University of Tartu

Faculty of Social Sciences

Institute of Psychology

Triin Kesküla

MEASUREMENT INVARIANCE AND COGNITIVE PERFORMANCE OF THE CAMBRIDGE NEUROPSYCHOLOGICAL TEST AUTOMATED BATTERY (CANTAB) IN PATIENTS WITH EARLY AND LATER PHASE OF THE SCHIZOPHRENIA SPECTRUM DISORDERS COMPARED TO CONTROL SUBJECTS

Master´s Thesis

Supervisors: Kätlin Anni (MSc), Liina Haring (MD, PhD)

Tartu 2021

Measurement invariance and cognitive performance of the Cambridge Neuropsychological Test Automated Battery (CANTAB) in patients with early and later phase of the schizophrenia spectrum disorders compared to control subjects

Abstract

The aim of this study was to use selected CANTAB subtests to investigate the structure of cognitive abilities among the early phase of the schizophrenia spectrum disorders (E-SSD), the later phase of the schizophrenia spectrum disorders (L-SSD), and the entire clinical sample compared to the control subjects. A total of 97 E-SSD patients, 102 L-SSD patients and 188 control subjects were administered six CANTAB subtests of cognitive functioning. To use the most appropriate level of analyses, we first identified the factor structure of selected CANTAB subtests. Multi-group CFA (MG-CFA) was used to test the measurement invariance of factors between groups. General linear models (GLMs) were used to evaluate the nature and severity of cognitive dysfunction among the patients compared to the controls. CFA models fitted almost all data well, except for the young controls. In MG-CFAs the subtest scores corresponded to different latent trait levels indicating that the comparison of mean latent scores between most of the subsamples is meaningless. At subtest score level the patients exhibited extensive cognitive dysfunction when compared with the control subjects ($p < .001$). Current study demonstrates that the patients exhibit widespread cognitive dysfunction and that the structure of underlying cognitive abilities is not the same for the clinical sample compared to the control sample as well as for the E-SSD compared to the younger control subjects.

Keywords: CANTAB, cognitive function, early phase of the schizophrenia spectrum disorders, later phase of the schizophrenia spectrum disorders, measurement invariance

Neuropsühholoogilise testi CANTAB mõõtmise invariantsus ja kognitiivse sooritusprofiili võrdlus varajases ja kroonilises skisofreeniaspektri häire faasis olevate patsientide ning kontrollisikute vahel

Kokkuvõte

Käesoleva töö eesmärk oli CANTAB allteste kasutades hinnata kognitiivsete võimete struktuuri varajases skisofreeniaspektri häire faasis (E-SSD) olevatel patsientidel, kroonilises skisofreeniaspektri häire faasis (L-SSD) olevatel patsientidel ja kogu kliinilisel valimil võrreldes kontrollisikutega. Üheksakümne seitsme E-SSD patsiendi, 102 L-SSD patsiendi ja 188 kontrollisiku kognitiivset funktsiooni hinnati kuut CANTAB alltesti kasutades. Eesmärgiga kasutada kõige sobivamat statistiliste analüüside taset selgitati esmalt välja CANTAB alltestide faktorstruktuur. Mõõtmise invariantsust (MI) testiti mitme-grupi kinnitava faktoranalüüsiga (MG-CFA). Patsientide kognitiivse düsfunktsiooni ulatuse hindamiseks võrreldes kontrollisikutega kasutati üldisi lineaarseid mudeleid (GLM). Kinnitava faktoranalüüsi (CFA) mudelid sobitusid enamusele alagruppidele, välja arvatud noortele kontrollisikutele. MG-CFAs vastasid alltestide skoorid erinevatele latentsetele tunnustele viidates, et keskmiste skooride võrdlemine enamik alagruppide vahel on ebamõistlik. Võrreldes kontrollisikutega omasid patsiendid alltestide skoorides ulatuslikku kognitiivset düsfunktsiooni ($p < .001$). Töö tulemused osutavad, et patsiendid omavad ulatuslikku kognitiivset düsfunktsiooni ja et kliinilise valimi ja kontrollisikute ning E-SSD ja noorte kontrollisikute kognitiivsete võimete aluseks olev struktuur ei ole samasugune.

Märksõnad: CANTAB, kognitiivsete võimete struktuur, skisofreeniaspektri häire, mõõtmise invariantsus

Introduction

Cognitive dysfunction has been widely observed in schizophrenia spectrum disorders (SSD). It has been characterised as one of the key features in the SSD population. It is shown that cognitive dysfunction is exhibited in the early stages of the illness and is present even before the first psychotic episode (Sheffield et al., 2018). Cognitive dysfunction in the SSD tends to be relatively persistent during and after the treatment (Mucci et al., 2017) and it also tends to be more independent of symptomatic state than the dysfunction seen in other neuropsychiatric disorders (Keefe & Fenton, 2007). Schizophrenia is believed to be a disorder of abnormal neurodevelopment, which might be responsible for the cognitive dysfunction in the SSD (Tripathi et al., 2018). The changes in the structure and in the function of the brain neuronal networks in schizophrenia can be explained as the cumulative effect of change in neuroplasticity, neurodevelopmental abnormality, and alteration in neuronal maturation (Gourion et al., 2004).

In general, the SSD patients perform 1.5 to 2.0 standard deviation below control subjects on a variety of cognitive ability assessment tests (Nuechterlein et al., 2004). Comparison of the cognitive performance of an early phase of the schizophrenia spectrum disorders (E-SSD) and the later phase of the schizophrenia spectrum disorders (L-SSD) allows to assess whether cognitive dysfunction appears through a degenerating process following an E-SSD or is dysfunction already present when the psychosis emerges (Sheffield et al., 2018). Meta-analysis (Mesholam-Gately et al., 2009) across 47 studies of over 2000 first-episode patients has shown medium-to-large dysfunction across 10 cognitive domains. Mesholam-Gately et al. compared these results with the L-SSD patients, and it was found that the impairment pattern was quite similar to the dysfunction found in the E-SSD patients. The effect size for the E-SSD patients ranged from -.64 to -1.20 and for the L-SSD patients the effect size ranged from -.46 to -1.41. Therefore, data indicates that broad cognitive dysfunction is relatively stable during the course of the SSD (Sheffield et al., 2018; Heilbronner et al., 2016). In contrast, there are studies which have reported that the cognitive dysfunction may become more severe in the L-SSD (Wu et al., 2016) and that the magnitude across cognitive dysfunctions varies in the E-SSD patients compared to the L-SSD patients (Braw et al., 2008). Braw et al. (2008) found that chronic SSD patients are significantly more impaired with dysfunction related to pattern memory, psychomotor speed, and executive functions. Therefore, more research on the question about the evolution of cognitive dysfunction over the course of the SSD is needed.

The most common cognitive deficits in the SSD are associated with attention, working memory, memory, problem solving, and processing speed. Factor analytic studies have shown that there might be a broad mechanism that represents the cognitive dysfunction seen in the SSD, which appears in one-factor solution (Barch & Ceaser, 2012; Dickinson et al., 2006). However, there are also studies that have found up to seven separable dimensions representing the cognitive dysfunction in the SSD (McCleery et al., 2015; Nuechterlein et al., 2004). The reason for that could be in the number of cognitive functioning assessment tests included and in the variety of conceptualizations of the latent structure of cognition proposed in the SSD (McCleery et al., 2015). Therefore, the question about the underlying mechanism of the cognitive dysfunction in the SSD population remains a matter of interest.

There are studies that have investigated the differences in cognitive functioning between the E-SSD patients and the L-SSD patients (e.g. Sheffield et al., 2018; McCleery, et al., 2014), but to the best of our knowledge, there are no studies that compare the structure of cognitive abilities among the E-SSD patients and the L-SSD patients using the Cambridge Neuropsychological Test Automated Battery (CANTAB; Robbins & Sahakian, 1994; Sahakian & Owen, 1992), especially on the Estonian SSD population. Haring et al. (2015) presented an idea that the cognitive differences between clinical sample and control subjects may not be only related to general levels of cognition but there may also be structural differences. Thus, it is important to investigate if the cognitive dysfunction change in the course of the illness and whether there may be structural differences underneath it.

The aim of the current study was to investigate the structure of cognitive abilities among patients with the E-SSD, the L-SSD and the entire clinical sample compared to the age-matched control subjects. Some of the E-SSD patients and matched younger control subjects (Y-CS) participating in this study have previously been included to our research group prior measurement invariance (MI) study (Haring et al., 2015). We extended present study by including the L-SSD patients and their age-matched older control subjects (O-CS) to investigate the cognitive performance of the SSD patients in the different stages of the illness in order to gain better understanding of the nature of the dysfunction seen in the SSD. Our study was based on the CANTAB, which has been widely used over the last decades to assess neurocognitive dysfunction in the SSD (Levaux et al., 2007; Pantelis et al., 1997). We selected six CANTAB tests (paired associates learning, PAL; spatial span, SSP; spatial working memory, SWM; Stockings of Cambridge, SOC; intra/extra-dimensional shift, IED; and rapid visual information processing, RVP) in order to reflect a broad spectrum of cognitive function among the sample.

In cognitive functioning studies, the age of the participant is an important factor because of the normal age-related cognitive change. Normal aging is associated with impairments in the cognitive processes (Morrison & Baxter, 2012). Therefore, it is necessary not only to assess the cognitive functioning among the whole clinical sample but also to assess the cognitive functioning separately in the early phase of the illness and the later phase of the illness. Besides investigating the participant scores on cognitive functioning assessment measures, it is important to examine the internal structure of the cognitive processes. In other words, if we intend to understand the cognitive dysfunction seen in the SSD, it is necessary to investigate the nature of the dysfunction. One way of meaningfully compare the results of measures between different groups is establishing the MI (Wicherts, 2016). MI allows to examine whether the results of patients from different stages of illness as well as in different age groups, can be interpreted in a similar way (Lee, 2018; Bialosiewicz et al., 2013). It is important to understand the basic principles of MI testing to produce more comprehensive and applicable results in research and in clinical practice (Lee, 2018). To test MI across the subjects from various subsamples, a multigroup confirmatory factor analysis (MG-CFA) should be conducted (Milfont & Fischer, 2010). MG-CFA is an extension of the typical CFA. In MG-CFA, the researchers divide the data set into groups, determine the model fit for each group separately and conduct the multi-group comparisons (Bialosiewicz et al., 2013). Establishing the MI in the E-SSD patients compared to the Y-CS and in the L-SSD patients compared to the O-CS is important because it helps to assess whether the cognition may qualitatively differ between the selected subgroups. The MI testing is typically conducted in three phases: the configural invariance testing, the metric invariance testing, and the scalar invariance testing. Thus, the MI testing is relevant addition to the statistical analyses helping to increase the validity of the research (Lee, 2018).

In line with previous literature, we hypothesised that MI would not be supported between the clinical sample and the control sample. In addition, we investigated whether the CANTAB subtest scores correspond to different latent trait levels in the E-SSD subgroup compared to the Y-CS and in the L-SSD subgroup compared to the O-CS. We also hypothesised that patients would have significantly poorer cognitive performance in the selected CANTAB tests compared to the control subjects and that the structure of cognitive dysfunction in an E-SSD subgroup might be quite similar compared with the structure of cognitive dysfunction in an L-SSD subgroup.

The author of the present study contributed by sample gathering, organising, and analysing the data, doing systematic literature search and synthesis of the previous publications, and writing the thesis.

Method

Participants

The overall sample of the study was 344 individuals from whom 45.3% were included in the clinical sample (N = 156) and 54.7% were included in the control sample (N = 188). The clinical sample was recruited from the Psychiatry Clinic of Tartu University Hospital, Estonia and consisted of SSD in-patients or outpatients (53.2% males). Ninety-seven patients from the clinical sample (aged between 18 and 45 years) were recruited with the E-SSD and 102 patients (aged between 18 and 65 years) were recruited with the L-SSD. In the L-SSD subgroup 43 patients were included from the E-SSD subgroup because their illness had been monitored over 5 years (therefore we could use longitudinal data for the analyses). The mean clinical sample age was 33.1 years (SD = 11.7, range 18–65 years) and 92.3% were right-handed. In the clinical sample, the mean years of education was 13.2 (SD = 2.7, range 8–21 years). The E-SSD subgroup fulfilled the following inclusion criteria: aged between 18 and 45 years; admitted for the first time to hospital for psychotic disorder (fulfilment of ICD-10 diagnoses F20 – F29). In the E-SSD subgroup the diagnoses were F23 (N = 68), F20.09 (N = 21), F20.39 (N = 2), F25 (N = 2), F20.29 (N = 1), F21 (N = 1), F22 (N = 1), and F28 (N = 1). The L-SSD subgroup fulfilled the following inclusion criteria: aged between 18 and 65 years; experience of a chronic phase of the SSD. In the L-SSD subgroup the diagnoses were F20 (N = 79), F25 (N = 20), F22 (N = 2), and F29 (N = 1). Diagnoses were based on a clinical interview according to the tenth revision of the International Classification of Diseases, ICD-10 (World Health Organization, 1992) criteria and medical chart review. When recruited, all patients were on antipsychotic treatment. The patients were in a stable phase and were capable of taking part of the study.

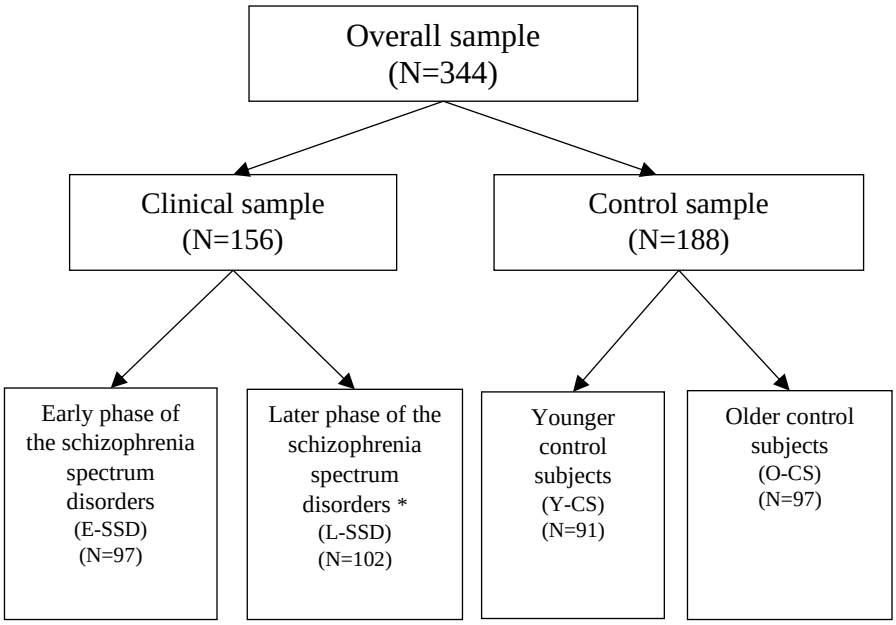
A sample of control subjects (N = 188, age range 18–69 years, 41.5% males) was recruited by advertisement from a similar geographical area. The mean age of the control subjects was 29.1 years (SD = 11.1) and 95.7% were right-handed. In the control sample, the mean years of education was 14.1 (SD = 2.3, range 9–19 years). Due to the mean age difference between the E-SSD and the L-SSD group members, we decided to divide the control sample into the Y-CS, N = 91 and the O-CS, N = 97. The mean age of the Y-CS was

21.5 years (SD = 1.8, range 18–25 years) and 94.5% were right-handed. The mean age of the O-CS was 36.2 years (SD = 11.4 range 23–69 years) and 96.7% were right-handed. Exclusion criteria for the control subjects included the presence of mental disorder and the presence of psychotic disorder among close relatives.

Exclusion criteria included problematic drug use, neurological disorders, presence of organic brain disease, mental retardation, significant learning disorder, major sight and hearing impairment or physical conditions severe enough to interfere with cognitive performance. All participants were required to speak the Estonian language. All participants gave written informed consent to participate in the study and did not receive compensation. Ethical approvals (no. 252/M-31; no. 176/T4; no. 177/T-2; no. 211/M-22; no. 96/16) were granted by the Ethics Review Committee on Human Research, University of Tartu, Estonia.

Figure 1

Sample Scheme



Note. *43 patients were included from the E-SSD subgroup. Patients cognitive functioning had been assessed twice. First assessment took place when patients were in the E-SSD and second assessment took place 5.3 years later, when patients were in the L-SSD.

Measures and procedures

For the current study, we chose six tests from CANTAB. The CANTAB is computerized neuropsychological assessment measure, which subtests have demonstrated great responsiveness in detecting changes in neuropsychological functioning and are known to correlate to neural networks (Cambridge Cognition, 2021). From the attention and psychomotor speed tasks we used Rapid Visual Information Processing A´(RVP). The RVP measures sustained attention. A prime assesses how good is the subject at detecting target sequences. From memory tasks we used Paired Associates Learning First Attempt Memory Score (PAL) and Spatial Span Forward Span Length (SSP). The PAL assesses visual memory and new learning. It measures how many times the subject choses the right box on their first attempt while recalling the pattern locations. The SSP assesses visuospatial working memory capability. For PAL and SSP, higher score indicated better performance. From executive functions, we used Intra-Extra Dimensional Set Shift EDS Errors (IED), Spatial Working Memory Between Errors (SWM) and Stockings of Cambridge Problems Solved in Minimum Moves Total (All moves) (SOC). The IED measures attentional set formation maintenance, visual discrimination, and attention flexibility with the quantity of trials for which the subject´s outcome was an incorrect response. The SWM measures subject´s strategy and working memory errors. It measures the quantity of times the subject incorrectly revisits a box from a token has previously been found. The outcome is calculated across four, six and eight token trials. For IED and SWM, higher score indicated worse performance. The SOC measures the subject´s ability to use problem-solving strategies when matching two sets of stimuli. The outcome is the quantity of successfully completed problems in the minimum possible number of moves. It is calculated over all assessed trials. For SOC, higher score indicated better subtest outcome. For further information, see the CANTAB® website (http://www.cambridgecognition.com).

The tests we chose are sensitive to detect participant performance profile peculiarity in the SSD (Barnett et al., 2010; Robbins, 2005). All used tasks are computerized and were run on a personal computer with a high-resolution touchscreen. The CANTAB tests use non-verbal stimuli. Instructions were given in Estonian from a translation of the CANTAB test manual. The task administration took about one hour. If participants needed, they were offered short breaks during the testing.

Statistical analyses

First, we examined all groups in terms of patients and controls demographic characteristics using an independent samples *t*-test for continuous variables and unpaired two-sample Wilcoxon test for not normally distributed data. We used Pearson's chi-squared test to analyse group differences by gender. Second, an analysis of the covariance structure of the measured CANTAB subtests was performed using principal components analyses (PCAs). We used parallel analysis to determine the most appropriate number of components. Because the components were expected to be correlated, each PCA was followed by an oblique (oblimin) rotation. Third, to see if PCA results were plausible reflections of latent cognitive constructs, confirmatory factor analysis (CFA) was conducted. PCAs and CFAs were done separately in the clinical sample, the control sample, the E-SSD subgroup, the Y-CS, the L-SSD subgroup and the O-CS. Fourth, multi-group CFA (MG-CFA) was used to test the MI for assessing whether the mean scores of each group could be meaningfully compared. For establishing MI, we tested a set of increasingly constrained structural equation models and tested the differences between them (Putnick & Bornstein, 2016; van de Schoot et al., 2012). The following models were tested: (1) the configural invariance model, with loadings and intercepts free to vary across groups, but the same factorial pattern was specified for each group; (2) the metric (weak) invariance model, with loadings constrained to be equal across the groups; (3) the scalar (strong factorial) invariance model, with factor loadings and intercepts constrained to be equal across the groups; (4) the residual (strict) invariance model, with factor loadings, intercepts and residual variances constrained to be equal across the groups. Differences in model fit were estimated using the χ^2 difference test (Horn & McArdle, 1992), where a statistically significant ($p < .05$) $\Delta\chi^2$ indicated a difference in fit. As the χ^2 difference test is found to be too sensitive to small deviations (Chen, 2007), we additionally estimated the difference between CFIs (ΔCFI). We used a cutoff of -.002 for ΔCFI (more constrained model minus less constrained model), meaning the change of smaller than or equal to -.002 indicates that the null hypothesis of invariance should not be rejected (Meade et al., 2008).

General linear models (GLMs) were used to investigate group differences in subtest scores between groups. One comparison was made with GLM repeated measures. Subtest scores were standardized. In order to standardize the subtest scores, we calculated the mean scores and standard deviations for the control sample, the Y-CS, the O-CS and for the 43 patients from the E-SSD and then determined the *z* scores for the clinical sample, the E-SSD, the L-SSD and for the 43 patients from the L-SSD. In the GLMs we used reversed tests scores

for IED and SWM because they measure the number of incorrect responses and therefore higher score indicates worse performance. Age, gender and years in education were used as covariates in comparisons of cognitive functioning. We used Cohen's *d* to describe the standardized mean difference of an effect. Regarding to Cohen (1988) we interpreted the effect sizes as small ($d = .2$), medium ($d = .5$), and large ($d = .8$). We elucidated effect size as clinically meaningful at $d = .3$. We interpreted differences as statistically significant at $p < .05$ and not significant at $p > .05$ (Pearson, 1900). Statistical analyses were conducted using the R Statistical software (R Development Core Team, 2020).

Results

For the means and standard deviations of participant age and years of education see Table 1. In the control sample, age was not normally distributed. Wilcoxon test showed that there was a significant difference ($U = 11056$, $p < .001$, $d = -.43$) between age in the clinical sample compared to the control sample. The median age for the clinical sample was 29.6 and for the control sample it was 25.1. There was a significant difference in the years of education between the clinical sample and the control sample, $t(342) = 3.32$, $p = .001$, $d = .36$. The number of participants by gender from the clinical sample compared to the control sample was significantly different, χ^2 (1, N = 344) = 4.28, $p < .05$. For the E-SSD subgroup compared to the Y-CS, there was a significant difference between age in the E-SSD subgroup and the Y-CS, $t(186) = -8.37$, $p < .001$, $d = 1.19$. There was no significant difference in the years of education between the E-SSD subgroup and the Y-CS, $t(186) = 1.23$, $p = .22$, $d = .18$. The number of participants from the E-SSD subgroup compared to the Y-CS did not differ by gender, χ^2 (1, N = 188) = 2.64, $p > .05$. For the L-SSD subgroup compared to the O-CS, there was no significant difference between age in the L-SSD subgroup and the O-CS, $t(197) = -1.1$, $p = .27$, $d = .16$. There was a significant difference in the years of education between the L-SSD subgroup and the O-CS, $t(197) = 3.36$, $p < .001$, $d = .51$. The number of participants by gender from the L-SSD subgroup compared to the O-CS was significantly different, χ^2 (1, N = 199) = 4.24, $p < .05$. Due to significant differences between the groups demographic data, age and years of education were treated as covariates in subsequent analyses.

Table 1

Means and Standard Deviations (SD) of Participant Age and Years of Education

	Overall sample		Clinical sample		Control sample	
	Clinical sample	Control sample	E-SSD	L-SSD	Y-CS	O-CS
Variable	Mean ± SD		Mean ± SD		Mean ± SD	
Age	33.1 ± 11.69	29.1 ± 11.06	27.4 ± 6.77	38.0 ± 11.21	21.5 ± 1.75	36.2 ± 11.36
Years of education	13.16 ± 2.69	14.06 ± 2.30	13.04 ± 2.55	13.27 ± 2.83	13.43 ± 1.78	14.66 ± 2.57

Note. E-SSD, early phase of the schizophrenia spectrum disorders; Y-CS, younger control subjects; L-SSD, later phase of the schizophrenia spectrum disorders; O-CS, older control subjects.

PCA

The scree plot test and parallel analysis suggested using one component solution for all groups. In all groups, six CANTAB subtest variables (PAL, IED, SOC, SSP, SWM, RVP) defined a component representing general cognitive ability as the subtests tended to cluster into one general cognitive functioning domain for all groups. The factor loadings for the clinical sample were 0.44 – 0.81 and the factor loadings for the control sample were 0.37 – 0.69. The solution accounted for 45% of the total variance among the indicators in the patients and 36% of the total variance among the indicators for the controls. The factor loadings for the E-SSD subgroup were 0.26 – 0.81 and factor loadings for the Y-CS were 0.18 – 0.70. The solution accounted for 39% of the total variance among the indicators in the E-SSD subgroup and 31% of the total variance among the indicators for the Y-CS. The factor loadings for the L-SSD subgroup were 0.59 – 0.82 and the factor loadings for the O-CS were 0.37 – 0.68. The solution accounted for 53% of the total variance among the indicators in the L-SSD subgroup and 38% of the total variance among the indicators for the O-CS. The mean item complexity for all the groups and models was 1.

CFA

One-component model derived from the PCA results was converted to a latent trait model for single-group CFAs. We executed the CFA for the clinical sample compared to the

control sample, the E-SSD subgroup compared to the Y-CS, the L-SSD subgroup compared to the O-CS as well as separate subgroups of controls and patients. All factor loadings were significant at least $p < .05$ in each of the analyses. The one-factor model did not fit well for the Y-CS (CFI = .886; RMSEA = .064), which indicates that this solution may not be the most appropriate for this group. However, as the fit indices of the model were good for all other subgroups (Table 2), we decided to continue with MG-CFA.

Table 2

Goodness-of-Fit Statistics for the Confirmatory Factor Analysis

	χ^2 (df)	CFI	RMSEA estimate (90% CI)
Clinical sample vs control sample (N=344)	12.603 (9)	0.993	0.035 (0.000-0.075)
Clinical sample (N=156)	6.482 (9)	1.000	0.000 (0.000-0.072)
Control sample (N=188)	6.115 (9)	1.000	0.000 (0.000-0.061)
E-SSD vs Y-CS (N=188)	11.343 (9)	0.992	0.038 (0.000-0.096)
E-SSD (N=97)	3.434 (9)	1.000	0.000 (0.000-0.018)
Y-CS (N=91)	12.349 (9)	0.886	0.064 (0.000-0.144)
L-SSD vs O-CS (N=199)	6.177 (9)	1.000	0.000 (0.000-0.060)
L-SSD (N=102)	4.646 (9)	1.000	0.000 (0.000-0.060)
O-CS (N=97)	5.772 (9)	1.000	0.000 (0.000-0.080)

Note. df, Degrees of freedom; CFI, comparative fit index; RMSEA, root mean square error of approximation; CI, confidence intervall; E-SSD, early phase of the schizophrenia spectrum disorders; Y-CS, younger control subjects; L-SSD, later phase of the schizophrenia spectrum disorders; O-CS, older control subjects.

MI

The results of the MI testing are shown in Table 3. Comparing the clinical sample and the control sample, the fit of the configural invariance model was good ($\chi2 = 12.598$, df = 18; RMSEA = 0.000; CFI = 1.000). This justified the evaluation of more restrictive models. The metric invariance was supported, as the equality of subtest loadings did not result in a significant degradation of model fit ($p = .130$; $\Delta CFI < .001$). The scalar invariance was not

supported because of the significant degradation of the model fit (p = .002; ΔCFI = .051). As scalar MI was not met, we discontinued the further invariance testing.

In the E-SSD subgroup compared to the Y-CS, the configural invariance was met, as CFI and RMSEA indicated a good fit [$\chi 2$ = 15.783, df = 18; RMSEA = .000; CFI = 1.000]. The metric invariance was also tenable (p = .472; ΔCFI < .001). However, the scalar invariance was not supported, as constraining intercepts equal did yield a significant degradation of the model fit (p = .015; ΔCFI = .072). As scalar MI was not met, testing for stricter forms of MI was not justified.

In the L-SSD subgroup compared to the O-CS, the fit of the configural invariance model was good [$\chi 2$ = 10.418, df = 18; RMSEA = 0.000; CFI = 1.000]. The metric invariance was also supported (p = .280; ΔCFI < .001), indicating that factor loadings were similar across the L-SSD subgroup and the O-CS. The scalar invariance was also tenable as the equality of the intercepts did not result in a significant degradation of the model fit (p = .682; ΔCFI < .001). We continued and tested the strict invariance as well, although this was not met, as constraining the residual invariances resulted in significant degradation of the model fit (p = .01, ΔCFI = .014).

The scalar noninvariance indicated that CANTAB subtest scores corresponded to different latent trait levels, hence the comparison of mean latent scores between the clinical sample and the control sample and between the E-SSD subgroup and the Y-CS is meaningless.

Table 3

Goodness-of-Fit Indices for Testing Measurement Invariance between groups with Multi-Group Confirmatory Factor Analysis

Invariance model	χ^2 *(df)*	$\Delta\chi^2$ *(df)*	*p*	RMSEA	CFI	ΔCFI
Clinical sample vs control sample						
1. Configural	12.598 (18)			.000	1.000	
2. Metric	21.102 (23)	8.504 (5)	.130	.000	1.000	<.001
3. Scalar	40.399 (28)	19.297 (5)	.002	.051	.949	.051
E-SSD vs Y-CS						
1. Configural	15.783 (18)			.000	1.000	
2. Metric	20.344 (23)	4.561 (5)	.472	.000	1.000	<.001
3. Scalar	34.465 (28)	14.121 (5)	.015	.050	.982	.072
L-SSD vs O-CS						
1. Configural	10.418 (18)			.000	1.000	
2. Metric	16.694 (23)	6.276 (5)	.280	.000	1.000	<.001
3. Scalar	19.805 (28)	3.111 (5)	.682	.000	1.000	<.001
4. Strict	36.565 (34)	16.760 (6)	.01	.028	.986	.014

Note. The metric models were compared to the configural models; the scalar models were compared to the metric models; the strict model was compared to the scalar model. CFI = comparative fit index; df = degrees of freedom; RMSEA = root mean square error of approximation; E-SSD, early phase of the schizophrenia spectrum disorders; Y-CS, younger control subjects; L-SSD, later phase of the schizophrenia spectrum disorders; O-CS, older control subjects.

Comparison of cognitive performance

We tested differences in cognitive performance using GLM based on observed test scores in the clinical sample compared to the control sample, the E-SSD subgroup compared to the Y-CS and the L-SSD subgroup compared to the O-CS. Due to differences between the groups within the factor structure, we used CANTAB subtest scores to assess the participant performance. For each CANTAB subtest score, age, gender, and years in education were

included as covariates in between-group comparisons (for the results, see Table 4, Table 5, and Table 6). For the comparison of cognitive performance of 43 patients within two assessments over the time period of 5.3 years, see Table 7. In general, the patients exhibited extensive cognitive dysfunction in assessed subtests when compared with the control subjects (see Figure 2). Age and years in education had a meaningful impact as covariates.

Table 4

General Linear Models. The Clinical Sample (N = 156) Compared to the Control Sample (N = 188)

	PAL	IED	SOC	SSP	SWM	RVP
Parameter	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
Clinical sample vs control sample	-0.89 (-1.11 to -0.67) ***	-1.15 (0.88-1.43) ***	-1.31 (-1.56 to -1.06) ***	-0.77 (-0.97 to -0.57) ***	-1.24 (0.98-1.50) ***	-0.89 (-1.12 to -0.66) ***
Education in years	0.08 (0.03-0.12)***	-0.10 (-0.15 to -0.04)***	0.08 (0.03-0.13)**	0.06 (0.02-0.10)**	-0.08 (-0.13 to -0.03)**	0.11 (0.06-0.15)***
Age	-0.05 (-0.06 to -0.04)***	0.03 (0.02-0.04)***	-0.01 (-0.02 to -0.00)*	-0.03 (-0.04 to -0.02)***	0.03 (0.02-0.04)***	-0.02 (-0.03 to -0.01)***
Gender	-0.03 (-0.24 to 0.18)	0.24 (-0.03 to 0.51)	-0.16 (-0.41 to 0.08)	-0.28 (-0.48 to -0.08)**	0.41 (0.16-0.67)**	-0.17 (-0.40 to 0.05)

Note. * p < .05, ** p < .01, *** p < .001; CI, confidence interval; PAL, Paired Associates Learning; IED, Intra-Extra Dimensional Shift; SOC, Stockings of Cambridge; SSP, Spatial Span; SWM, Spatial Working Memory; RVP, Rapid Visual Information Processing; estimates for IED and SWM are reversed.

Table 5

General Linear Models. The E-SSD Subgroup (N = 97) Compared to the Y-CS (N = 91)

	PAL	IED	SOC	SSP	SWM	RVP
Parameter	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
E-SSD vs Y-CS	-1.09 (-1.49 to -0.69) ***	-1.48 (1.98 to 0.98) ***	-1.71 (-2.15 to -1.27) ***	-0.80 (-1.15 to -0.45) ***	-1.54 (1.08-2.00) ***	-0.97 (-1.35 to -0.58) ***
Education in years	0.14 (0.06-0.22)***	-0.12 (-0.23 to -0.02)*	0.03 (-0.06 to 0.12)	0.08 (0.01-0.15)*	-0.07 (-0.16 to 0.02)	0.16 (0.08-0.23)***
Age	-0.06 (-0.09 to -0.02)**	0.03 (-0.2 to 0.07)	0.03 (-0.01 to 0.07)	-0.04 (-0.07 to -0.01)**	0.02 (-0.02 to 0.06)	-0.02 (-0.05 to 0.01)
Gender	0.03 (-0.30 to 0.37)	0.61 (0.19-1.04)**	-0.43 (-0.80 to -0.06)*	-0.35 (-0.64 to -0.05)*	0.43 (0.05-0.82)*	-0.28 (-0.61 to 0.05)

Note. * p < .05, ** p < .01, *** p < .001; CI, confidence interval; PAL, Paired Associates Learning; IED, Intra-Extra Dimensional Shift; SOC, Stockings of Cambridge; SSP, Spatial Span; SWM, Spatial Working Memory; RVP, Rapid Visual Information Processing; E-SSD, early phase of the schizophrenia spectrum disorders; Y-CS, younger control subjects; estimates for IED and SWM are reversed.

Table 6

General Linear Models. The L-SSD Subgroup (N = 102) Compared to the O-CS (N = 97)

	PAL	IED	SOC	SSP	SWM	RVP
Parameter	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
L-SSD vs O-CS	-0.81 (-1.11 to -0.51) ***	-0.80 (0.44-1.15) ***	-0.84 (-1.14 to -0.53) ***	-0.74 (-1.00 to -0.48) ***	-0.91 (0.58-1.23) ***	-0.69 (-0.98 to -0.39) ***
Education in years	0.06 (0.01-0.11)*	-0.10 (-0.16 to -0.02)**	0.10 (0.04-0.15)***	0.08 (0.03-0.13)**	-0.08 (-0.14 to -0.02)**	0.11 (0.05-0.16)***
Age	-0.04 (-0.06 to -0.03)***	0.03 (0.01-0.04)**	-0.02 (-0.04 to -0.01)***	-0.03 (-0.04 to -0.02)***	0.04 (0.02-0.05)***	-0.02 (-0.03 to -0.00)*
Gender	0.04 (-0.26 to 0.35)	0.03 (-0.33 to 0.39)	0.15 (-0.16 to 0.46)	-0.29 (-0.55 to -0.02)*	0.47 (0.14-0.80)**	-0.07 (-0.38 to 0.23)

Note. * p < .05, ** p < .01, *** p < .001; CI, confidence interval; PAL, Paired Associates Learning; IED, Intra-Extra Dimensional Shift; SOC, Stockings of Cambridge; SSP, Spatial Span; SWM, Spatial Working Memory; RVP, Rapid Visual Information Processing; L-SSD, later phase of the schizophrenia spectrum disorders; O-CS, older control subjects; estimates for IED and SWM are reversed.

Table 7

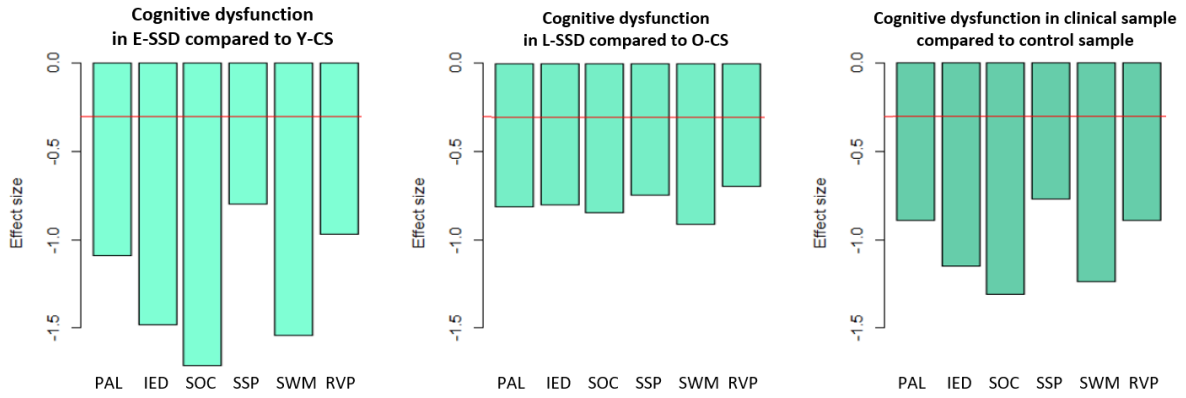
General Linear Models. Cognitive Assessment of the 43 Patients from the E-SSD Subgroup Whose Cognitive Functioning Has Been Assessed Over the Time Period of 5.3 Years

	PAL	IED	SOC	SSP	SWM	RVP
Parameter	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)	Estimate (95% CI)
E-SSD vs L-SSD	0.18 (-0.30 to 0.65)	0.27 (-0.77 to 0.24)	0.42 (-0.00 to 0.85)	0.03 (-0.43 to 0.49)	0.43 (-0.85 to -0.01)*	0.34 (-0.08 to 0.76)
Education in years	0.05 (-0.04 to 0.14)	-0.07 (-0.17 to 0.03)	0.12 (0.04-0.21) **	0.15 (0.06-0.24) **	-0.10 (-0.18 to -0.02)*	0.17 (0.09-0.25) ***
Age	-0.05 (-0.08 to -0.01)*	0.01 (-0.03 to 0.05)	0.04 (0.01-0.08)*	-0.04 (-0.08 to 0.00)	0.01 (-0.02 to 0.05)	-0.02 (-0.06 to 0.01)
Gender	0.51 (0.03-0.99) *	0.14 (-0.37 to 0.65)	-0.24 (-0.67 to 0.19)	-0.33 (-0.79 to 0.14)	0.35 (-0.07 to 0.78)	-0.12 (-0.54 to 0.31)

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. CI, confidence interval; PAL, Paired Associates Learning; IED, Intra-Extra Dimensional Shift; SOC, Stockings of Cambridge; SSP, Spatial Span; SWM, Spatial Working Memory; RVP, Rapid Visual Information Processing; E-SSD, early phase of the schizophrenia spectrum disorders; L-SSD, later phase of the schizophrenia spectrum disorders; estimates for IED and SWM are reversed.

Figure 2

Cognitive Dysfunction Profiles



Note. This figure demonstrates the cognitive performance of the E-SSD subgroup compared to the Y-CS, the L-SSD subgroup compared to the O-CS and the clinical sample compared to the control sample expressed as effect sizes. The red line on effect size at 0.3 represents clinically meaningful dysfunction. PAL, Paired Associates Learning; IED, Intra-Extra Dimensional Shift; SOC, Stockings of Cambridge; SSP, Spatial Span; SWM, Spatial Working Memory; RVP, Rapid Visual Information Processing; E-SSD, early phase of the schizophrenia spectrum disorders; Y-CS, younger control subjects; L-SSD, later phase of the schizophrenia spectrum disorders; O-CS, older control subjects; estimates for IED and SWM are reversed.

Discussion

The study aimed to assess the cognitive dysfunction among the clinical sample compared to the control sample, the E-SSD subgroup compared to the Y-CS, and the L-SSD subgroup compared to the O-CS. To use the most appropriate level of analyses, we first identified the factor structure of selected CANTAB tests (PAL, IED, SOC, SSP, SWM, and RVP) across the subgroups of the sample. CFAs support the conclusion that the single-factor model of cognitive tests were appropriate in most of the subgroups, except for the young control subjects. We used MG-CFA to test the MI and assess whether the mean latent scores of each group could be meaningfully compared. Our results indicated that the comparison of mean latent scores across most of the subsamples is meaningless. The scalar invariance, which enables to the comparison of the mean latent scores across groups were tenable only in comparison of the L-SSD subgroup and the O-CS. Next, we used CANTAB tests scores to investigate the dysfunction of cognitive abilities in the clinical sample compared to the control sample, the E-SSD subgroup compared to the Y-CS and the L-SSD subgroup compared to the O-CS. All used CANTAB tests scores indicated that patients exhibit clinically meaningful cognitive dysfunction when compared with control subjects. In this study, we extended the previous research of Haring et al. (2015) by including the L-SSD patients. Haring et al. (2015) examined the structure of cognitive dysfunction in the first episode psychosis patients

compared to the control subjects. For the current study we used partly same data as Haring et al. (2015) as we also compared the cognitive functioning of the patients in the E-SSD subgroup and Y-CS. Similar to Haring et al. (2015), we used MG-CFA to test the MI.

In SSD studies, various models for the underlying structure of cognition can be made. The models can be categorized as single-factor, correlated-factor, and hierarchical models (McCleery et al., 2015). Similar to previous CFA studies (Barch & Ceaser, 2012; Dickinson et al., 2006) single-factor model provided a good fit for our data. In this paper, we used six cognitive functioning tests, which represented three cognitive functioning domains (memory, executive functions, attention and psychomotor speed). We selected these subtests based on previous studies on the SSD population and due to fact that these tests are well known to detect cognitive changes in the SSD population (Cambridge Cognition, 2021). To compare the latent structure of a neuropsychological battery in the SSD and the controls, Dickinson et al. (2006) used CFAs including 17 tests which represented six cognitive functioning domains (verbal comprehension, perceptual organization, verbal learning and memory, visual learning and memory, information processing speed, and executive/working memory). In Dickinson et al. (2006) study, the CFAs supported the hierarchical model, where cognitive functioning tests loaded on six cognitive functioning domain factors, and those factors loaded on the general cognitive functioning factor. These results support the idea of general cognitive ability factor decidedly influencing a number of distinct domains of cognitive functioning (Dickinson et al., 2006). Although our data indicated to single-factor model for clinical sample in different stages of the SSD, some studies have supported models with up to seven factors (McCleery et al., 2015; Nuechterlein et al., 2004). There are many possible reasons for these incongruity findings. For example, the number of cognitive functioning tests included, domains assessed, and models proposed. McCleery et al. (2015) compared seven-factor model to different cognitive functioning models (a single-factor model, a correlated-factors model, and a hierarchical-model) using 16 tests representing seven cognitive functioning domains. McCleery et al. (2015) concluded, that in their study, the seven correlated-factors model was the best.

During the validation process of neurocognitive test batteries, it is common to conduct factor analysis, but if we intend to assess the adequacy of the test norms for certain subgroups, we need to evaluate if the factor model fits for the subgroups and whether the measurement parameters of the model are invariant across these subgroups. If cognitive functioning assessment test fails the invariance testing, then there are other factors besides the targeted underlying cognitive functioning that affect the performance differently across these groups.

Lack of invariance underlines that the measurement model relating to the assessed subtest scores differs across the groups (Wicherts, 2016). Therefore, it is important to establish the MI, because it can detect the possible differences in the structure of measured constructs (Haring et al., 2015). Establishing the MI is especially relevant in the clinical studies, where the validity of the tests is essential (Wicherts, 2016). Our results on MI testing support Haring et al. (2015) results on the clinical sample compared to the control sample and on the E-SSD subgroup compared to the Y-CS. Similar to Haring et al. (2015), in our study it was not possible to compare the clinical sample and the control sample in terms of their mean latent cognitive factors because the structure among the tests were different between groups. Interestingly, the mean latent scores in the L-SSD subgroup compared to the O-CS could be meaningfully compared regarding scalar invariance. Showing that the structure of the cognitive abilities in the L-SSD subgroup compared to the O-CS was more similar than the structure in the E-SSD subgroup compared to the Y-CS.

Our hypothesis, that the clinical sample would exhibit meaningful cognitive dysfunction compared with the control sample was supported and is consistent with other authors (e.g Nuechterlein et al., 2004). Our results on cognitive dysfunction seen in the E-SSD subgroup compared to the Y-CS and in the L-SSD subgroup compared to the O-CS, indicated that patients in both stages of the illness exhibit deeper dysfunction on the executive function domain than in other measured domains. These findings support Brawn et al. (2008), who also found that executive functions is one of the domains which is more impaired in the SSD population. Surprisingly, the dysfunction seen in the L-SSD subgroup compared to the O-CS was not on the same magnitude as in the E-SSD subgroup compared to the Y-CS. The reason for that could lie in the cognitive performance of the Y-CS. The recruitment of the control subjects was based on convenience sampling. Therefore, the results in the Y-CS may not be extrapolatable to their peers living in Estonia. We believe that most of the Y-CS were willing to participate in the study due to their curiosity in psychology. It is possible that the cognitive functioning of the Y-CS in our study may be better than the cognitive functioning of their peers living in the similar geographical area. Recruited E-SSD subgroup participants had a larger degree of heterogeneity in age than the Y-CS. Although the years of education was similar in the E-SSD subgroup and in the Y-CS, the participants in the E-SSD subgroup were almost 6 years older than the Y-CS. Showing, that the Y-CS were still on the rise of their educational performance compared to the participants in the E-SSD subgroup, whose peak years of education was likely achieved. The E-SSD usually emerges during early adulthood or late adolescence. It is prime time for youngsters to establish their identity and functioning as

members of the society. It is well noted that majority of individuals in the E-SSD experience disruptions on their educational journey (Shinn et al., 2020).

We also made cognitive performance comparison for the 43 patients, whose performance had been assessed within two assessments over the time period of 5.3 years. The results indicated that the comparison of the two assessments was statistically not significant for most of the subtests, except for the SWM ($p < .05$). This would suggest that the cognitive dysfunction of the participants in this subsample was quite stable over the assessed time period. This finding is in line with studies that suggest stable cognitive performance profile following one´s first psychotic episode (Sheffield et al., 2018; Heilbronner et al., 2016). Therefore, the cognitive dysfunction in the E-SSD is believed to be close to its final form. The cognitive dysfunction is also believed to be one of the core features in the SSD (Sheffield et al., 2018). Surprisingly, there was a statistically significant change in the participants´ performance of the SWM subtest as the outcome was better for the second assessment. In neuropsychological assessment, retesting is necessary to observe the course of the disorder, but it is also important to consider the role of practice effects on the participant performance on follow-up testing (Beglinger et al., 2005). The practice effect is defined as subject´s improvement in the test performance (Hausknecht et al., 2007). Thus, it may be possible that in the course of the SSD, the cognitive dysfunction may stay stable or may become more severe, but the change may be hidden under the learning effect.

One limitation of this study could be the sample size. Although it has the benefit of representing an early phase and the later phase of the SSD, a larger sample might give a broader understanding of the cognitive performance of the subjects. Second, since we used six cognitive functioning subtests from CANTAB, it is possible, that studies with more subtests may provide a different outcome. Studies including more subtests might give more universal picture of the cognitive dysfunction seen in the SSD. Third, we assessed patients who were willing to participate, were clinically stable and on antipsychotic medication. Therefore, the results may not represent the overall cognitive functioning profile of the SSD patients in Estonia.

Despite the possible limitations, we believe that the current study provides interesting results which can be transferred to psychiatric practice. The methodical strength of this study is in providing data from different subsamples representing distinct stages of the disorder and in establishing the MI. We believe that it is important to establish the MI while comparing psychometric data between clinical samples and control samples. To our knowledge, this is the first study comparing the structure of cognitive abilities among the E-SSD patients and the

L-SSD patients using CANTAB on the Estonian SSD population. With present study, we would like to emphasize the broader use of CANTAB in the cognitive assessment in the different stages of the SSD.

In conclusion, the CFA analyses provide support for the single-factor model of cognitive functioning in the SSD subgroups and the control subjects. We extended our analyses with MI which indicate that the cognitive functioning structure in the SSD and the population without the SSD could be different. Our results highlight that patients with the SSD perform significantly poorer in cognitive performance, such as executive functions, memory, attention, and psychomotor speed than individuals without the SSD.

Acknowledgements

I am very grateful to my supervisors Kätlin Anni and Dr Liina Haring for endless inspiration, encouragement, and valuable guidance during the research process. I would also like to thank all the participants of the present study for contributing their time and effort.

References

Barch, D. M., & Ceaser, A. (2012). Cognition in schizophrenia: Core psychological and neural mechanisms. *Trends in Cognitive Sciences*, *16*(1), 27-34.

Barnett, J. H., Robbins, T. W., Leeson, V. C., Sahakian, B. J., Joyce, E. M., & Blackwell, A. D. (2010). Assessing cognitive function in clinical trials of schizophrenia. *Neuroscience & Biobehavioral Reviews*, *34*(8), 1161-1177.

Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J., Fastenau, P. S., & Siemers, E. R. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*, *20*(4), 517-529.

Bialosiewicz, S., Murphy, K., & Berry, T. (2013). An introduction to measurement invariance testing: Resource packet for participants. *American Evaluation Association*, 1-37.

Braw, Y., Bloch, Y., Mendelovich, S., Ratzoni, G., Gal, G., Harari, H., Tripto, A., & Levkovitz, Y. (2008). Cognition in young schizophrenia outpatients: Comparison of first-episode with multiepisode patients. *Schizophrenia Bulletin*, *34*(3), 544-554.

Cambridge Cognition Ltd. (2021). CANTAB Cognitive Research Software. Cambridge Cognition. https://www.cambridgecognition.com/cantab

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464-504.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillsdale, NJ: Laurence Erlbaum Associates.

Dickinson, D., Ragland, J. D., Calkins, M. E., Gold, J. M., & Gur, R. C. (2006). A comparison of cognitive structure in schizophrenia patients and healthy controls using confirmatory factor analysis. *Schizophrenia Research*, *85*(1-3), 20-29.

Gourion, D., Gourevitch, R., Leprovost, J. B., & Krebs, M. O. (2004). Neurodevelopmental hypothesis in schizophrenia. *L'encephale*, *30*(2), 109-118.

Haring, L., Mõttus, R., Koch, K., Trei, M., & Maron, E. (2015). Factorial validity, measurement equivalence and cognitive performance of the Cambridge Neuropsychological Test Automated Battery (CANTAB) between patients with first-episode psychosis and healthy volunteers. *Psychological Medicine*, *45*(9), 1919.

Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., & Moriarty Gerrard, M. O. (2007). Retesting in selection: a meta-analysis of coaching and practice effects for tests of cognitive ability. *Journal of Applied Psychology*, *92*(2), 373.

Heilbronner, U., Samara, M., Leucht, S., Falkai, P., & Schulze, T. G. (2016). The longitudinal course of schizophrenia across the lifespan: Clinical, cognitive, and neurobiological aspects. *Harvard Review of Psychiatry*, *24*(2), 118.

Horn, J. L. & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117–144.

Keefe, R. S., & Fenton, W. S. (2007). How should DSM-V criteria for schizophrenia include cognitive impairment?. *Schizophrenia Bulletin*, *33*(4), 912-920.

Lee, S. T. (2018). Testing for measurement invariance: Does your measure mean the same thing for different participants?. *APS Observer*, *31*(8).

Levaux, M. N., Potvin, S., Sepehry, A. A., Sablier, J., Mendrek, A., & Stip, E. (2007). Computerized assessment of cognition in schizophrenia: Promises and pitfalls of CANTAB. *European Psychiatry, 22*(2), 104-115.

McCleery, A., Ventura, J., Kern, R. S., Subotnik, K. L., Gretchen-Doorly, D., Green, M. F., Hellemann, G. S., & Nuechterlein, K. H. (2014). Cognitive functioning in first-episode schizophrenia: MATRICS consensus cognitive battery (MCCB) profile of impairment. *Schizophrenia Research, 157*(1-3), 33-39.

McCleery, A., Green, M. F., Hellemann, G. S., Baade, L. E., Gold, J. M., Keefe, R. S., Kern, R. S., Mesholam-Gately, R. I., Seidman, L. J., Subotnik, K. L., Ventura, J., & Nuechterlein, K. H. (2015). Latent structure of cognition in schizophrenia: a confirmatory factor analysis of the MATRICS consensus cognitive battery (MCCB). *Psychological Medicine*, *45*(12), 2657.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of Applied Psychology, 93*(3), 568-92.

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*, *23*(3), 315–336.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: applications in cross-cultural research. *International Journal of Psychological Research*, *3*(1), 111-130.

Morrison, J. H., & Baxter, M. G. (2012). The aging cortical synapse: Hallmarks and implications for cognitive decline. *Nature Reviews Neuroscience*, *13*(4), 240-250.

Mucci, A., Merlotti, E., Üçok, A., Aleman, A., & Galderisi, S. (2017). Primary and persistent negative symptoms: Concepts, assessments and neurobiological bases. *Schizophrenia Research, 186*, 19-28.

Murante, T., & Cohen, C. I. (2017). Cognitive functioning in older adults with schizophrenia. *Focus*, *15*(1), 26-34.

Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., & Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research*, *72*(1), 29-39.

Pantelis, C., Barnes, T. R. E., Nelson, H. E., Tanner, S., Weatherley, L., Owen, A. M., & Robbins, T. W. (1997). Frontal–striatal cognitive deficits in patients with chronic schizophrenia. *Brain: A Journal of Neurology*, *120*(10), 1823–1843.

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157-175.

Putnick, D. L. & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review, 41*, 71-90.

R Core Team (2020). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. (https://www.R-project.org/).

Robbins TW, Sahakian BJ (1994). Computer methods of assessment of cognitive function. In Principles and Practice of Geriatric Psychiatry (ed. J. R. M. Copeland, M. T. Abou-Saleh and D. G. Blazer), pp. 205–209. John Wiley & Sons, Ltd:Chichester, UK.

Robbins, T. W. (2005). Synthesizing schizophrenia: a bottom-up, symptomatic approach. *Schizophrenia Bulletin*, *31*(4), 854-864.

Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, *85*(7), 399.

Sheffield, J. M., Karcher, N. R., & Barch, D. M. (2018). Cognitive deficits in psychotic disorders: a lifespan perspective. *Neuropsychology Review*, *28*(4), 509-533.

Shinn, A. K., Cawkwell, P. B., Bolton, K., Healy, B. C., Karmacharya, R., Yip, A. G., Öngür, D., & Pinder-Amaker, S. (2020). Return to college after a first episode of psychosis. *Schizophrenia Bulletin Open*, *1*(1), sgaa041.

Tripathi, A., Kar, S. K., & Shukla, R. (2018). Cognitive deficits in schizophrenia: Understanding the biological correlates and remediation strategies. *Clinical Psychopharmacology and Neuroscience*, *16*(1), 7.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, *9*(4), 486-492.

Wicherts, J. M. (2016). The importance of measurement invariance in neurocognitive ability testing. *The Clinical Neuropsychologist*, *30*(7), 1006-1016.

Wu, J. Q., Tan, Y. L., Xiu, M. H., De Yang, F., Soares, J. C., & Zhang, X. Y. (2016). Cognitive impairments in first-episode drug-naive and chronic medicated schizophrenia: MATRICS consensus cognitive battery in a chinese han population. *Psychiatry Research*, *238*, 196-202.

Non-exclusive licence to reproduce thesis and make thesis public

1. I, Triin Kesküla, herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, Measurement invariance and cognitive performance of the Cambridge Neuropsychological Test Automated Battery (CANTAB) in patients with early and later phase of the schizophrenia spectrum disorders compared to control subjects, supervised by Kätlin Anni and dr Liina Haring.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Triin Kesküla

31/05/2021