (CIs) used) to establish (non-) significance of differences

between the western and our sample. However, the lack of overlap between the youngest age groups of the two samples on the PRM might implicate early education. Compulsory schooling starts from 6 years old (Ukraine Channel 2017). There are kindergarten facilities for 5-year-old children in Ukraine, but availability, schooling of educators, and children's attendance do not fall under educational law.

Effects of schooling on cognitive development have been reported. Small but significant differences were found in cognition (albeit different tests) when comparing children from Flanders (Belgium) and the Netherlands for instance. These neighboring countries share a similar language. Resing and Bleichroth (2015, p. 151) found a difference on subtests of the RAKIT intelligence test (6–12.5 years old) between the two nationalities, which, among others, concerned processing and recognizing patterns. Differences in findings in the WPPSI intelligence test in a sample from this mixed population supported this claim; 4–7-year-old Dutch children who habitually attend school from 4 years old had higher scores than Flemish children of the same age (Hurks and Hendriksen 2015, p. 121). There were no differences in performance on the same tests in children younger than 4 years old. Results were explained for one by differences in school systems; kindergarten in Belgium (despite starting from 2.5 years old) focuses more on social and general school skills, whereas the focus in the Netherlands is more on cognitive skills (Schittekatte and Hurks 2015, p. 171). Neighboring countries which may be perceived as similar (given the shared language) thus show differences in cognitive development at specific ages. The PRM showed lower scores in the youngest Ukrainian age group compared to the western norm. If *similar* countries (Belgian Flanders and Holland) as discussed show differences in cognitive development in early years, research might benefit from a focus on differences in early years of education including dissimilar populations such as western samples versus Ukraine. Additionally, this was a cross-sectional study. Longitudinal design would lend greater credence to future research as this may reveal the developmental trajectory of test scores in one sample (the same children) over time. Including younger children up until young adulthood is therefore recommended.

Final Conclusion

Demographically corrected norms for three computerized VSTM tasks were established for a sample of Ukrainian schoolchildren. The regression-based analyses differ from traditional approaches to norming in allowing the inclusion of multiple demographic variables simultaneously. Age and level of parental education influenced VSTM

task scores, while the demographic variable sex did not influence test performance in this sample. The regression-based normative method included all data points in the sample simultaneously and thus results in more suitable and individualized norms. Comparison of the VSTM task for patterns only, between the Ukrainian sample and the western norm, indicated no significant differences except for the youngest age group, which may point toward differences in early education. Demographically corrected norms increase the relevance of the interpretation of test scores for clinical settings.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: are they separable? *Child Development*, *77*(6), 1698–1716.
- Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133–137.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82.
- Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*(1), 143–150.
- Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, *15*(4), 185.
- Ardila, A., Rosselli, M., Matute, E., & Guajardo, S. (2005). The influence of the parents' educational level on the development of executive functions. *Developmental Neuropsychology*, *28*(1), 539–560.
- Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: the role of the episodic buffer. *Neuropsychologia*, *49*(6), 1393–1400. <https://doi.org/10.1016/j.neuropsychologia.2010.12.042>.
- Barona, A., Reynolds, C. R., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS—R. *Journal of Consulting and Clinical Psychology*, *52*(5), 885.
- Bechger, T., Hemker, B., & Maris, G. (2009). < cito_pok_memorandum_2009-01-3.pdf >. Retrieved from the provided pdf link on 10-05-2016.
- Bigelow, J., & Poremba, A. (2012). Comparing short-term memory among sensory modalities. *Visual Cognition*, *20*(9), 1012–1016.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*(1), 371–399.
- Cambridge Cognition. (2012). *Neuropsychological Test Automated Battery (CANTABeclipse) manual*. Cambridge: Cambridge Cognition Ltd.
- Cambridge Cognition. (2014). *Neuropsychological Test Automated Battery (CANTAB®) pattern recognition memory, spatial span, & paired associates learning child norms*. Cambridge: Cambridge Cognition Ltd.

- Cherner, M., Suarez, P., Lazzaretto, D., Fortuny, L. A., Mindt, M. R., Dawes, S., ... & HNRC group. (2007). Demographically corrected norms for the brief visuospatial memory test-revised and Hopkins verbal learning test-revised in monolingual Spanish speakers from the US–Mexico border region. *Archives of Clinical Neuropsychology*, 22(3), 343–353.
- Courtney, S. M., Ungerleider, L. G., Keil, K., & Haxby, J. V. (1996). Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex*, 6(1), 39–49.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037–2078.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19(2), 294.
- De Graaf, T. (2016). *Attention: phenomena, forms, and theories of attention*. Maastricht: Maastricht University.
- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the Cantab. I: development of executive function over the lifespan. *Journal of Clinical & Experimental Neuropsychology*, 25(2), 242.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
- Farrell Pagulayan, 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>.
- Field, A. (2009). *Discovering statistics using SPSS*. Thousand Oaks: Sage.
- Fisher, S. D., Gray, J. P., Black, M. J., Davies, J. R., Bednark, J. G., Redgrave, P., ... Reynolds, J. N. (2014). A behavioral task for investigating action discovery, selection and switching: comparison between types of reinforcer. *Frontiers in Behavioral Neuroscience*, 8, 398. doi:<https://doi.org/10.3389/fnbeh.2014.00398>.
- Fuster, J. M. (2002). Physiology of executive functions: the perception-action cycle. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 96–108). New York: Oxford University Press.
- Green, R., Till, C., Al-Hakeem, H., Cribbie, R., Téllez-Rojo, M. M., Osorio, E., et al. (2019). Assessment of neuropsychological performance in Mexico City youth using the Cambridge Neuropsychological Test Automated Battery (CANTAB). *Journal of Clinical and Experimental Neuropsychology*, 41(3), 246–256.
- Gruhn, W., & Rauscher, F. (2002). The neurobiology of music cognition and learning. The new handbook of research on music teaching and learning, 445–460. <http://www.posgrado.unam.mx/musica/lecturas/educacion/Neurobiologia%20of%20learning.pdf>
- Hackman, D. A., Betancourt, L. M., Gallop, R., Romer, D., Brodsky, N. L., Hurt, H., & Farah, M. J. (2014). Mapping the trajectory of socioeconomic disparity in working memory: parental and neighborhood factors. *Child Development*, 85(4), 1433–1445.
- Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Chapter 5—demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In *Clinical interpretation of the WAIS-III and WMS-III* (pp. 181–210). San Diego: Academic.
- Hitch, G. J., Towse, J. N., & Hutton, U. (2001). What limits children's working memory span? Theoretical accounts and applications for scholastic development. *Journal of Experimental Psychology: General*, 130(2), 184.
- Hurks, P., & Hendriksen, J. (2015). WPPSI-III-NL Intelligentiemeting bij kinderen en adolescenten. In W. C. M. Resing (Ed.), *Handboek intelligentietheorie en testgebruik* (pp. 113–140). Amsterdam: Pearson.
- Johnson, M. H., Mareschal, D., & Csibra, G. (2008). The development and integration of the dorsal and ventral visual pathways in object processing. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 467–478). Cambridge: MIT.
- Kaplan, G. A., Turrell, G., Lynch, J. W., Everson, S. A., Helkala, E.-L., & Salonen, J. T. (2001). Childhood socioeconomic position and cognitive function in adulthood. *International Journal of Epidemiology*, 30(2), 256–263.
- Kessels, R. P., van Zandvoort, M. J., Postma, A., Kappelle, L. J., & de Haan, E. H. (2000). The Corsi Block-Tapping task: standardization and normative data. *Applied Neuropsychology*, 7(4), 252–258. https://doi.org/10.1207/s15324826an0704_8.
- Koball, H., & Jiang, Y. (2018). *Basic facts about low-income children: children under 18 years, 2016*. National Center for Children in Poverty, Columbia University Mailman School of Public Health Retrieved from http://www.nccp.org/publications/pub_1194.htm. Accessed Jan 2018.
- Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, 17(1), 26–49.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: evidence from children. *British Journal of Developmental Psychology*, 21(1), 59–80.
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment*. New York: Oxford University Press.
- Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: visual, spatial or central executive. *Mental Images in Human Cognition*, 80, 105Å.
- Lowe, C., & Rabbitt, P. (1998). Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. *Neuropsychologia*, 36(9), 915–923.
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, 76(3), 697–712.
- Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four-to eight-year-old children. *Neuropsychologia*, 36(3), 273–293.
- Luciana, M., & Nelson, C. (2002). Assessment of neuropsychological function through use of the Cambridge Neuropsychological Testing Automated Battery: performance in 4-to 12-year-old children. *Developmental Neuropsychology*, 22(3), 595–624.
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, 17(8), 391–400.
- Luria, R., Sessa, P., Gotler, A., Jolicœur, P., & Dell'Acqua, R. (2009). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience*, 22(3), 496–512. <https://doi.org/10.1162/jocn.2009.21214>.
- Mammarella, I. C., Pazzaglia, F., & Cornoldi, C. (2008). Evidence for different components in children's visuospatial working memory. *British Journal of Developmental Psychology*, 26(3), 337–355.
- Mattson, S. N., Roesch, S. C., Fagerlund, A., Autti-Ramo, I., Jones, K. L., May, P. A., ... Collaborative Initiative on Fetal Alcohol Spectrum, D. (2010). Toward a neurobehavioral profile of fetal alcohol spectrum disorders. *Alcohol Clinical and Experimental Research*, 34(9), 1640–1650. doi:<https://doi.org/10.1111/j.1530-0277.2010.01250.x>.
- Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, 27(3), 272–277.
- Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: developmental and educational implications. *Advances in child development and behavior*, 42, 197–243 JAI.

- National Center for Education Statistics (2014). *Digest of education statistics*. Retrieved from https://nces.ed.gov/programs/digest/d14/tables/dt14_219.70.asp. Accessed Jul 2019.
- Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., et al. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, *18*(5), 773.
- Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*(4), 464–480. <https://doi.org/10.1111/j.1467-7687.2007.00600.x>.
- Norman, M. A., Evans, J. D., Miller, W. S., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. *Journal of Clinical and Experimental Neuropsychology*, *22*(1), 80–94.
- Norman, M. A., Moore, D. J., Taylor, M., Franklin Jr, D., Cysique, L., Ake, C., ... & Hnrc Group. (2011). Demographically corrected norms for African Americans and Caucasians on the Hopkins verbal learning test–revised, brief visuospatial memory test–revised, stroop color and word test, and Wisconsin card sorting test 64–card version. *Journal of Clinical and Experimental Neuropsychology*, *33*(7), 793–804.
- Osborne, J. W. (2010). Improving your data transformations: applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation*, *15*(12), 1–9.
- Quinn, J. G. (1994). Towards a clarification of spatial processing. *The Quarterly Journal of Experimental Psychology*, *47*(2), 465–480.
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: a review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122.
- Rauscher, F. H., Shaw, G., Levine, L., Wright, E., Dennis, W., & Newcomb, R. (1997). Music training causes long-term enhancement of preschool children's spatial-temporal reasoning. *Neurological Research*, *19*(1), 2–8.
- Resing, W. C. M., & Bleichroth, N. (2015). RAKIT-2 Intelligentiemeting bij kinderen in de basisschool leeftijd. In W. C. M. Resing (Ed.), *Handboek intelligentietheorie en testgebruik* (pp. 141–168). Amsterdam: Pearson.
- Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia*, *5*(5), 266–281.
- Sahakian, B., & Owen, A. (1992). Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *Journal of the Royal Society of Medicine*, *85*(7), 399.
- Schittekatte, M., & Hurks, P. (2015). WISC-III Intelligentiemeting bij kinderen en adolescenten. In W. C. M. Resing (Ed.), *Handboek intelligentietheorie en testgebruik* (pp. 169–198). Amsterdam: Pearson.
- Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., et al. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, *440*(7084), 676–679.
- Smith, E. E., Jonides, J., & Koeppel, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral Cortex*, *6*(1), 11–20.
- Subtelny, O. (2009). *Ukraine: a history*. Toronto: University of Toronto Press.
- Teixeira, R. A. A., Zachi, E. C., Roque, D. T., Taub, A., & Ventura, D. F. (2011). Memory span measured by the spatial span tests of the Cambridge Neuropsychological Test Automated Battery in a group of Brazilian children and adolescents. *Dementia & Neuropsychologia*, *5*(2), 129–134.
- Toornstra, A., Hurks, P. P. M., Van der Elst, W., Kok, G., & Curfs, L. M. G. (2019). Measuring visual matching and short-term recognition memory with the CANTAB® delayed matching to sample task in schoolchildren: effects of demographic influences, multiple outcome measures and regression based normative data. <https://doi.org/10.1080/09297049.2019.1642316>.
- Ukraine Channel (2017). Retrieved from <http://www.ukraine.com/education/>. Accessed Nov 2018.
- Van Breukelen, G. J. P., & Vlaeyen, J. W. S. (2005). Norming clinical questionnaires with multiple regression: the pain cognition list. *Psychological Assessment*, *17*(3), 336–344. <https://doi.org/10.1037/1040-3590.17.3.336>.
- Van der Elst, W., Hurks, P., Wassenberg, R., Meijis, C., & Jolles, J. (2011). Animal verbal fluency and design fluency in school-aged children: effects of age, sex, and mean level of parental education, and regression-based normative data. *Journal of Clinical and Experimental Neuropsychology*, *33*(9), 1005–1015. <https://doi.org/10.1080/13803395.2011.589509>.
- Van der Elst, W., Molenberghs, G., Hilgers, R. D., Verbeke, G., & Heussen, N. (2016). Estimating the reliability of repeated measured endpoints based on linear mixed-effects models. A tutorial. *Pharmaceutical Statistics*, *15*(6), 486–493.
- Van der Elst, W., Ouweland, C., van der Werf, G., Kuyper, H., Lee, N., & Jolles, J. (2012). The Amsterdam Executive Function Inventory (AEFI): psychometric properties and demographically corrected normative data for adolescents aged between 15 and 18 years. *Journal of clinical and experimental neuropsychology*, *34*(2), 160–171.
- Van der Elst, W., van Boxtel, M. P., van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, *11*(3), 290–302. <https://doi.org/10.1017/S1355617705050344>.
- Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *24*(2), 307–334. <https://doi.org/10.3758/s13423-016-1085-7>.
- World Medical Association Declaration of Helsinki. (2013). Ethical principles for medical research involving human subjects. *JAMA*, *310*, 2191–2194.
- Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: implications for the WAIS-R. *Journal of Clinical Psychology*, *41*(1), 86–94.
- Zasenko, O.E., Stebelsky, I., Hajda, L.A., Kryzhaniysky, S.A., Makuch, A., & Yerofeyev, I.A. (2017). Ukraine. *Encyclopaedia Britannica*. Retrieved from <https://www.britannica.com/place/Ukraine>. Accessed Feb 2018.